QUANTITATIVE TRAIT LOCI ANALYSIS TO IDENTIFY MODIFIER GENES OF THE GENE *Opaque2* IN MAIZE ENDOSPERM

A Dissertation

by

LIBARDO ANDRES GUTIERREZ ROJAS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2007

Major Subject: Molecular and Environmental Plant Sciences

QUANTITATIVE TRAIT LOCI ANALYSIS TO IDENTIFY MODIFIER GENES OF THE GENE *Opaque2* IN MAIZE ENDOSPERM

A Dissertation

by

LIBARDO ANDRES GUTIERREZ ROJAS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Monica Menz
	Alan Pepper
Committee Members,	Javier Betran
	Scott Finlayson
Chair of Molecular and Environmental Plant Sciences Faculty,	Jean Gould

December 2007

Major Subject: Molecular and Environmental Plant Sciences

ABSTRACT

Quantitative Trait Loci Analysis to Identify Modifier Genes of

the Gene *Opaque2* in Maize Endosperm. (December 2007)

Libardo Andres Gutierrez Rojas, B.S., Universidad Nacional de Colombia;

M.S., Universidad Nacional de Colombia

Co-Chairs of Advisory Committee: Dr. Monica Menz
Dr. Alan Pepper

The protein quality of maize can be improved by replacing normal *Opaque2* alleles with non-functional recessive alleles *opaque2* (*o2*). The allele *o2* produces a severe phenotype with soft endosperm enhancing its protein quality but decreasing its agronomical value. Plant breeders have restored a desirable ratio of hard to soft endosperm in *o2* germplasm known as Quality Protein Maize (QPM). Neither the mechanism nor the genetic components by which the modification of the endosperm in QPM lines occurs are well understood. To increase the understanding of the genetics of endosperm modification, a population of 146 recombinant inbred lines derived from a cross between the *o2* inbred line B73*o2* and the QPM inbred line CML161 was evaluated in two Texas locations from 2004 to 2006. Four traits related to endosperm texture were measured and showed significant effect of the inbred lines, high heritability estimates and high genetic correlations. Relative content of the essential amino acids lysine, tryptophan and methionine were measured and showed significant effects of the

lines and considerable high genetic correlations and heritabilities. Negative correlation was observed between endosperm texture traits and amino acid content. Quantitative trait loci (QTL) were mapped for traits related to the modification of endosperm texture and the content of lysine, tryptophan and methionine. QTLs clusters for endosperm texture traits were detected on chromosomes 3, 5, 6 and 8 explaining 62-68% of the variation. QTLs clusters for amino acid contents were located on chromosomes 7 and 8 that explained up to 39% of the observed variation. The product of the *O2* gene is a transcription factor that affects the expression of a number of endosperm genes. A group of 29 endosperm genes associated with the O2 activity were evaluated in developing endosperm of the recombinant inbred lines. Genomic regions controlling gene transcript abundance in developing endosperm were identified by expression QTL mapping. Evidence is presented of QTL hot spots that segregate in association with endosperm texture modification or amino acid contents and are associated with the regulation of the expression of a group of endosperm genes.

DEDICATION

This dissertation is dedicated to my parents, Maria Stella and Libardo Antonio, to my wife Cristina, and to my daughter Paula.

ACKNOWLEDGEMENTS

I hold a very special feeling of gratitude to my mentor and co-chair, Dr. Monica Menz. She showed immense confidence in me and my capabilities as a scientist. Beyond the academic guidance, she has been a counselor and a friend in all aspects of my life. I express my sincere thanks and gratitude to Dr. Javier Betran, my former co-chair and committee member. Dr. Betran has been an example of dedication, hard work, enthusiasm, and wisdom. I am grateful to Dr. Alan Pepper for being my co-chair and sharing with me his broad vision of plant molecular biology. Special thanks to Dr. Scott Finlayson for serving on my committee and for his opportune input.

Thanks to Dr. William Russell and Dr. Palmy Jesudhasan for their technical support. I would also like to thank Dr. Jean Gould and Dr. Marla Binzel who served as chairs of the Molecular and Environmental Plant Sciences program and to Dr. Marty Dickman, director of the Institute for Plant Genomics and Biotechnology. Sincere gratitude towards our collaborators outside Texas A&M, Paul Scott and Jim Coors.

The constant support of Susan Hall was invaluable for the successful completion of my studies. I must recognize the role of Mrs. Shiela Dainello who was in charge of too many details of my life as a student of Texas A&M University. Thanks to Dr. No in the Laboratory for Plant Genome Technology, and to Kathryn Rocheli who was a friend and support. Thanks to Kerry Mayfield who assisted me in many occasions. Thanks are also due to all fellow graduate students, former post docs, student workers, and visiting scientists of our lab and in the Maize breeding program: Bryan, Nilesh, Vivek, Hsiao-Hi,

Jorge, Chandra, Bryan, Ryan, Kareen, Poonam, Lacy, Cody, Otto, Halima, Rebecca, and many others. It would be difficult to name all those who collaborated during this research, but their efforts will be always greatly appreciated.

I thank Pioneer Hi-Bred International for providing me a scholarship to pursue doctoral studies. I also appreciate the logistic and financial support of the Texas Agricultural Experiment Station.

None of this would have been possible without my spouse, friend and colleague Cristina. It is impossible to describe how important Cristina was for the entire Ph.D. process, not only as the perfect companion, but as a coworker in the laboratory and in the field. Cristina also gave me Paula, my ultimate motivation. Thanks to my parents, brothers, and sister; far away but always close to my heart. Thanks to my whole Gutierrez-Cruz family. To my good old friends from high-school and college, specially to Juan and Martha. Thanks God.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

CHAPTER I

INTRODUCTION

Maize *(Zea mays* L.) is one of the most important crops in the world. In 2005, 712,334.380 tons of maize were produced in the world. The U.S produced 40% of the total, making it the largest maize producer in the world (FAOSTAT, 2007). In developed countries, maize is mainly used as feed for livestock and industrial processes to manufacture products such as plastics, fabrics, syrups, and alcohol. In many developing countries in Latin America, Africa and Asia, maize constitutes a staple food that contributes up to 15% of the protein intake and 20% of the calories (National Research Council, 1988).

Maize belong to the *Poaceae* family, the monophyletic family of the grasses with approximately 10,000 species including other cereal crops such as rice, wheat, barley, and sorghum (Clark et al., 1995). Maize has 10 chromosomes (n=10, 2n=20). The size of the maize genome (~2,400 Mb) is similar to the human genome, nearly twenty times the one of *Arabidopsis*, and six times that of rice (Timmermans et al., 2004). The genome of maize is particularly rich in repetitive DNA, accounting for up to 66% of the total genome (Haberer et al., 2005). Despite decades of intensive research by botanists, geneticists, archeologists and molecular biologists, the origin an early history of maize is still subject of debate. One of the most accepted theories proposes that maize descends

---

This dissertation follows the style of Crop Science.

from an annual species of Teosinte (*Zea mays* ssp. *parviglumis*), native to the Balsas River Valley of southern Mexico (Piperno and Flannery, 2001). Additional research suggests that maize was domesticated approximately 7,500 years ago in central or southern Mexico (Hilton and Gaut, 1998). Maize was a staple food for the Mayan and Aztecan civilizations and played also an important role in their cultural traditions and religious beliefs. Maize was also important for the pre-Columbian cultures of South America who depended heavily on maize as source of calories (Mesa, 1957).

Christopher Columbus wrote in his journal on October 16 1492, 4 days after his first landing in San Salvador (eastern Bahamas), that the people who inhabited the islands cultivated maize, which he recorded as panizo grass (Morison, 1963). Columbus transported maize seeds in his following expeditions, and the first European description of maize appeared in 1494. Maize was initially cultivated en European gardens as an ornamental plant, but soon its value as food was recognized (Purseglove, 1972). Spaniard and Portuguese explorers spread maize during their voyages. Maize was reported in China in 1573 (Purseglove, 1972). Maize is now grown around the world and a crop of maize is harvested somewhere every month of the year (Mangelsdorf and Reeves, 1939).

Maize is a very variable species. One of the earliest systems to classify maize was based on its separation in six groups: pop, flint, flour, dent, sweet and pod (Sturtenvant, 1899). Except from pod maize, the separation in these groups is based on the composition and characteristics of the endosperm of the kernel (Glover and Mertz, 1987).

The Kernel of Maize

The grain or kernel is the main harvested product of maize. Maize kernels are produced in the female inflorescence or ear. The kernel of maize is botanically classified as a caryopsis and it is composed of three main parts: embryo, endosperm, and pericarp (Watson, 2003). The pericarp is the outermost cover of the kernel and constitutes about 5-6 % of the kernel dry weight. The embryo comprises all the tissues that will give rise to the seedling structure after germination and constitutes approximately 1.1 % of the kernel weight. The endosperm is the main storage tissue of the kernel and constitutes up to 85% of the kernel weight (Watson, 2003).

Endosperm

Because of its function as an storage tissue, the endosperm of cereals is an invaluable renewable source of food for humankind (Olsen, 2004). The endosperm of angiosperms is a unique triploid tissue originating after a male gamete fertilizes the diploid central cell in a parallel process to the fertilization of the egg cell that gives origin to the diploid zygote (Nawaschin, 1898; Guignard, 1899). The most recognized role of the endosperm is the synthesis and accumulation of storage products to nurture the embryo during initial stages of germination and seedling development. Reserves in the endosperm are accumulated in the form of lipids, carbohydrates and proteins. The proportions of these components vary between species (Demason, 1997). In the case of maize, the main structural component of the endosperm is starch, accounting for about 71% of the grain (Prassana et al., 2001). The protein fraction represents 9.5% of the

kernel, while other plants like soybean or beans contain 38% and 23% of protein, respectively (National Research Council, 1988).

The Prolamins of Maize: The Zeins

According to the classification system of cereal proteins proposed by Osborne et al. (1914), there are four classes of proteins depending on their solubility: albumins (proteins soluble in pure water), globulins (protein soluble in diluted salt solutions but insoluble in pure water), prolamins (proteins soluble in 70% ethanol), and glutelins (proteins soluble in diluted acid or base). Approximately 60% of the proteins present in the endosperm of maize belong to the prolamin type (also known as zeins), while glutelins account for 34%, and albumins and globulins for 3% each. The zeins are particularly deficient in the essential amino acids lysine and tryptophan, and their prevalence in the endosperm is the primary cause of poor protein quality in maize (Vasal, 2000). Lysine is the first limiting amino acid in the diets of non-ruminants and humans, followed by tryptophan and threonine (Shimada and Cline, 1974). Increasing the levels of the nutritionally limiting amino acids is an important objective of plant breeding programs.

Based on their related structure and differential solubility, the zeins are classified in α-, β-, δ-, and γ-zeins (Coleman et al., 1997; Esen, 1987). The α-zein is the most abundant class of the zeins comprising about 70% of the total. When separated on SDS-polyacrilamide gels, the α-zeins appear as two bands of apparent molecular masses of 19- and 22 kDa. These bands correspond to a mixture of several polypeptides, because

the α-zeins are a multigene family. The γ-zeins are the second most abundant class of zeins making up about 20% of the total (Thompson and Larkins, 1989). The γ-zeins are cisteine-rich proteins that in gels show approximate molecular size of 16-, 27- and 50-kDa (Coleman et al., 1997; Woo et al., 2001). The only member of the β-zein class is a polypeptide of approximately 15 kDa, accounting for about 5% of the total zein content (Thompson and Larkins, 1989). The δ-zeins are evidenced as two bands of 10- and 18-kDa. The δ-zeins account for less than 5% of the total zeins (Thompson and Larkins, 1989).

The mRNAs of the zeins are detected in the endosperm tissue from 8-10 dap and reach their maximum expression at approximately 16-18 dap (Dolfini et al., 1992). The same pattern of gene expression is shared by several zein genes, suggesting a coordinated regulation of transcription (Dolfini et al., 1992).The zeins are synthesized by membrane-bound polyribosomes and transported into the lumen of the endoplasmic reticulum where they are packaged in protein bodies during endosperm development (Lending and Larkins, 1989). The spatial distribution of the zeins is crucial for the formation and maturation of the protein bodies (Kim et al., 2002). The γ-zeins and β-zeins seem to be important in leading the formation of protein bodies by associating with α-zeins and δ-zeins and promoting their accumulation in the lumen of the rough endoplasmic reticulum (Bagga et al., 1997; Coleman et al., 1996; Geli et al., 1994; Kim et al., 2002; Lee, 1998; Woo et al., 2001). The 50- and 27-kDa γ-zeins remain at the surface of the protein bodies, while the 15-kDa β-zein and the 16-kD γ-zein would be

important in determining the organization of α-zeins and δ-zeins in the core of the protein body (Lending and Larkins, 1989; Woo et al., 2001).

The α-zeins are encoded by a large multigene family that can be subdivided into four subfamilies: z1A, z1B, and z1D (all three 19 kDa α-zeins), and z1C (22-kDa α-zein) according to DNA sequence similarity (Song and Messing, 2002; Song et al., 2001). The inbred B73 carries 25 genes for the three 19 kDa α-zein subfamilies. The z1A and z1B subfamilies share clusters of genes in chromosomes 4S and 7S. The z1D is the smallest subfamily and is located on chromosome 1L. (Song and Messing, 2002; Woo et al., 2001). B73 carries 15 genes of the z1C subfamily that are located in chromosome 4S (Song and Messing, 2003). Variable copy number has been observed for the α-zein subfamilies in different genetic backgrounds (Song and Messing, 2003). The variability in the gene structure of the α-zeins may be explained by mechanisms of intergenic exchange, amplification, and transposition (Viotti et al., 1985). Despite the high number of α-zein genes in the maize genome, only a small percentage of them are expressed in the endosperm, while the rest are inactive or expressed at low levels (Song and Messing, 2003; Woo et al., 2001).

The gene organization of the β-, γ, and δ-zeins is simpler than the one of the α-zeins. The 15 kDa β-zein gene is located on chromosome 6S (bin 6.01) (Weerakoon et al., 1993). The gene encoding the 10kDa d-zein is located in chromosome 9L near the centromere (bin 9.03) (Benner et al., 1989), and the 18 kDa δ-zein gene mapped to the chromosome 6L (bin 6.04) (Swarup et al., 1995). The gene encoding the 16 kDa γ-zein

was mapped to chromosome 2L (bin 2.07), and the gene for the 50 kDa γ-zein to chromosome 7L (7.05) (Woo et al., 2001). The 27 kDa γ-zein has one or two gene copies, depending on the genetic background, located in tandem in chromosome 7 (bin 7.02-7.05) (Das and Messing, 1987; Geetha et al., 1991).

The Opaque2 Transcription Factor

The mutant gene *opaque2* (*o2*) was initially discovered by Jones and Singleton in the early 1920s (Emerson et al., 1935). However, it was not until the 1960s that Mertz and coworkers at the University of Purdue noticed that the presence of this recessive allele in a homozygous form caused a significant increase in the proportion of lysine and tryptophan in the grain (Mertz et al., 1964). Plant breeders started introgression of the *o2 allele* into different varieties, but they soon found that undesirable traits were associated with the presence of this gene. The kernel turned very soft and floury, the yield decreased 8-15% and the plants were more susceptible to fungi and insects (National Research Council, 1988). The floury endosperm of *o2* materials was found directly associated with the reduction in the proportion of α-zein proteins. The *O2* gene encodes a basic leucine-zipper (bZIP) class transcription factor (Schmidt et al., 1990). It has been shown that *O2* in maize endosperm induces the transcription of α-zeins and β-zeins by binding specific sequences in their promoter elements (Schmidt et al., 1992; Cord-Neto et al., 1995). Therefore, homozygous plants for a non-functional *o2* allele show a remarkable reduction in the proportion of zeins. It has been postulated that fewer zeins cause the development of smaller protein bodies (one-fifth of the normal size) that alter

the packing of the starch fraction during seed desiccation (Coleman et al., 1997). However, due to the duplication and amplification events of the 22 kDa α-zein genes in the maize genome, some genotypes carry copies of α-zein genes that do not require the O2 protein as transcriptional activator (Song et al., 2001). The O2 protein could also interacts with a prolamin box binding factor (PBF) to regulate the expression of zein genes that lack O2 binding sites (Vicente-Carbajosa et al., 1997). The O2 protein also influences the expression of other genes such as *b-32* (a type I ribosome inactivating protein), and *CyPdk1(*a cytosolic pyruvate orthophosphate dikinase) (Bass et al., 1992; Schmidt et al, 1992; Maddaloni et al., 1996).

Additional research at the International Maize and Wheat Improvement Center (CIMMYT) discovered that the soft endosperm, caused by *o2*, could be restored into normal type endosperm by the activity of some genes called *modifiers* (Bjarnason and Vasal 1992). The genotypes that were developed containing the *o2* modifiers were called 'quality protein maize' (QPM), because they retain higher levels of lysine and tryptophan when compared to normal maize materials (Vasal, 2000). Several groups have attempted to identify those genes responsible for the endosperm modification, but no single modifier gene has been fully characterized to date. Pollacsek and coworkers genetically proved the presence of two genes causing *o2* suppression (Pollacsek, 1970; Pollacsek et al., 1972). Genetic analyses on QPM material suggest that there are two independent loci controlling endosperm modification (Lopes and Larkins, 1995). Bulk segregant analysis (BSA) of segregating populations developed from *o2* and modified varieties found two loci associated with the modification of soft endosperm (Lopes et al.,

1995). Some studies suggest that the γ-zeins and their regulation are involved in the modification of *o2* endosperm (Geetha et al., 1991; Burnett and Larkins, 1999).

QPM genotypes have been introduced to different production systems in several tropical and sub-tropical countries, where their quality advantages have been acknowledged (National Research Council, 1988). Nevertheless, the actual combination of genetic features that define a QPM line is yet to be established. A detailed knowledge about the genomic location of the genes, the characterization of the modifier genes and specific alleles involved, in addition to expression profiles and regulatory aspects will provide fundamental information to continue increasing the benefits offered by maize.

CHAPTER II

PHENOTYPIC CHARACTERIZATION OF ENDOSPERM MODIFICATION AND AMINO ACID CONTENTS

Overview

Homozygous *o2* kernels have increased levels of lysine and tryptophan. Unfortunately, the associated soft texture of the *o2* kernels causes poor yield and susceptibility to diseases and insects. Breeding programs have restored the desirable ratio of hard vs. soft endosperm in *o2* genotypes developing modified endosperm versions called Quality Protein Maize (QPM). QPM germplasm, developed mostly by the International Maize and Wheat Improvement Center, is currently utilized in breeding programs worldwide and has been demonstrated to be competitive in yield trials when compared to local checks. To contribute to an increased understanding of the genetics of endosperm modification, a population of 146 recombinant inbred lines (S5 to S7) derived from a cross between the *o2* inbred line B73*o2* and the QPM inbred line CML161 was evaluated in two Texas locations from 2004 to 2006. Four traits related to endosperm texture showed significant effect of the inbred lines, high heritability estimates (h^2 = 0.83-0.94) and high genetic correlations (GC) among them (GC= 0.81-1.0). Relative content of the essential amino acids lysine, tryptophan and methionine also showed significant effects of the lines, genetic correlations GC = 0.63-0.71, and heritabilities h^2 = 0.71-0.80. Negative correlation was observed between endosperm texture traits and amino acid content. Nevertheless, favorable responses to selection

would be expected for both traits if selection strategies include monitoring of both endosperm texture modification and relative content of the essential amino acids.

Introduction

The structure and content of the maize endosperm influences the expression of traits targeted for genetic improvement such as grain yield (Salamini et al., 1970; Vyn and Tollenaar, 1998), grain quality (Mazur et al., 1999), suitability for industrial processing (Paulsen and Hill, 1985; Chandrashekar and Mazhar, 1999), ruminal digestibility (Corona et al., 2006), and tolerance to mycotoxin accumulation (Bhatnagar et al., 2003). The protein fraction constitutes only 8-9% of the endosperm (Lawton and Wilson, 1987), while starch accounts for about 71% of the kernel (Prasanna et al., 2001). In normal maize, 50 to 70% of the endosperm proteins are of the prolamin type (zeins). The zein proteins are particularly deficient in the essential amino acids lysine and tryptophan. The high proportion of zeins in the endosperm is the primary reason for the poor protein quality of maize (Vasal, 2000).

Several mutations, both spontaneous and induced, have been identified affecting the composition of the maize endosperm. Among them, those of the *Opaque2* (*O2*) gene are the most extensively studied (Lazzari et al., 2002; Henry et al., 2005). The *O2* gene encodes a leucine-zipper class transcription factor (Schmidt et al., 1990) that induces the transcription of a group of α-zeins and influences expression of other genes such as b-32 (a type I ribosome inactivating protein) and CyPdk1 (a cytosolic pyruvate orthophosphate dikinase) (Bass et al., 1992; Schmidt et al., 1992; Maddaloni et al.,

1996). In the 1960s, *o2* received considerable attention as it was associated with a significant increase in the proportion of lysine and tryptophan in the grain (Mertz et al., 1964). Plant breeders transferred the *o2* allele into different germplasm, aiming to improve nutritional quality, but undesirable traits were associated with the presence of the *o2* allele. In particular, the kernel became soft and more prone to mechanical damage, the yield decreased by 8-15%, and the plants were more susceptible to fungi and insects (Lambert et al., 1969; Salamini et al., 1970).

Fortunately, the soft endosperm of *o2* maize can be altered to resemble normal endosperm by the activity of modifier genes (Paez et al., 1969; Vasal, 1971). Modified *o2* genotypes developed at the International Maize and Wheat Improvement Center (CIMMYT) are called Quality Protein Maize (QPM) (Vasal, 2000). The general strategy for the development of QPM initially included backcrossing to develop *o2* varieties and hybrids from normal genotypes, and later, recurrent selection to develop specific genetic backgrounds that restored the vitreous portion of the endosperm. QPM genetic pools, populations and inbred lines have resulted from these breeding efforts. Protein quality testing during the breeding process was used in an attempt to avoid excessive losses of lysine and tryptophan. In general, QPM genotypes retain higher levels of lysine and tryptophan than normal maize materials (Ortega and Bates, 1983). Yield trials have shown that some QPM germplasm can be competitive with local checks (Pixley and Bjarnason, 1993). QPM genotypes have been introduced to production systems in several tropical and sub-tropical countries, where their nutritional advantages have been acknowledged (National Research Council, 1988; Vietmeyer, 2000).

Nevertheless, there are major challenges to develop and use QPM germplasm because of the unknown number of modifier genes required to restore the desired hard-to-soft endosperm ratio, the need to evaluate the grain quality during the breeding process, and the effects specific to the genetic background (Belousov, 1987; Ciceri et al., 2000; Huang et al., 2004; Gibbon and Larkins, 2005).

The objectives of this study were to utilize a population of recombinant inbred lines derived from an *o2* line and a QPM line to estimate variance components, phenotypic and genotypic correlations, heritabilities and to identify selection strategies for endosperm texture modification and relative content of the essential amino acids lysine, tryptophan and methionine.

Materials and Methods

Plant Material

A population of RILs was derived from a cross between B73*o2* and CML161. CML161 is an exotic subtropical inbred classified as QPM and released by CIMMYT. B73*o2* is an *opaque2* conversion of B73, an Iowa Stiff Stalk inbred. A group of 146 RILs were used for field trials (S_5, S_6, or S_7 generations), and genotypic analysis (S_6 generation).

Field Design

The RIL population was evaluated in two Texas locations during the years 2004, 2005, and 2006. The College Station (CS) location (latitude 30° 37, elevation 96 masl) at the Texas Agricultural Experiment Station in Burleson County has a humid subtropical climate and 99.3 cm of annual precipitation on average, while the Weslaco (WE) location (latitude 26° 09, elevation 22.5 masl) has a semiarid subtropical climate and 58.4 cm of annual precipitation on average (Griffiths and Bryan, 1987). The combination of location and years produced the environments WE04 (Weslaco in 2004), WE05 (Weslaco in 2005), CS05 (College Station in 2005), CS06A (College Station in 2006 first planting) and CS06B (College Station in 2006 second planting). A randomized complete block design was used in all environments: WE04 (one replication), WE05 (three replications), CS05 (two replications), CS06A (two replications), and CS06B (two replications). Each plot consisted of one row 5.2 m in length, with 0.76 (CS) or 1 m (WE) of distance between rows. Trials received common management practices according to each research station. At least 10 plants per plot were self-pollinated. All plots were manually harvested.

Endosperm Texture Modification

Endosperm texture was evaluated with four different measurements (TEXT-F, TEXT-L, OPAC and VITR) associated to the extent of modification of the endosperm in the RIL population. Endosperm texture (TEXT) was a visual rating from 1 (modified = flint-type round crown kernel and vitreous appearance) to 5 (opaque = dent-type kernels

with very high proportion of floury endosperm) with increments of 0.5. A value of TEXT was assigned to all the ears self-pollinated and harvested from each plot. Data was taken independently by one observer in the field during harvesting (TEXT-F), and two additional observers in the laboratory before shelling (TEXT-L). Opacity (OPAC) was scored in a light box using a scale 1 (modified = light passes trough the whole kernel) to 5 (opaque = no light transmission due to completely opaque kernels) (Bjarnason and Vasal, 1992). For the trait vitreousness (VITR), an image analysis-based method was adapted (Leyva-Ovalle et al., 2002). Briefly, eight kernels from bulked ears per plot were arranged embryo down in a metallic grid and sanded with a 114 mm orbital sander (Hitachi Koki, Tokyo, Japan), until approximately one third of the wide of the kernel was removed. An 8-bit black and white image was obtained by scanning the kernels in a tabletop scanner (Hewlett Packard ScanJet 3970, Palo Alto, CA) using a dark blue background. The negative of the image was used to estimate the area of soft (black) and hard (white) endosperm by the pixel counting option of the UTHSCSA Image Tool 3.0 software (Wilcox et al., 2002). VITR was defined as the percentage of the area of hard endosperm to the total endosperm area.

Amino Acid Composition

Tryptophan (Trp), Methionine (Met) and Lysine (Lys) were quantified using a microbiological method based on E. coli strains auxotrophic for Trp, Met or Lys as described by Scott et al. (2004). Kernels from bulked ears per plot were ground and

measured in triplicate. The concentration of Met, Trp or in each analysis was calculated using linear regression onto a line fitted to the standards.

Statistical Analysis

Analysis of variance (ANOVA) was conducted within and across environments for each trait separately, considering all effects in the statistical model as random, in PROC GLM of SAS 9.1 (SAS Institute, 2003). Genotype least square means were estimated per environment for each trait. Homogeneity of variances between environments was tested by the Levene's test (Glass, 1966). The Best Linear Unbiased Prediction (BLUP) procedure was used to predict the effects of each RIL across environments using univariate mixed model analysis in PROC MIXED of SAS 9.1 (SAS Institute, 2003). Heritabilities across environments for each trait were calculated on a plot and entry-mean basis using univariate mixed model analysis (all effects random) in PROC MIXED of SAS 9.1 (Holland et al., 2003; SAS Institute, 2003). Because VITR, Lys, Trp and Met measurements were taken on a single repetition in 2 or 4 locations, the heritability estimates are presented as repeatabilities. The Phenotypic (PC) and genotypic correlations (GC) among pair of traits were computed across environments using multivariate restricted maximum likelihood (REML) estimation in PROC MIXED (Holland, 2006; SAS Institute, 2003). The relationship between traits and RILs was assessed by Single Value Decomposition (SVD) of standardized variables using Biplot v1.1 (Lipkovich and Smith, 2002). Data were presented as a biplot with two principal

components. Direct response to selection and correlated response were calculated according to Falconer and Mackay (1996).

Results and Discussion

A population of RILs was derived from crossing an *o2* temperate line (B73*o2*) with a QPM tropical line (CML161). Inbred CML161 is a high yielding QPM line commonly used in tropical and semi-tropical breeding programs where it has been used in multiple hybrid combinations (Prasanna et al., 2001; Xingming et al., 2001; Bhatnagar et al., 2003). Inbred B73 has become the most popular maize line in maize genomic studies and played a key role in the development of U.S hybrids.

Trait Means and Variation

Significant ($p < 0.001$) main effects of the RILs were observed in the ANOVA for all traits (Table 2.1). Significant ($p < 0.001$) main effects of the environments were observed for TEXT-L, OPAC, VITR, Lys, Trp and Met. In addition, significant genotype x environment interaction (GEI) ($p < 0.001$) was observed for the traits TEXT-F, TEXT-L, and OPAC. Significant GEI have been reported for other QPM materials for endosperm modification and other traits such as grain yield, protein in grain, Trp in grain and protein (Pixley and Bjarnason, 2002; Lou et al., 2005). In normal maize it has been shown that environmental factors like availability of nutrients in the soil, water stress and heat stress influence the process of kernel filling and the accumulation of

Table 2.1. Combined analysis of variance for endosperm texture in the field (TEXT-F), endosperm texture in the laboratory (TEXT-L), opacity (OPAC), vitreousness (VITR), lysine content (Lys), tryptophan content (Trp), and methionine content (Met) in an *o2* x QPM RIL population evaluated in Texas between 2004 and 2006.

Trait	Sources of variation	df	Mean square	P > F	Variance component
TEXT-F	Env	2	0.0924	0.7328	_
(score 1-5)	Rep/Env	3	0.4042	0.2547	_
	Lines	142	6.3021	<.0001	1.150
	Env*Lines	278	0.4748	<.0001	0.101
TEXT-L	Env	3	3.2146	<.0001	0.009
(score 1-5)	Rep/Env	5	0.5522	0.2038	_
	Lines	142	4.6312	<.0001	0.535
	Env*Lines	419	0.528	0.0001	0.072
OPAC	Env	3	27.7953	<.0001	0.121
(score 1-5)	Rep/Env	4	0.526	0.088	
	Lines	142	4.7937	<.0001	0.719
	Env*Lines	421	0.435	<.0001	0.104
VITR	Env	1	1223.94	<.0001	9.123
(%hard/total area)	Lines	142	264.0364	<.0001	114.091
Lys	Env	3	4.15	<.0001	0.027
(rel. units)	Lines	142	1.3461	<.0001	0.256
Trp	Env	3	179.7089	<.0001	1.260
(rel. units)	Lines	142	10.4832	<.0001	2.121
Met	Env	3	53.6035	<.0001	0.374
(rel. units)	Lines	142	2.7528	<.0001	0.491

storage proteins (Hamilton et al., 1951; Hadi, 2004; Monjardino et al., 2005). The relative content of the amino acids lysine, tryptophan and methionine were measured. The parental line B73*o2* showed consistently higher values of TEXT-F, TEXT-L, OPAC, Met, Trp and Lys; whereas CML161 had higher values of VITR (Fig. 2.1, Fig.2.2). As expected, B73*o2* had consistent and significantly softer endosperm than CML161.

The mean relative Lys content was 0.100 for B73*o2* and 0.087 for CML161. The mean relative Trp content was 0.200 for B73*o2* and 0.136 for CML161. The mean relative Met content was 0.129 for B73*o2* and 0.101 for CML161. Inbred B73*o2* had consistently higher relative content of these amino acids than inbred CML161. These results are consistent with previous estimates were B73*o2* had high values of Lys, Trp and Met (Scott et al., 2004).

The frequency distribution of each trait in the population of RILs approached a normal distribution for all traits except for TEXT-F, which showed bimodal distribution with peaks towards the semi-modified and opaque categories, and for OPAC, which intermediate categories had fewer individuals (Fig. 2.2). Transgressive segregation was not significantly observed for any trait. Both parental inbreds were among the groups with most extreme expression of these traits. Transgressive segregation has been currently observed in several traits in mapping populations (DeVicente and Tanskley, 1993; Rieseberg, 1999). The nature of the traits addressed in this research and the selection of two inbreds with very extreme expression of endosperm traits could explain the lack of transgressive segregation.

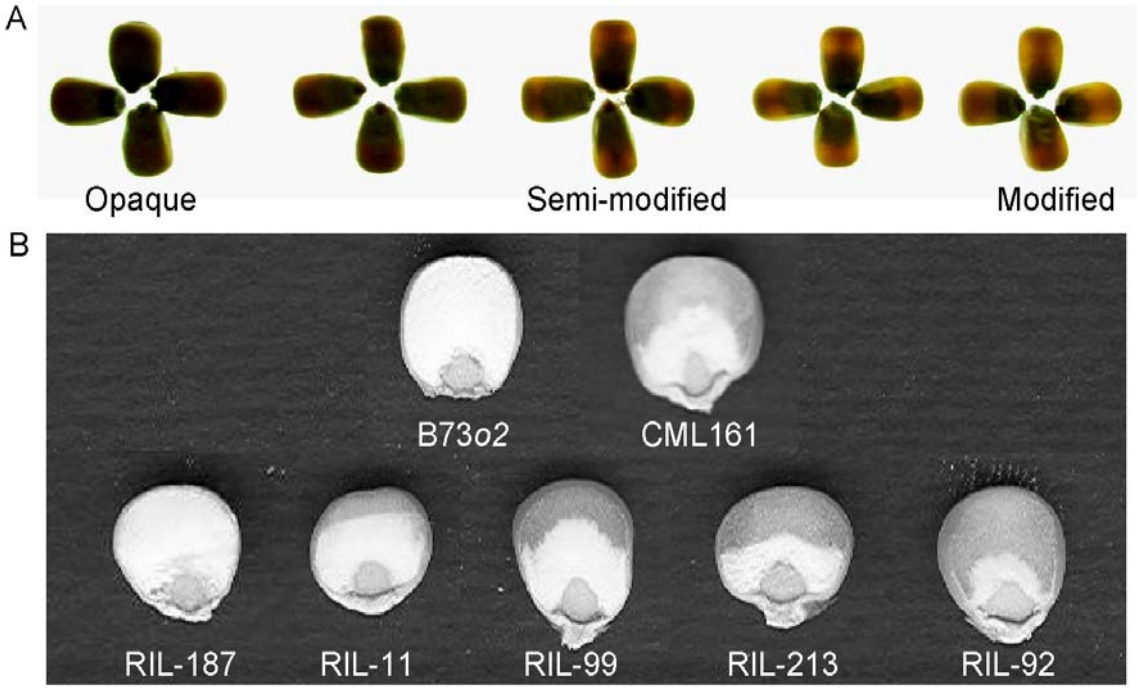

Figure 2.1. Endosperm texture modification. (A) Photograph of backlit kernels showing segregation of the trait opacity (OPAC) in the F_2 of B73*o2* x CML161. (B) Kernel sections (8-bit images) of B73*o2*, CML161 and five recombinant inbred lines used for measuring the trait vitreousness (VITR).

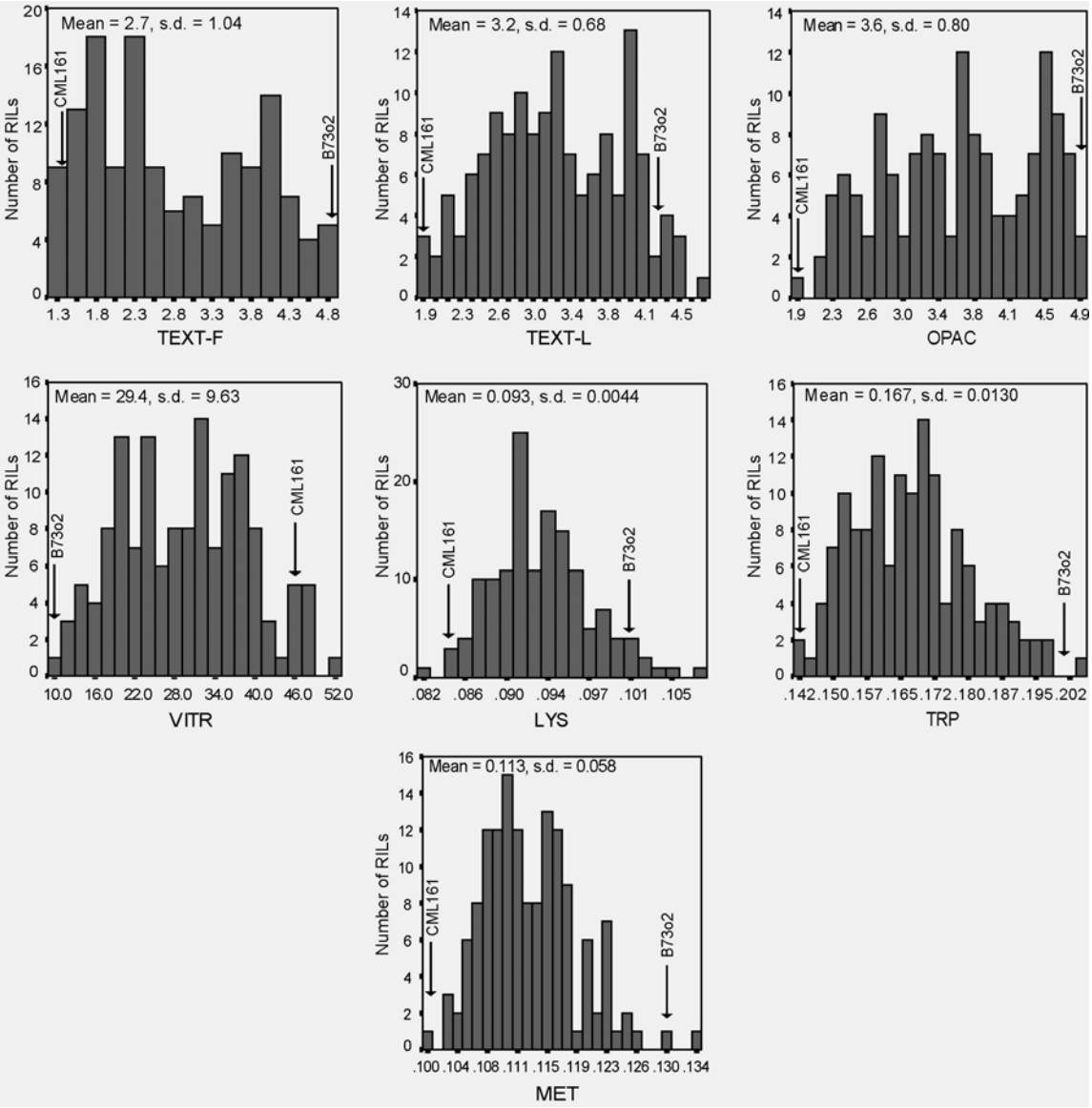

Figure 2.2. Frequency distributions of endosperm texture modification (TEXT-F, TEXT-L, OPAC and VITR), and amino acid content (Lys, Trp and Met) in an *o2* x QPM RIL population evaluated in Texas between 2004 and 2006. Mean values of the parents B73*o2* and CML161 are showed by arrows.

The Levene's test suggested homogeneity of variances for all traits between the different environments except for Trp, in which the two CS locations showed a shift in the distribution towards a higher mean value. Both TEXT-F and TEXT-L correspond to visual estimates of the endosperm texture of whole ears measured under different conditions by different observers. However, the test for homogeneity of variance suggested heterogeneous variances between TEXT-F and TEXT-L, and therefore they were treated as different traits.

Given the contrasting differences between these two inbreds, CML 161 and B73*o2*, the amount of variation observed was high for all the traits measured. Hence, this RIL population was used as mapping population for a subsequent study (Chapter III).

Phenotypic Correlations, Genotypic Correlations and Traits Biplot

It has been suggested that the BLUP procedure is useful to analyze unbalanced data (Bernardo, 2002). Since not all traits were evaluated in all the environments, BLUP predictions were used for SVD of traits. In SVD biplots, an angle of 90° between any two trait vectors suggest that the traits are independent, an angle of >90° indicates negative correlation, and an angle of <90° suggests a positive correlation. SVD illustrated the negative relationship between the opacity-based measurements (TEXT and OPAC) and the vitreousness-based measurement (VITR). In addition, the SVD biplot showed how the three amino acid traits are closely correlated and their positive but low correlation with the endosperm texture traits (Fig. 2.3).

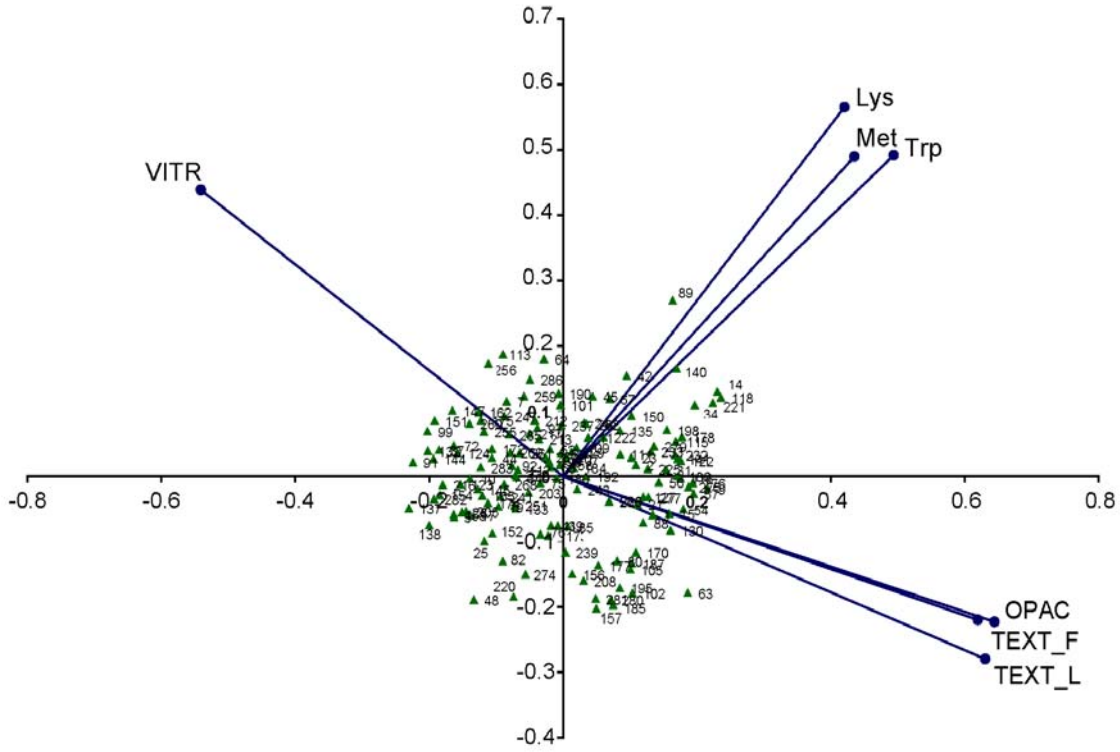

Figure 2.3. Single value decomposition biplot of first two principal components for endosperm texture in the field (TEXT-F), endosperm texture in the laboratory (TEXT-L), opacity (OPAC), vitreousness (VITR), lysine content (Lys), tryptophan content (Trp), and methionine content (Met) in an *o2* x QPM RIL population evaluated in Texas between 2004 and 2006.

The phenotypic correlations between all the four texture-related traits ranged between 0.66 and 0.75. The genotypic correlations ranged between 0.81 and 1.00. The phenotypic correlation between TEXT-F and TEXT-L was 0.70 and the genetic correlation was 0.95. Due to the scale use in the measurements, VITR showed negative

correlation with TEXT-F, TEXT-L and OPAC (Table 2.2, Fig. 2.3). The phenotypic correlation between Lys and Met was 0.41, whereas the genotypic correlation was 0.63. The phenotypic correlation between Lys and Trp was 0.62, and the genetic correlation was 0.71. The phenotypic correlation between Trp and Met was 0.54, and the genetic correlation was 0.68. Scott et al. (2004) analyzed a set of opaque and QPM lines and hybrids and observed a general positive correlation between lysine and tryptophan, but a negative correlation with methionine. The positive correlation among amino acids content in this work could relate to the observation that the *o2* parent contained higher relative amounts of lysine, tryptophan and methionine while the modified parent CML161 showed lower content of these amino acids. In addition, the nature of diversity displayed by the RIL population used in this study and that displayed among the lines used by Scott et al (2004) is different.

In the population of RILs derived from B73*o2* x CML161, there was a positive correlation between endosperm opacity and relative content of lysine, tryptophan and methionine. The relative values of these amino acids decreased as the inbred lines tended from opaque towards flint/dent modified kernels. However, there were RILs with high content of amino acids and degree of modification suggesting the possibility to select for both simultaneously. Contradictory reports have described both neutral (Pixley and Bjarnason 2002; Bantte and Prasanna, 2004) and negative (Paez et al., 1969; Robutti et al., 1974a; Wessel- Beaver 1985) correlations for endosperm

Table 2.2. Phenotypic (PC) and genotypic correlation (GC) estimates for endosperm texture in the field (TEXT-F), endosperm texture in the laboratory (TEXT-L), opacity (OPAC), vitreousness (VITR), lysine content (Lys), tryptophan content (Trp), and methionine content (Met) in an *o2* x QPM RIL population evaluated in Texas between 2004 and 2006.

	TEXT-F		TEXT-L		OPAC		VITR		Lys		Trp	
	PC	**GC**	**PC**	**GC**	**PC**	**GC**	**PC**	**GC**	**PC**	**GC**	**PC**	**GC**
TEXT-F	_	_										
SE												
TEXT-L	0.695	0.948	_	_								
SE	0.027	0.018										
OPAC	0.782	0.940	0.673	1.009	_	_						
SE	0.023	0.017	0.028	0.013								
VITR	-0.675	-0.811	-0.664	-0.947	-0.754	-0.942	_	_				
SE	0.022	0.032	0.038	0.032	0.030	0.057						
Lys	0.318	0.455	0.236	0.429	0.322	0.503	-0.180	-0.297	_	_		
SE	0.057	0.083	0.053	0.088	0.050	0.080	0.014	0.045				
Trp	0.390	0.551	0.307	0.576	0.413	0.632	-0.343	-0.325	0.620	0.713	_	_
SE	0.054	0.073	0.051	0.073	0.048	0.065	0.004	0.007	0.022	0.025		
Met	0.355	0.540	0.242	0.524	0.271	0.509	-0.207	-0.305	0.412	0.633	0.538	0.681
SE	0.055	0.081	0.054	0.085	0.093	0.146	0.008	0.024	0.038	0.047	0.014	0.016

modification and amino acid content. Here, grain yield analysis was not included although there is evidence for lack of phenotypic correlations between yield and endosperm modification, but negative correlation between yield and amino acid content (Pixley and Bjarnason, 2002).

Broad-Sense Heritability and Gain from Selection

The heritabilities (entry mean basis) calculated across environments for TEXT-F, TEXT-L and OPAC were 0.94, 0.90 and 0.92, respectively. The heritabilities (plot basis) calculated for TEXT-F, TEXT-L and OPAC were 0.75, 0.54 and 0.66, respectively (Table 2.3). Due to the intrinsic nature of the measurements, only one repetition per environment was used for the quantification of VITR and amino acid content (Lys, Trp and Met). This could introduce a bias in the estimation of heritabilities, hence these estimates are referred as estimates of repeatability. The estimates of repeatability for VITR were 0.83 (entry mean basis) and 0.71 (plot basis). Repeatabilities on plot basis calculated across environments for Lys, Trp and Met were 0.44, 0.50 and 0.38, respectively. Repeatabilities on entry mean basis calculated across environments for Lys, Trp and Met were 0.76, 0.80 and 0.71, respectively (Table 2.3). When inbred lines are developed from an F_2 population, the variance among inbred lines is equal to twice the original additive variance in the starting F_2 population (Bernardo, 2002). Therefore, the heritability/repeatability estimates calculated from inbred lines could be higher than estimates made using other mating designs.

Table 2.3. Heritability and repeatability estimates for endosperm texture in the field (TEXT-F), endosperm texture in the laboratory (TEXT-L), opacity (OPAC), vitreousness (VITR), lysine content (Lys), tryptophan content (Trp), and methionine content (Met) in an *o2* x QPM RIL population evaluated in Texas between 2004 and 2006.

Trait (environments, repetitions)	Plot basis	Entry mean basis
TEXT-F (e=3, r=6)	**0.748**	**0.935**
SE	0.027	0.010
TEXT-L (e=4, r=9)	**0.536**	**0.897**
SE	0.036	0.014
OPAC (e=4, r=8)	**0.659**	**0.924**
SE	0.032	0.107
VITR (e=2, r=2)	**0.707**	**0.829**
SE	0.043	0.030
Lys (e=4, r=4)	**0.435**	**0.755**
SE	0.045	0.034
Trp (e=4, r=4)	**0.503**	**0.802**
SE	0.043	0.027
Met (e=4, r=4)	**0.377**	**0.708**
SE	0.046	0.040

With the estimates of heritability and the additive variance it is possible to predict the response to selection (R) (Falconer and Mackay, 1996). The direct selection response was calculated using a selection intensity of 20%. As expected, the results suggested that a substantial gain can be expected when phenotypic selection is used for endosperm texture. Gains were higher when entry means were used as selection units. The maximum estimated gain was about 40%, for the trait TEXT-F with entry means selection (Table 2.4). For amino acid content, the gain was lower with a maximum of 7.7% of gain for Trp. Calculated gain in Lys was 4.7%. As for endosperm texture, selection based on entry means produced higher gain in amino acid content (Table 2.4). As mentioned before, amino acid content and endosperm modification traits were negatively correlated in the population of RILs. The correlated response to selection (CR) can be predicted to anticipate the change in amino acid content when selecting for endosperm modification (Table 2.5). Selection for any of the four texture-related traits produced a decrease in the content of the three amino acids. Approximately 3% decrease in the content of lysine was observed when selecting for endosperm texture (Table 2.5).

Studies of the genetic control of the observed variation on amino acid content and endosperm texture in the population of RILs will be described in chapter III. The results suggested that there are both common and independent genomic regions controlling endosperm texture and amino acid content. As a consequence, if the correlated loss in amino acid content caused by selection for endosperm is found to be significant for the purposes of the breeding program, then evaluation for both traits must be carried out when developing QPM genotypes (Wessel-Beaver at al., 1985). However,

Table 2.4. Predicted response to selection for endosperm texture in the field (TEXT-F), endosperm texture in the laboratory (TEXT-L), opacity (OPAC), vitrousness (VITR), lysine content (Lys), tryptophan content (Trp), and methionine content (Met) with both plot basis and entry means-based selection units.

Trait	Selection Response†		Selection Response as percentage of the mean	
	plot basis	Entry mean basis	plot basis	entry mean basis
			---------- %	----------
TEXT-F	0.92	1.03	35.49	39.68
TEXT-L	0.53	0.69	16.51	21.36
OPAC	0.68	0.81	18.90	22.39
VITR	8.89	9.63	30.05	32.52
LYS	0.33	0.44	3.55	4.67
TRP	1.02	1.29	6.12	7.72
MET	0.43	0.58	3.78	5.18

† Directed selection response $Rx = i*hx*\sigma_A X$ ((i= 1.4, s = 20%) (Falconer and Mackay, 1996).

Table 2.5. Correlated response of lysine content (Lys) when selecting for endosperm texture in the field (TEXT-F), endosperm texture in the laboratory (TEXT-L), opacity (OPAC), vitreousness (VITR), lysine content (Lys), tryptophan content (Trp), and methionine content (Met) with entry means-based selection units. (GC= genetic correlation).

Selecting for	$GC_{X\text{-}Lys}$	Entry mean basis	
		-------%	------
TEXT-F	**0.455**	0.31	3.35
TEXT-L	**0.429**	0.29	3.09
OPAC	**0.503**	0.34	3.68
VITR	**-0.297**	-0.19	-2.06

selection for amino acid content based on direct measurements of each amino acid is not a feasible alternative for most breeding programs (Prasanna et al., 2001). Alternatives such as specific calibrations of Near Infrared Reflectance Spectroscopy (NIRS) for grain quality traits have shown that thousands of samples can be processed in a relative short period of time at reasonable cost (Fontaine et al., 2002; Montes et al., 2006).

Conclusion

Sufficient genetic variation for endosperm and amino acids content was observed in this study. This variation supported intermediate to high heritabilities and the existence of RILs with desirable combination of high amino acids content and endosperm vitreousness. The development of segregating population between opaque with high amino acids content and QPM material could be a breeding approach to

enhance further the protein quality of QPM lines (e.g., CML161). The lack of significant transgressive segregation may be the consequence of using parents with very extreme expression of these traits. Parents with intermediate expression should provide segregating populations showing transgressive segregation.

High heritabilitilies and genetic correlations suggest high direct and indirect genetic gain possible by single traits. Nevertheless, negative correlation between desirable endosperm traits and amino acids contents may indicate the need of monitoring both types of traits and the use of selection indices during the inbreeding and selection process.

The quantification of endosperm vitreousness by image analysis was conducted successfully and can be used in breeding programs. It was highly correlated with visual endosperm ratings but some differences were appreciated among RILs ranking.

Based on the results presented here, it may be possible to further enhanced the protein and nutritional value of QPM by measuring endosperm characteristics and levels of Lys, Trp, and Met during breeding and selection of recombinant lines (i.e., conventional or double haploids) and select those having vitreous endosperm and high levels of these amino acids.

CHAPTER III

QUANTITATIVE TRAIT LOCI FOR ENDOSPERM MODIFICATION AND AMINO ACID CONTENTS

Overview

The deficient protein quality of corn grain can be improved by replacing the normal *Opaque2 (O2)* alleles with non-functional mutant alleles *o2*. Unfortunately, *o2* alleles are associated with a very soft endosperm texture, poor yield and susceptibility to diseases and insects. Plant breeders have been able to restore a desirable ratio of hard to soft endosperm in *o2* germplasm. These modified genotypes are known as Quality Protein Maize (QPM). Neither the mechanism nor the genetic components by which the modification of the endosperm in QPM lines occurs are well understood. Using a population of Recombinant Inbred Lines (RILs), derived from a cross between an *o2* line and a QPM line and a novel evaluation method for endosperm modification, quantitative trait loci (QTL) were mapped for traits related to the modification of endosperm texture and the content of the essential amino acids lysine, tryptophan and methionine. For endosperm texture traits, QTLs clusters were detected on chromosomes 3, 5, 6 and 8 accounting together for 62-68% of the observed variation. For traits related to amino acid contents, QTLs clusters were located on chromosomes 7 and 8 that explained up to 39% of the observed variation. The elucidation of the genetic mechanisms of the modification of *o2* endosperm and essential amino acid contents will provide valuable

information and important tools to plant breeders as well as plant scientists interested in improving the quality of the cereal seeds.

Introduction

Maize, a major crop, is a source of food, feed, and raw materials for several industries. The grain is the major harvested product of maize. The main components of the kernel are the endosperm, the embryo, and the pericarp (Watson, 2003). The endosperm is a triploid tissue originating after a male gamete fertilizes the diploid central cell in a process parallel to the fertilization of the egg cell that gives origin to the diploid zygote (Olsen, 2004). The main role of the endosperm is the synthesis and accumulation of storage products to nurture the embryo during initial stages of germination and seedling development (Costa et al., 2004). Reserves in the endosperm accumulate in the form of lipids, carbohydrates, and proteins. The structure and content of the endosperm affects traits targeted for genetic improvement such as grain yield (Salamini et al., 1970; Vyn and Tollenaar, 1998), grain quality (Mazur et al., 1999), suitability for industrial processing (Paulsen and Hill, 1985; Chandrashekar and Mazhar, 1999), ruminal digestibility (Corona et al., 2006), and tolerance to mycotoxin accumulation (Bhatnagar et al., 2003).

Protein constitutes less than 10% of the kernel, while starch accounts for about 70% (Lawton and Wilson, 1987; Prasanna et al., 2001). In normal maize, 50-70% of the endosperm proteins are of the prolamin type (zeins). Zein proteins are often classified by differential solubility and mobility through gels as α-zein, β-zein, δ-zein, and γ-zein

(Coleman et al., 1997). Zeins accumulate in sub-cellular compartments derived from the endoplasmic reticulum known as protein bodies (Lending and Larkins, 1989). The zein fraction is particularly deficient in the essential amino acids lysine (Lys) and tryptophan (Trp). The high proportion of zeins in the endosperm is the primary reason for the poor protein quality of maize (Vasal, 2000).

Several genes affecting the composition and structure of the maize endosperm have been identified. The *Opaque2* (*O2*) gene is one of the most extensively studied (Lazzari et al., 2002; Henry et al., 2005). Mertz et al. (1964) realized that maize genotypes homozygous for the mutant allele *o2* had a considerable higher content of Lys and Trp in the grain when compared to wild type genotypes. After this discovery, plant breeders started to introgress *o2* alleles into different germplasm, trying to improve their nutritional quality, but undesirable traits were associated with *o2*. In particular, the kernel became soft and very prone to mechanical damage, the yield decreased 8-15%, and the plants were more susceptible to fungi and insects (Lambert et al., 1969; Salamini et al., 1970).

The *O2* gene encodes a leucine-zipper class transcription factor (Schmidt et al., 1990) required for efficient transcription of a group of α-zeins and influences expression of other genes such as *b-32* (a type I ribosome inactivating protein) and *CyPpdk1* (a cytosolic pyruvate orthophosphate dikinase) (Bass et al., 1992; Schmidt et al., 1992; Maddaloni et al., 1996). The soft endosperm texture of *o2* germplasm is associated with the reduction in the proportion of α-zein proteins (Huang et al., 2004; Huang et al., 2005). It has been postulated that the absence of specific zeins in *o2* genotypes causes

the formation of smaller protein bodies and therefore alters the packing of the starch fraction during seed desiccation, resulting in abnormally soft endosperm (Schmidt et al., 1990). However, the endosperm of *o2* maize can be restored to resemble normal endosperm by the activity of modifier genes (Paez et al., 1969; Vasal, 1971). Modified *o2* genotypes with hard endosperm developed at the International Maize and Wheat Improvement Center (CIMMYT) are called Quality Protein Maize (QPM) (Vasal, 2000). In general, QPM genotypes retain higher levels of Lys and Trp than normal maize materials (Ortega and Bates, 1983).

Several investigations have aimed to identify the molecular mechanisms responsible for the endosperm modification and grain quality in QPM maize (Gibbon and Larkins, 2005). Lopes et al. (1995) using bulk segregant analysis (BSA) in populations developed from *o2* and modified varieties, found two loci associated with the modification of endosperm. Another report based on BSA reported two loci associated with endosperm modification, one of them in the same chromosome region reported by Lopes et al. (1995), and another one in a different genome location (Lizarraga-Guerra et al., 2006). These and other studies suggest that the γ-zeins and their regulation are involved in the modification of *o2* endosperm (Burnett and Larkins, 1999; Geetha et al., 1991). The 27 kDa γ-zeins seem to be involved in the development of protein bodies and in the formation of a protein network that surrounds starch grains in vitreous endosperm (Dannenhoffer et al., 1995). In addition, the branching pattern of the starch and the genetic variability of starch biosynthesis enzymes have been associated with modified phenotypes (Gibbon et al., 2003).

QPM genotypes have been introduced into production systems in several tropical and sub-tropical countries, where their nutritional advantages have been acknowledged (National Research Council, 1988; Vietmeyer, 2000). Nevertheless, there are major challenges to develop and use QPM germplasm because of the unknown number of modifier genes required to restore the desired hard-to-soft endosperm ratio, the need to evaluate the grain quality during the breeding process, and the genetic background effects (Belousov, 1987; Ciceri et al., 2000; Huang et al., 2004; Gibbon and Larkins, 2005). Characterization of the genes involved in the modification of *o2* genotypes will provide valuable information to understand the genetic basis of endosperm composition and to facilitate the improvement of grain-related traits in maize. The objective of this study was to identify QTLs for several traits associated with the modification of endosperm, including endosperm texture and amino acid content, using a population of recombinant inbred lines (RILs) developed from the cross of an *o2* line and a QPM line.

Materials and Methods

Plant Material

Details of the development of the population of RILs and the experimental design were as described in Chapter II. A population of RILs was derived from a cross between B73*o2* and CML161. CML161 is an exotic subtropical inbred classified as QPM and released by CIMMYT. B73*o2* is an *opaque2* conversion of B73, an Iowa Stiff

Stalk inbred. A group of 146 RILs were used for field trials (S_5, S_6, or S_7 generations), and genotypic analysis (S_6 generation).

Field Design

The RIL population was evaluated in two Texas locations during the years 2004-2006, producing the environments WE04 (Weslaco in 2004, one replication), WE05 (Weslaco in 2005, three replications), CS05 (College Station in 2005, two replications), CS06A (College Station in 2006 first planting, two replications) and CS06B (College Station in 2006 second planting, two replications). A randomized complete block design was used in all environments. Trials received common management practices according to each research station. At least 10 plants per plot were self-pollinated. All plots were manually harvested.

Endosperm Texture Modification

Endosperm texture was evaluated with four different measurements (TEXT-F, TEXT-L, OPAC and VITR) associated with the extent of modification of the *o2* endosperm in the RIL population. Endosperm texture (TEXT) was based on a visual rating from 1 (modified = flint-type round crown kernel and vitreous appearance) to 5 (opaque = dent-type kernels with very high proportion of floury endosperm) with increments of 0.5. A value of TEXT was assigned to self-pollinated ears that were harvested from each plot. Data was taken independently by one observer in the field during harvesting (TEXT-F), and two additional observers in the laboratory before

shelling (TEXT-L). Opacity (OPAC) was scored in a light box using a scale 1 (modified = light passes trough the whole kernel) to 5 (opaque = no light transmission due to completely opaque kernels) (Bjarnason and Vasal, 1992). For the trait vitreousness (VITR), an image analysis-based method was adapted (Leyva-Ovalle et al., 2002). Eight-bit black and white images were obtained by scanning longitudinal dissected kernels in a tabletop scanner (Hewlett Packard ScanJet 3970, Palo Alto, CA). The negative of the image was used to estimate the area of soft (black) and hard (white) endosperm using the pixel counting option of the UTHSCSA Image Tool 3.0 software (Wilcox et al., 2002). VITR was defined as the percentage of the area of hard endosperm to the total endosperm area.

Amino Acid Composition

Tryptophan (Trp), methionine (Met) and lysine (Lys) were quantified using a microbiological method based on *E. coli* strains auxotrophic for Trp, Met or Lys as described by Scott et al. (2004). Kernels from bulked ears from each plot were ground and measured in triplicate. The concentration of Met, Trp or Lys in each analysis was calculated using linear regression onto a standard curve developed using known amounts of pure amino acid standards.

Genotyping and Linkage Map

DNA was extracted from hypocotyl tissue obtained in bulk from 5 to 10 seedlings per RIL using a CTAB/sorbitol extraction buffer. Simple sequence repeat

(SSR) markers were selected from the Maize Genetics and Genomics Database (Lawrence et al., 2004) based on consensus map location. SSR markers were amplified by polymerase chain reaction (PCR) and separated by electrophoresis in 4% (w/v) Super Fine Resolution agarose gels (Amresco, Solon, OH) using standard protocols (CIMMYT, 2005). A linkage map was constructed in MapMaker/EXP 3.0b (Lander et al., 1987) using a set of 180 SSR markers. The Kosambi's mapping function was used to transform recombination frequencies into map distances in centimorgans (cM).

Statistical Analysis

Analysis of variance (ANOVA), phenotypic and genetic correlations among the traits, and heritability estimates were reported in Chapter II. For the analysis of QTLs, genotype least square means were estimated for each environment for each trait, and the best linear unbiased prediction (BLUP) procedure was used to predict the effects of each RIL across environments using univariate mixed model analysis in PROC MIXED of SAS 9.1. (SAS Institute, 2003).

Quantitative Trait Loci Analysis

Analysis of QTLs was conducted in Windows QTL Cartographer version 2.5 (Wang et al., 2006) for each trait using least square means calculated for each environment, and BLUPs estimated across environments. Single marker analysis (SMA) was used to find associations between individual SRR markers and the endosperm traits means. Composite interval mapping (CIM) was implemented using the forward and

backward regression method (probability in = 0.1, probability out = 0.1), 5 cM window size, and 1 cM walk speed. Significance thresholds for the LOD scores corresponding to a type I error rate of 5% and 10% were determined by permutation tests (n=1,000 permutations) (Churchill and Doerge, 1994). Significant QTLs detected by CIM were incorporated in multiple interval mapping (MIM) to estimate their effects and to investigate possible QTL x QTL interactions or epistasis. In addition, novel models were developed in MIM by creating initial MIM models by forward regression on markers with significance level 0.005, based on probability of partial R^2. Subsequent MIM models were accepted when decreasing the Bayesian Information Criteria (BIC). QTL detection results are reported for the MIM model with lower BIC after several rounds of search, optimization and testing.

Results

One-hundred forty-six RILs derived from B73*o2* x CML161were screened for the presence of a mutant allele *o2* using the *O2* gene-specific SSR markers phi057 and umc1066. Three RILs were removed from the QTL analysis after showing unexpected fragment sizes for these two markers. One hundred-eighty (180) polymorphic SSR markers were used to genotype the population of RILs. The length of the linkage map constructed was 1798.1 cM, and the average distance between markers was 10.05 cM (Fig. 3.1). The order of the loci in the linkage map was consistent with the consensus map of maize (Schaeffer at al., 2006; Lawrence et al., 2004). The statistical analysis of

the traits, phenotypic and genotypic correlations, and heritability estimates were reported in chapter II.

Endosperm Texture Modification

Given the contrasting differences between the two parental inbreds, CML 161 and B73*o2*, the amount of variation observed was high for all the traits measured. Considerably high heritability estimates (h^2 entry-mean basis = 0.83-0.94) were calculated for the four measurements of endosperm modification as described in chapter II. CIM analysis detected QTLs for the four traits in chromosomes 1, 3, 4, 5, 6, 8, and 10 (Fig. 3.1 and Table 3.1). Both TEXT traits corresponded to visual ratings of endosperm texture taken on whole ears. Seven QTLs were identified for the trait TEXT-F over all environments. These QTLs explained 77% of the phenotypic variation. In the QTLs of bins 3.02 (phi374118), 4.05 (bnlg1755), 5.05 (umc2026), 6.00 (umc1143) and 8.05 (umc1562), the allele increasing endosperm modification was the one derived from the QPM parent. The QTL on bin 8.05 near to the marker umc1562 explained 27.5% of the variation. In the QTLs of bins 1.02 (umc1976) and 6.06 (umc2170), the allele that decreased the proportion of opaque endosperm (i.e., more vitreous endosperm) was the one derived from the *o2* parent. A novel MIM model containing six QTLs and two QTL

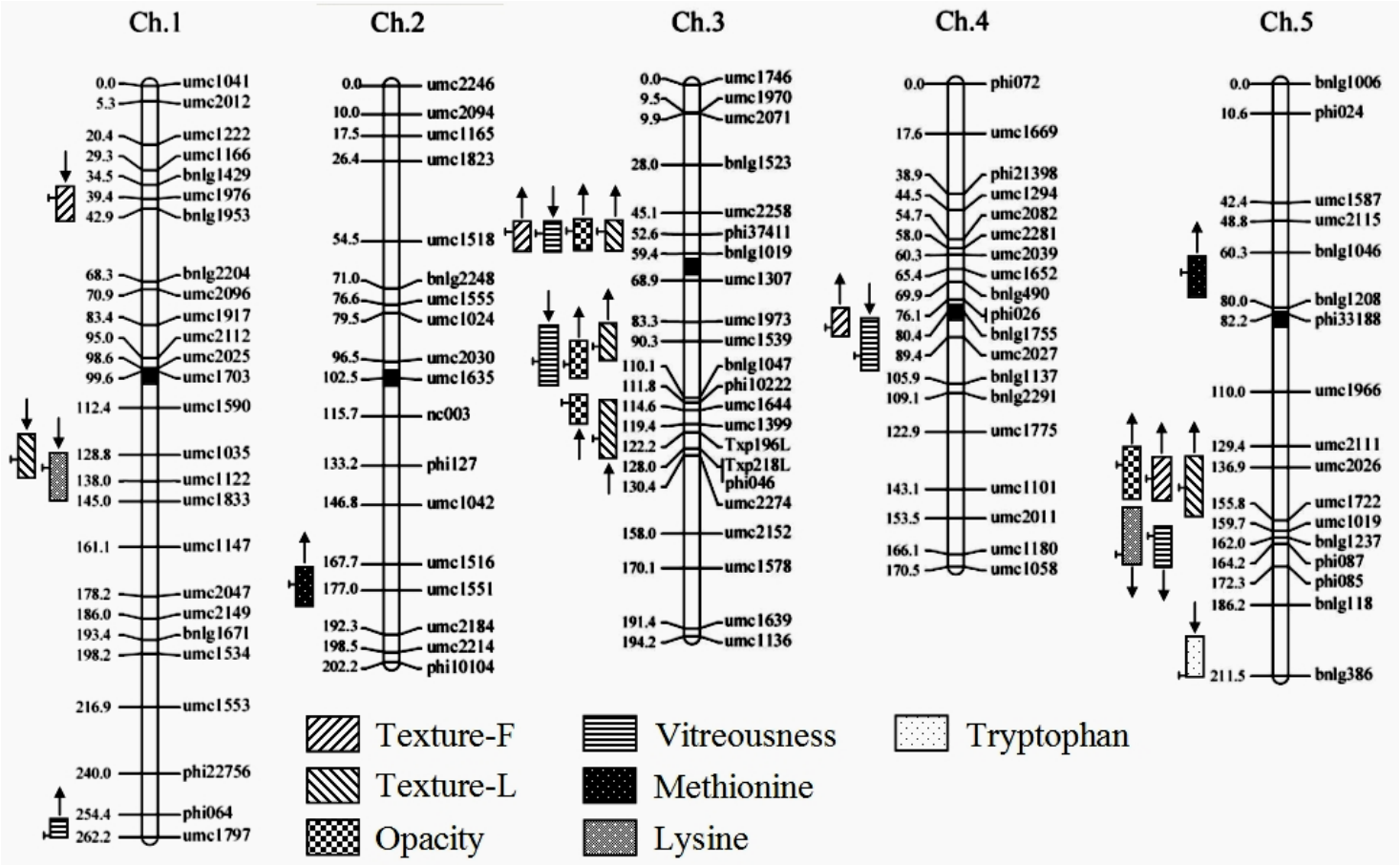

Figure 3.1. QTLs detected with CIM analysis for traits related with endosperm texture modification (TEXT-F, TEXT-L, OPAC, and VITR) and amino acids content (Lys, Trp and Met) in a population of RILs derived from the cross between maize lines CML161 and B73*o2* evaluated in Texas during 2004-2006. Arrows indicate the direction of B73*o2* allele phenotypic effect (up, increasing; down, decreasing). The length of the QTL boxes depicts the 2-LOD support intervals.

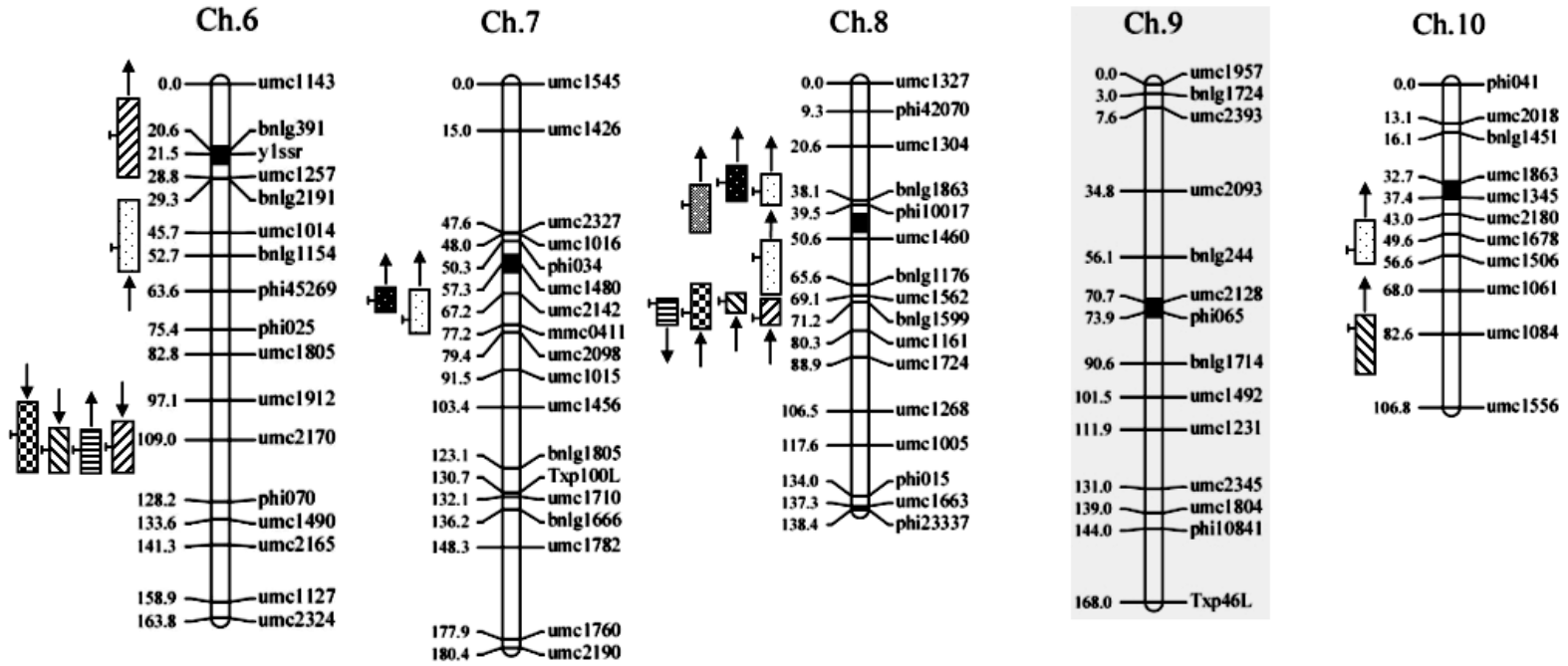

Figure 3.1. Continued.

Table 3.1. Quantitative trait loci (QTLs) identified by composite interval mapping analysis across environments for the traits related to endosperm texture modification: TEXT-F, TEXT-L, OPAC, and VITR. QTL peaks or chromosome position with the highest LOD value, closest marker to the peak, and the bin of the marker according to the consensus maize map are shown.

Trait	Bin	QTL peak	Marker	QTL interval	LOD	QTL Effect‡	R2 (%)§
Texture field (TEXT-F)	1.02	39.38	umc1976	37-44	5.14*	-0.24	4.2
	3.02	52.6	phi374118	49-58	4.16*	0.31	8.2
	4.05	84.45	bnlg1755	78-88	3.01†	0.17	3.6
	5.05	142.92	umc2026	134-150	4.53*	0.33	14.7
	6.00	17.01	umc1143	8-29	3.11†	0.22	5.9
	6.06	113.03	umc2170	101-120	5.32*	-0.33	12.9
	8.05	70.06	umc1562	69-73	15.58*	0.49	27.5
Texture lab (TEXT-L)	1.06	132.81	umc1035	123-137	5.00*	-0.11	5.1
	3.02	52.6	phi374119	49-57	5.18*	0.17	7.8
	3.05	93.28	umc1539	85-103	3.91*	0.09	5.2
	3.07	124.25	Txp196L	113-129	2.60	0.08	4.8
	5.05	143.92	umc2026	133-153	3.90*	0.15	8.2
	6.06	115.03	umc2170	106-121	8.84*	-0.30	23.1
	8.05	70.06	umc1562	69-71	10.62*	0.25	19.5
	10.06	81	umc1061	73-93	2.50†	0.09	3.8
Opacity (OPAC)	3.02	52.6	phi374120	48-59	2.88†	0.18	5.7
	3.05	98.28	umc1539	90-105	5.50*	0.21	12.1
	3.06	113.78	phi102228	111-119	2.90†	0.12	6.2
	5.05	136.42	umc2111	130-150	3.66*	0.20	7.8
	6.06	107.12	umc1912	102-119	7.93*	-0.26	13.4
	8.05	74.2	bnlg1599	66-78	6.04*	0.27	17.3
Vitrousness (VITR)	1.11	261.38	phi064	256-261	2.81†	1.78	0.6
	3.02	53.6	phi374121	50-57	4.97*	-2.84	9.3
	3.05	98.28	umc1539	86-106	2.87†	-2.54	11.6
	4.06	94.38	umc2027	82-102	3.05†	-1.93	3.7
	5.05	161.97	bnlg1237	157-172	4.49*	-2.33	7.8
	6.06	113.03	umc2170	108-118	6.35*	4.55	27.5
	8.05	71.2	bnlg1599	70-76	5.82*	-2.73	11.4

* QTLs with LOD value above the threshold obtained after 1000 permutations for a significance level = 0.05.

† QTLs with LOD value above the threshold obtained after 1000 permutations for a significance level = 0.1

‡ The QTL effect of the allele derived from B73o2

§ Percentage of the phenotypic variance explained by the QTL estimated by creating a MIM model with significant QTLs detected by CIM.

interactions explained 81.2% of the phenotypic variation in TEXT-F. Five of the six QTLs detected by CIM were present in the MIM model. The interaction among QTLs that explained more variation was that of bins 1.02 and 6.01with only 2.2%.

For the trait TEXT-L, nine QTLs showed significant LOD scores across all environments. These QTLs together explained 76.9% of the phenotypic variation (Fig. 3.1 and Table 3.1). In the QTLs of bins 3.02 (phi374119), 3.05 (umc1539), 3.07 (Txp196L), 5.05 (umc2026), 8.05 (umc1562) and 10.06 (umc1061), the allele increasing modification was the one derived from the QPM parent. In the QTLs of bins 1.06 (umc1035) and 6.06 (umc2170), the allele that decreased the proportion of opaque endosperm derived from the *o2* parent. The QTL on bin 6.06 near to the marker umc2170 explained 23.1% of the phenotypic variation, and the QTL on bin 8.05 near to the marker umc1562 explained 19.5%. A MIM model with nine QTLs and two QTL interactions explained 88.7% of the phenotypic variation observed for TEXT-L. Six of the QTLs detected by CIM were present in the MIM model. The interaction among the QTLs of bins 6.06 and 8.04 explained 3.2% of the variation.

The trait opacity (OPAC) was based on the observation that opaque kernels transmitted less light than the modified kernels. CIM analysis detected six significant QTLs for OPAC (Fig. 3.1 and Table 3.1). These QTLs explained 62.5% of the observed variation. The alleles that decreased opacity in the QTLs of bins 3.02 (phi374119), 3.05 (umc1539), 3.06 (phi102228), 5.05 (umc2111) and 8.05 (bnlg1599), were derived from the QPM parent. The QTL on bin 8.05 (bnlg1599) explained 17.3% of the variation, whereas the QTL on bin 3.05 (umc1539) explained 12.1%. The allele that decreased the

proportion of opaque endosperm in the QTL of bin 6.06 (umc1912) was derived from the *o2* parent B73*o2*. The QTL on bin 6.06 (umc1912) explained 13.4% of the observed variation. A MIM model with nine QTLs and two QTL interactions explained 88.9% of the phenotypic variation observed in OPAC. For five of the 6 QTLs detected by CIM, there was a correspondent QTL in the MIM model. The interaction that explained more observed variation was that of the QTLs in bins 1.00 and 6.06 with 3.8%.

The trait vitreousness (VITR) was measured through analysis of digital images of dissected endosperms that allowed a quantitative measure of the areas corresponding to soft and hard endosperm in representative kernels of each RIL. CIM analysis detected seven significant QTLs for VITR (Fig. 3.1 and Table 3.1). These QTLs explained 71.9% of the observed variation. In the QTLs of bins 3.02 (phi374121), 3.05 (umc1539), 4.06 (umc2027), 5.05 (bnlg1237) and 8.05 (bnlg1599) the alleles that increased the proportion of hard endosperm were derived from the QPM parent. The alleles that increased the proportion of hard endosperm in QTLs of bins 1.11 (phi064) and 6.06 (umc2170) were derived from the *o2* parent. The QTL of bin 3.05 (bnlg1539) explained 11.6% of the observed variation. The QTL on bin 6.6 (umc2170) explained 27.5% of the observed variation. A MIM model with five QTLs explained 70% of the observed variation in VITR. All the five QTLs in this model were also detected by CIM analysis.

Taking together the QTLs identified for the four endosperm texture traits (TEXT-F, TEXT-L, OPAC and VITR), there were five QTL clusters or “hot spots” regions where CIM analysis detected QTLs for at least three traits. There was a QTL cluster in chromosome 3S, between markers umc2258 and bnlg1019 (bins 3.03-3.04), with one

well-defined QTL for endosperm texture modification. There was a second region in chromosome 3L, between markers umc1973 and phi046 (bins 3.05-3.08), that contained at least two QTLs for endosperm modification. In chromosome 5L, between markers umc2111 and umc1722 (bins 5.04-5.05), there was another cluster. In chromosome 6L, between markers umc1912 and phi070 (bins 6.06-6.07), there was a cluster with one well-defined QTL peak for which the allele that contributes to texture modification came consistently from the *o2* parent. In chromosome 8L, between markers bnlg1176 and umc1161 (bins 8.05-8.06), there was another cluster for QTLs related to endosperm modification.

Amino Acid Contents

The inbred parents showed significant differences for the relative content of the essential amino acids lysine, tryptophan, and methionine. The *o2* parent showed higher levels than the QPM parent for all the three amino acids. Across environments, CIM analysis detected three significant QTLs for Lys representing 32.9% of the observed variation (Fig. 3.1 and Table 3.2). One QTL in bin 8.03 near to the marker bnlg1863, explaining 14.8% of the variation, had the favorable allele derived from the *o2* parent. Interestingly, the alleles that increased the content of Lys in the QTLs of bins 1.06 (umc1035) and 5.06 (phi087) were derived from the QPM parent. The percentage of the observed variation explained by these QTLs was 12 and 6.1%, respectively. A MIM model with five QTLs explained 45.2% of the phenotypic variation for Lys. The three QTLs detected by CIM were present in the MIM model.

For the content of Trp, CIM identified six QTLs (Fig. 3.1 and Table 3.2). These QTLs explained 49.1% of the observed variation. The alleles that increased the content of Trp in QTLs of bins 6.04 (umc1014), 7.02 (umc2142), 8.02 (umc1304), 8.04 (umc1460), and 10.04 (umc1678) were derived from the *o2* parent. In the QTL of bin 5.07 (bnlg118), the allele that increased the content of Trp was derived from the QPM parent. The QTL in bin 8.02 (umc1304) explained 9.5% of the variation, whereas the QTL in bin 10.04 (umc1678) explained 16.4% of the variation. A MIM model with 7 QTLs and 3 QTL interactions explained 64.8% of the observed variation in Trp content. Three of the six QTLs detected by CIM were included in this MIM model. The interaction between the QTLs of the bins 3.03 and 5.06 explained 5.6% of the observed variation.

Four QTLs were identified for Met content (Fig. 3.1 and Table 3.2). These QTLs together explained 57.3% of the observed variation. In all the QTLs, bins 2.08 (umc1516), 5.03 (bnlg1046), 7.02 (umc2142), and 8.02 (umc1304), the favorable alleles that increased the content of Met were derived from the *o2* parent B73o2. The QTL on

Table 3.2. Quantitative trait loci (QTLs) identified by composite interval mapping analysis across environments for the content of the aminoacids: lysine (Lys), tryptophan (Trp) and methionine (Met). QTL peaks or chromosome position with the highest LOD value, closest marker to the peak, and the bin of the marker according to the consensus maize map are shown.

Trait	Bin	QTL peak	Marker	QTL interval	LOD	QTL Effect‡	R2 (%) §
Lysine (Lys)	1.06	133.81	umc1035	129-144	3.33*	-1.53	12.0
	5.06	168.22	phi087	150-171	4.27*	-1.07	6.1
	8.03	38.14	bnlg1863	32-47	6.51*	1.60	14.8
Tryptophan (Trp)	5.07	211.17	bnlg118	197-211	2.98†	-0.28	5.0
	6.04	50.7	umc1014	36-57	2.20	0.27	4.7
	7.02	74.2	umc2142	65-78	3.86*	0.31	7.9
	8.02	33.59	umc1304	29-39	5.45*	0.31	9.5
	8.04	59.57	umc1460	52-71	4.16*	0.23	5.6
	10.04	52.57	umc1678	45-56	5.85*	0.51	16.4
Methionine (Met)	2.08	174.69	umc1516	168-182	3.82*	1.99	9.4
	5.03	67.32	bnlg1046	60-76	3.76*	2.04	8.9
	7.02	67.2	umc2142	63-70	8.91*	2.46	20.7
	8.02	32.59	umc1304	27-38	5.07*	2.44	18.3

* QTLs with LOD value above the threshold obtained after 1000 permutations for a significance level = 0.05.

† QTLs with LOD value above the threshold obtained after 1000 permutations for a significance level = 0.1

‡ The QTL effect of the allele derived from B73o2

§ Percentage of the phenotypic variance explained by the QTL estimated by creating a MIM model with significant QTLs detected by CIM.

bin 7.02 (umc2142) explained 20.7% of the observed variation. A MIM model with 5 QTLs explained 59% of the phenotypic variation observed for Met. All the four QTLs detected by CIM were included in this MIM model.

There were two QTL clusters detected by CIM where QTLs were located for at least two of the three amino acid-related traits. There was a QTL cluster in chromosome 7L close to the centromere, between markers umc1480 and mmc0411 (bin 7.02-7.03), that contains QTLs for Met and Trp. The allele that increases the proportion of these two amino acids derived from the *o2* parent. There was a second QTL cluster in chromosome 8S, close to the centromere, that encloses QTLs for all three amino acids. This region was located between markers umc1304 and umc1460 (bins 8.02-8.04) and the favorable allele derived from the *o2* parent.

Discussion

Several of the efforts to improve the nutritional value of maize take advantage of the effect of *o2* mutant alleles in reducing the proportion of the low quality zeins in the endosperm. In most cases, the genetic background of the *o2* genotypes must be carefully selected to ensure the presence of modifier loci that overcome the undesirable effects caused by the *o2* allele. The genetic mechanisms behind the modification of the endosperm are complex and involve additive, dominant and recessive gene actions, as well as paternal (i.e., xenia) and maternal cytoplasm effects (Lopes and Larkins, 1995; Lou et al, 2005; Vasal et al., 1980; Wessel-Beaver and Lambert, 1982; Wessel-Beaver et al., 1985).

One of the key steps for understanding the genetic mechanisms of endosperm modification is to locate and identify the genes involved. In this work, we identified five QTL clusters in chromosomes 3S, 3L, 5L, 6L and 8L controlling traits related to endosperm texture modification. The percentage of the phenotypic variation explained by the QTLs clusters together ranged between 62-68% depending on the trait. Each trait corresponded to a different approach to measure the variation present in the populations of RILs. Common QTLs were expected for these traits due to the high phenotypic and genetic correlations between them. These QTL clusters were detected across environments in a multi-environment experimental design that included data from two locations and three years. The five QTL clusters were confirmed by novel MIM models. MIM analysis has the advantage over other QTL mapping procedures of using a multiple-QTL model in which the effects of all putative QTL are analyzed simultaneously so that certain epistatic interactions can be included (Kao, 1999; Zeng, 2000). MIM provides better precision and more power than single-QTL mapping methods, particularly in parameter estimation (Zeng, 2000). However, we are aware that because of the evaluation of a population of bigger size was not feasible; the risk of underestimation of QTL number and overestimation of QTL effects should be considered (Beavis, 1998; Melchinger, 1998).

Our results agree with previous works that have suggested that the modification of *o2* endosperm is quantitative with polygenic control (Annapurna and Reddy, 1971; Motto et al., 1978;Wessel-Beaver and Lambert, 1982; Lou et al., 2005). Other researchers have suggested that the modification is caused by one or few genes

(Pollacsek, 1970; Pollacsek et al., 1972; Lopes and Larkins, 1995). The quality and composition of the endosperm reflects the variation present in maize germplasm. Some of the oldest and most used systems to classify maize germplasm are based on endosperm texture and composition (Sturtenvant, 1899). Therefore, it is not surprising that modification of the *o2* endosperm is strongly dependent on the genetic background and that different extents of modification can be obtained by a range of genetic mechanisms. It is also plausible that the genes that are necessary to recover a desired ratio of hard-to-soft endosperm are few, but there are other genes responsible for fine tuning the phenotype.

In four of the clusters the source of the "modifier" QTL allele, which increased the proportion of hard endosperm, derived from the QPM parent. But, in chromosome 6L, there was a major effect QTL cluster in which the favorable allele originated from the *o2* parent. It is possible that B73 contains alleles that increase the proportion of hard endosperm, but when the *o2* recessive allele was introgressed, the phenotype caused by these genes was masked. Some endosperm-related genes mapped to this chromosome region are *starch synthase IIa* (ssIIa) or *Sugary2* (*Su2*) (bin ~6.04), *Opaque-14* (*O14*) (bin ~6.04), and *Pyruvate orthophosphate dikinase1* (*Ppdk1*) (bin 6.05) (Lawrence et al., 2004; Lawrence et al., 2007). It has been proposed that dent corn is a derivative of flint-flour crosses, and thus variation of the ratio of soft to hard endosperm is expected among dent lines (Hamilton 1951, Vyn and Tollenaar, 1998; Sene et al., 2001). Different genetic backgrounds caused by recombination with the QPM parent allowed the identification of this QTL derived from the dent parent in the population of RILs.

Previous research has aimed to map the location of the modifier genes for the gene *o2* too. Lopes et al. (2005) used restriction length polymorphism (RFLP) analysis to analyze bulks of DNA from opaque and modified F_2 individuals selected from two populations. They reported one RFLP probe (npi277), on chromosome 7 that was linked to kernel modification in one population and another probe (umc35), also in chromosome 7, linked to modification in a second F_2 population. The probe npi277 maps to bin 7.01-7.02, near several loci related to accumulation of zein proteins such as *zp27* (27-kDa zein protein), *de*-B30* (defective endosperm B30), and *o5* (*opaque endosperm 5*). The probe umc35 maps to the 7.05 and it has been suggested that it corresponds to the locus *o15* (*opaque endosperm 15)*, a mutation that reduces the synthesis of γ-zeins in the endosperm (Dannenhofer, 1995). Lizarraga-Guerra et al. (2006) reported the use of BSA for mapping markers linked to endosperm modification in a population derived from a cross of a South African QPM and an opaque inbred. They found two loci (bins 7.02 and 9.02) linked to the modified phenotypes. In the population of RILs derived from B73o2 and CML161, we did not find evidence of QTLs in these genome regions. SMA suggested significant associations between markers in chromosome 7S and the traits TEXT-F, TEXT-L, OPAC and VITR, but these associations were not picked as QTLs by CIM or MIM (data not shown). Although not mapping modifiers for *o2*, Sene et al. (2001) mapped QTLs for several kernel characteristics in a population of RILs derived from a cross between a flint line (F-2) and a semident line (Io). They identified three main effects QTLs for vitreousness in bins

2.03, 5.02 and 6.06. Interestingly, as reported here, the favorable allele for the QTL detected in the bin 6.06 originated from the dent parent and not from the flint one.

The lack of common QTLs with previous mapping efforts for *o2* modifier genes can be explained by differences in the genetic background of the populations, or by sampling effects associated with QTL mapping in populations of limited size (Beavis, 1998; Openshaw and Frascaroli, 1998). However, the use of a population of RILs with homozygous triploid genome offers several advantages over other experimental populations (e.g. F_2) used to map modifier genes in the past (Knapp and Bridges, 1990; Burr and Burr, 1991). In a population of RILs no dominance effects are expected, and the additional rounds of recombination in the development of RILs generate better map resolution of the detected QTLs (Lee et al., 2002). Moreover, we evaluated six environments with replicates, which allowed us to estimate least square means and BLUPs to exert control over the non-genetic variation. Finally, the use of different methodologies to evaluate endosperm modification allowed us to avoid confounding effects caused by the person evaluating or the measurement technique.

In *o2* genotypes the content of the limiting amino acids lysine and trypthopan are increased, together with histidine, aspartate, asparagine and glycine. On the contrary, glutamate, glutamine, alanine and leucine contents tend to decrease; whereas methionine levels seem to vary independently (Glover and Mertz, 1987). Contradictory reports have described both positive and negative correlations for endosperm modification and amino acid contents (Paez, 1969; Robutti et al., 1974a; Pixley and Bjarnason, 2002; Bantte and Prasanna, 2004). The effects of the endosperm texture modification on protein quality

seem to be greatly affected by the genetic background (Bantte and Prasanna, 2004). In general, modified *o2* endosperms have lower proportion of Lys than *o2* endosperms, but the increase in protein content compensates the loss (Robutti et al., 1974a). Modified *o2* genotypes have been found having more tryptophan and less methionine than normal inbreds (Bantte and Prasanna, 2004; Scott et al., 2004). Lysine and Methionine are synthesized in the same metabolic pathway, having aspartate as their common precursor (Azevedo et al., 2006). Tryptophan is synthesized from chorismate in a pathway that also provides precursors for important secondary metabolites like hormones and phytoalexins (Radwanski and Last, 1995). Though the metabolic pathways for the biosynthesis of these amino acids are well understood (Gaziola et al., 1999; Galili et al., 2002; Azevedo et al., 2006), literature regarding the elucidation of genetic control of variation of amino acid contents in maize is scarce. Wang and Larkins (2001) using IM detected QTLs in chromosomes 2L, 2S, 3S and 7L controlling free amino acid (FAA) contents in a F_2 population derived from two *o2* lines differing in FAA contents. Later, triploid specific models for QTL mapping suggested 11 QTLs in all ten chromosomes for FAA contents in the same population (Wu et al., 2002). Further research suggests that *Ask2*, which encodes for an aspartate kinase sensitive to lysine inhibition, could be part of the QTL in chromosome 2L that affects FAA content (Wang et al., 2001). We found significant differences between the inbred parents B73*o2* and CML161 for Lys, Trp and Met content, with the *o2* inbred being superior for all three amino acids. Heritability estimates for the content of these three amino acids were considerably high as reported in chapter II. Consequently, QTLs for the contents of Lys, Trp and Met were

detected in the population of RILs. Two QTL clusters were detected in chromosomes 7L and 8S and major effect QTLs in several other chromosomes, with favorable alleles increasing the amount of amino acids in the endosperm derived from both parents. QTLs controlling levels of Lys, Trp and Met were found to co-localize on 8S. Because these three amino acids are synthesized by two different pathways, it is possible that these QTLs control total protein content. A QTL controlling grain protein content was identified at a similar position in a previous study (Melchinger et al., 1998)

Possible candidates for genes controlling amino acid levels are genes involved in the biosynthetic pathways of the amino acids studied. For example, four QTLs controlling methionine content were identified. Those at bins 2.08 and 5.03 map close to the positions of BAC contigs (2.05 and 5.02) that are detected by hybridization using a probe corresponding to methionine synthase (Gardiner et al, 2004; Lawrence et al., 2007; Wei et al., 2007). Similarly, an overgo probe contained in a putative anthranilate synthase gene hybridizes to BACs in bin 10.04, co-localizing with the QTL explaining the most phenotypic variation in our study (Gardiner et al, 2004; Lawrence et al., 2007; Wei et al., 2007).

It was of interest to us to examine if QTLs for endosperm texture modification and amino acid contents co-localized. Co-localization of QTLs for two traits might reflect a pleiotropic effect of a single gene, or tight linkage of genes controlling the traits independently (Sene et al., 2001). Several reports, some of them very soon after the discovery of the enhancement on protein quality by the *o2* allele, mention that there is not a strict relationship between protein quality and endosperm texture (Paez, 1969;

Dudley et al., 1972; Pradilla et al., 1972; Moro et al., 1995). Nonetheless, a general negative correlation between endosperm texture and protein quality creates the requirement for simultaneous selection for texture and amino acid contents in the development of the QPMs (Bjarnason and Vasal, 1992; Vietmeyer, 2000; Prassana et al., 2001). We observed that QTLs explaining variation for endosperm texture and amino acids (Lys, Trp and Met) content did not overlap along the genetic map (Fig. 3.1). Nevertheless, it is noteworthy that a cluster that contains QTLs for the three amino acids and a cluster that contains QTLs for the four endosperm texture traits are in close proximity on chromosome 8 (Fig. 3.1). Some endosperm-related genes located in this genomic regions are the *b32 ribosome inactivating protein* (*b32/RiP*) (bin 8.03), glyceraldehyde-*3-phosphate dehydrogenase1* (*gpa1*) (bin 8.03), *Proline Responding 1* (*Pro1*) or *Opaque-6* (*O6*) (bin 8.04), *Floury-3* (*Fl3*) (bin 8.04), *pyruvate orthophosphate dikinase-2* (*ppdk2*) (bin 8.04), and *Opaque-16* (*O16*) (bin 8.05). The two QTL clusters were well defined by both CIM and MIM analysis, appear to be on different arms of the chromosome 8 and are linked in repulsion. The recombination fraction between the closest markers to the QTL peaks, bnlg1863 and umc1562, is r= 0.276. This observation might have breeding implications since selection for the favorable allele controlling endosperm texture could drag the unfavorable allele for amino acid content and reduce the amount of Lys, Trp and Met by at least 10%. Additional studies would be needed to establish the prevalence of this linkage drag effect in other populations.

The information generated here can readily be utilized to develop a modified B73*o2* line carrying the most important favorable QTL alleles derived from CML161

(i.e. texture QTLs 3S, 3L, 5L, and 8L.) and retain the favorable QTL alleles derived from B73*o2* (Texture QTL 6L and amino acid QTLs 7L and 8s). By using SSR markers flanking the QTL regions as markers for selection, maize breeders would ensure the selection of the favorable QTLs, and decrease the risk of linkage drag that can diminish its nutritional value.

To our knowledge, this is the first time that both endosperm texture and amino acid contents have been exhaustively evaluated in a multi-environment experiment with a population derived from *o2* x QPM materials. Despite the fact that QTLs identified here do not co-locate with those identified in previous works and therefore did not provide clear consensus regions for the location of modifiers for the gene *O2*, we are confident that the complexity of the endosperm modification can be unveiled with additional studies. For instance, the development of near isogenic lines (NILs) using an *o2* genotype (i.e. B73*o2*) as recurrent parent and the QPM CML161 as the donor genome will establish the effects of the individual QTLs for endosperm texture of chromosomes 3S, 3L, 5L, and 8L. A similar approach can be used to establish the effects of the QTL cluster in 6L using B73*o2* as the donor parent. The development of NILs will contribute to improved resolution of the QTL regions, and facilitate the identification of candidate genes. Moreover, results of a complementary expression QTL (eQTL) mapping study to analyze genetic control of mRNA abundance of endosperm transcripts (e.g. zein genes, enzymes in the starch biosynthesis, elongation factor 1A, etc.) suggest a role for differential regulation of some endosperm genes in the modification of *o2* endosperm. In addition, multi-population studies that include the evaluation of

endosperm traits will test the prevalence of the role of these genomic regions in the composition and structure of the maize grain in general (Buckler, 2006).

Conclusion

After the enthusiasm in the late 60's and early 70's for the use of the *o2* alleles in breeding programs, there was a general disappointment when the undesirable traits associated with the mutation were identified. However, the discovery of the effect of the *o2* in protein quality spurred research on the biochemistry, physiology, microstructure, and genetics of endosperm of maize and other crops like sorghum, rice and wheat. The development of the QPMs has demonstrated that positives responses from selection can be observed for both endosperm texture and amino acid contents with simultaneous selection schemes. So far, it seems that the modification of the endosperm in *o2* maize is rather complex with several genetic factors playing roles for both endosperm texture and amino acid contents, with the genetic background being one of the most important variables contributing to the variation observed so far. The broader applicability of the accumulated knowledge to more plant breeding programs depends on the understanding of how the *Opaque2* gene interacts with the rest of the genome in different genetic backgrounds to produce different phenotypes. In the mean time, a more practical approach using inbred lines relevant to breeding programs to develop QTL mapping populations takes advantage of both marker-based genetic analysis and more efficient evaluation methodologies for the development of enhanced germplasm.

CHAPTER IV

IDENTIFICATION OF MODIFIER LOCI FOR THE GENE *opaque2* BY EXPRESSION QUANTITATIVE TRAIT LOCI MAPPING

Overview

The main function of the maize endosperm is to accumulate reserves in form of starch, storage proteins and lipids to be used during germination and seedling development. The recessive allele *opaque2* (*o2*) produces a severe phenotype in the endosperm of maize enhancing its protein quality, but decreasing its agronomical value. The opaque phenotype can be altered to resemble normal endosperms by the action of modifier genes. The modification of the opaque phenotype has been studied for almost five decades without conclusive evidence on its genetic control. The modification of the *o2* phenotype is important for agriculture, industry and nutrition. The product of the *O2* gene is a transcription factor that affects the expression of a number of endosperm genes. We used a well-characterized population of recombinant inbred lines derived from an *opaque2* and a modified inbred line to identify genomic regions controlling gene transcript abundance in developing endosperm. We selected a group of 29 endosperm genes associated with the O2 activity, including members of zein protein family, enzymes of the starch pathway, the b32/RiP protein, a pyruvate orthophosphate dikinase, and the lysine-ketoglutarate reductase. We provide evidence that support the presence in the population of recombinant inbred lines of QTL hot spots that explain variation for endosperm texture modification or amino acid contents and are associated with the regulation of the expression of a group of endosperm genes. These results support an

interaction in the regulation of both starch and zein biosynthesis, and a role for gene products under regulation of O2 in the modification of opaque endosperm.

Introduction

Domestication and selection for 8,000 years have developed modern maize or corn. This important crop is cultivated in hundreds of countries worldwide and it is considered a staple food for many of them. Maize is also a major crop for livestock feed and has several industrial applications, including the biofuels industry. The mature grain of maize is composed mainly by starch (70-75%), protein (8-10%) and lipids (4-5%). The endosperm is the largest tissue of the grain and constitutes approximately 80% of the mature kernel dry weight (Watson, 2003). The structure and composition of the endosperm are important traits for maize producers and breeders. Some of the oldest and most often used system to classify maize germplasm is based on endosperm characteristics (Sturtenvant, 1899). The most common types of maize include flint, flour, dent, pop, sweet, and waxy. The physical appearance of each kernel type is determined by the quantity and quality of endosperm constituents. The majority of the maize grown in the U.S. belongs to the dent type, while flint maize is grown in South America, some parts of Africa and Europe (Watson, 2003).

The endosperm of a typical dent kernel is composed by two easily distinguishable portions, the soft or opaque endosperm with a very floury texture, and the hard or vitreous endosperm with a more glassy appearance. Several mutations have been identified affecting the structure and composition of the maize endosperm (for

review, see Coleman et al., 1997). Among them, the recessive mutations that increase the proportion of soft endosperm are known as *opaque* (*o*) mutations. The opacity of the kernel is produced by light refraction caused by gaps in between the starch granules in the endosperm. In normal maize, the starch granules are tightly packed with a continuous protein matrix and protein bodies that by filling these gaps prevent the refraction of light (Duvick, 1961). The protein bodies in maize endosperm are compartments derived from the endoplasmic reticulum (ER) in which the prolamin (zeins) proteins are deposited after synthesis. The zeins constitute 40-70% of the endosperm proteins in maize (Lawton and Wilson, 1987; Soave and Salamini, 1984), and about 50% of the total proteins in mature grains (Soave et al., 1981). The zein proteins are often classified by differential solubility and mobility through gels as: α-zein, β-zein, δ-zein, and γ-zein (Coleman et al., 1997). In normal maize, the protein bodies show a maturation gradient both in developmental time and location in the different layers of the endosperm (Khoo and Wolf, 1970; Lending and Larkins 1989). More mature protein bodies are located at the interior of the endosperm, larger in size, richer in α-zeins and poorer in 27 kDa γ-zein (Lending and Larkins 1989). Some evidence suggest that both the soft portion of dent maize and the whole soft *o2* endosperm are the result of an early arrest in the development towards a more physiologically mature vitreous endosperm with larger protein bodies (Dombrink-Kurtzman and Bietz, 1993; Landry et al., 2004). A more vitreous kernel, or at least a vitreous endosperm coat, would be needed to cope with challenges that the seed endures (e.g. insects, fungal infections, etc.) between seed maturation and germination.

One of the *opaque* mutations, the *opaque2* (*o2*), have received vast attention after Mertz et al. (1964) discovered that *o2* kernels contained more of the essential amino acids lysine (Lys) and tryptophan (Trp). This enhancement in nutritional quality of opaque mutants is associated to the reduction of zein proteins that are particularly deficient in Lys and Trp, and are the primary reason for the poor protein quality of maize (Vasal, 2000).

Additional research suggests that the phenotype of *o2* materials is affected by their genetic background, and therefore it is possible to obtain *o2* kernels resembling those of *O2* wild-type maize. This modification of the *o2* endosperm is controlled by the action of modifier genes of the *O2* gene, and has allowed the development of the Quality Protein Maize (QPM) (for review see, Bjarnason and Vasal, 1992; Prassana et al., 2001). QPMs are lines or cultivars homozygous for *o2* mutant alleles with enhanced nutritional value, with dent- or flint-like endosperms, and better agronomical performance than *o2* materials.

The genome of maize contains approximately 50,000-60,000 genes (Messing et al., 2004; Liu et al., 2007) and approximately 11-20% of them are transcribed in the endosperm (Lai et al., 2004; Verza et al., 2005). The *O2* is an endosperm gene that encodes a bZIP class transcription factor (Schmidt et al., 1990) that induces the transcription of a group of α-zeins (Schmidt et al., 1992) and influences expression of other genes such as *b-32* (a type I ribosome inactivating protein) (Bass et al., 1992) and *cyPPDK1 (*a cytosolic pyruvate orthophosphate dikinase) (Maddaloni et al., 1996). Global analysis of gene expression has been used to establish a strong pleiotropic effect

of the *o2* and other opaque mutants in the overall gene expression of endosperm (Hunter et al., 2002; Jia et al., 2007).

Advances in the development of experimental populations, molecular marker technologies, and statistical tools have proven to be important in unveiling the genetic basis of quantitative traits by dissecting complex traits into Mendelian quantitative trait loci (QTLs) (for review, see Asins, 2002; Salvi and Tuberosa, 2005). In addition, methodological advances such as microarray analysis or quantitative real-time RT-PCR (qRT-PCR) allow measuring the levels of expression of many genes in every individual of an experimental population (e.g. population of RILs). The gene expression levels can be treated as phenotypes and analyzed as quantitative traits to map expression QTLs (eQTLs) (Jia and Xu, 2007; Lan et al., 2003). Such strategies have been named 'genetical genomics' (Jansen and Nap, 2001). Schadt et al. (2003) published one of the first studies of microarrays-based eQTL mapping and included maize as one of their model organisms. Genetical genomics has several practical applications in biomedical and agricultural research, as well as in other aspects of basic biological research (Gibson and Weir, 2005).

Here, we used eQTL mapping analysis to identify genomic regions controlling variation in levels of mRNA abundance of genes associated with the modification of *o2* endosperm, and to identify eQTLs that colocalize with QTLs mapped for endosperm modification and essential amino acid contents. This approach may corroborate the identity of candidate genes or provide new candidate genes for modifiers for the gene *O2*.

Materials and Methods

Plant Material

The inbred lines B73, B73*o2* and CML161 were grown in a greenhouse (IPGB-Borlaug Center, Texas A&M University, College Station, TX) during winter 2005. Irrigation, fertilization and conventional greenhouse practices were carried out to ensure uniformity and homogeneity during growth. Each plant was self-pollinated by hand. Four ears per inbred line were collected 18 days after pollination (dap). The 93 recombinant inbred lines (RILs) evaluated were a subset of the population of RILs derived from B73*o2* and CML161 previously used to map QTLs associated with endosperm modification and amino acid contents (Chapter III). The RILs were planted in summer 2006 in the Texas Agricultural Experiment Station (latitude $30^{o}37$, elevation 96 m.a.s.l.) in Burleson County that has a humid subtropical climate and 99.3 cm of annual precipitation on average. Each RIL was planted in a single row and at least 10 plants per row were self-pollinated by hand. Three ears per RIL were collected 18 dap. Ears from the greenhouse and the field were harvested and immediately frozen in liquid nitrogen before storage at -80^{o}C. Endosperms were manually dissected from the embryo and pericarp. Three dissected endosperms from a single ear were pooled and treated as one sample (experimental unit).

Protein Extraction, Quantification and MS Analysis

Equal amounts of flour (120μg) of mature endosperm of the inbred lines B73, B73*o2* and CML161were defatted with 1 mL acetone (with agitation for 30 min) and air-dried. Protein fractions were extracted as described by Landry and Moureaux (1970), except that 70% ethanol was used instead of 55% isopropanol. Protein concentrations of Landry-Moureaux (LM) fractions II and III were estimated with the BCA (bicinchoninic acid) protein assay (Pierce, Rockford, IL), and of fractions I, IV, and V with the Quick Start Bradford protein assay (Bio-Rad, Hercules, CA), using Bovine Serum Albumin (BSA) as standard. Protein fractions were separated electrophoretically using precast Sodium dodecyl sulfate-polyacrilamide gel electrophoresis (SDS-PAGE) with 10-20% gradient gels (Bio-Rad, Hercules, CA) and visualized by Coomassie Brilliant Blue R250 staining. SigmaMarker low range weight (Sigma-Aldrich, St. Louis, MO) and prestained SDS-PAGE broad range (Bio-Rad, Hercules, CA) standards were loaded as mass references.

Selected protein bands were excised from stained SDS-PAGE gels, and digested with trypsin at 37˚C for 4 hrs following standard protocols. The sample was desalted using ZipTips (Millipore, Bedford, MA), and then combined 1:1 with MALDI matrix (5 mg mL^{-1} CHCA prepared in 50:50 ACN: ddH_20 containing 10-mM AP and 0.1 % TFA). MALDI-MS experiments were performed using a 4700 Proteomics Analyzer. MALDI-TOF/TOF (Applied Biosystems, Foster City, CA). PMF and MS/MS data were searched against the Swiss-Prot protein sequence database using the GPS Explorer (Applied Biosystems) software. For both the PMF and MS/MS searches, proteins were matched

with the identified peptides and each protein was assigned a Mascot score, which was a probability-based Mowse score (Perkins et al., 1999).

RNA Extraction

Total RNA was extracted from dissected endosperms using an adaptation of a method based on an SDS-based buffer and Trizol (Invitrogen, Carlsbad, CA) (Li and Trick, 2005). RNA was quantified in the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Rockland, DE). Additional purification of RNA was done for microarray samples using RNeasy MinElute Columns (Qiagen, Valencia, CA), followed by analysis in the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA).

Microarray Analysis

Global analysis of gene expression was done for inbreds B73, B73*o2* and CML161 using arrays from the Maize Oligonucleotide Array Project (http://www.maizearray.org). The array version used in this study contained ~57,000 array elements or features derived from automated printing of 70-mer oligomers. The Amino Allyl MessageAmp II aRNA Amplification kit (Ambion, Austin, TX) was used for labeling of the RNA probes using the fluorescent dyes Cyanine-3 (Cy3) and Cyanine-5 (Cy5). Labeled RNA was used to hybridize microarray slides following an interconnecting loop design (four hybridizations for each line: two with Cy3 and two with Cy5 labeling) (Fig. 4.1). The entire experiment was replicated twice, once as a hybridization service offered by the Maize Oligonucleotide Array Project

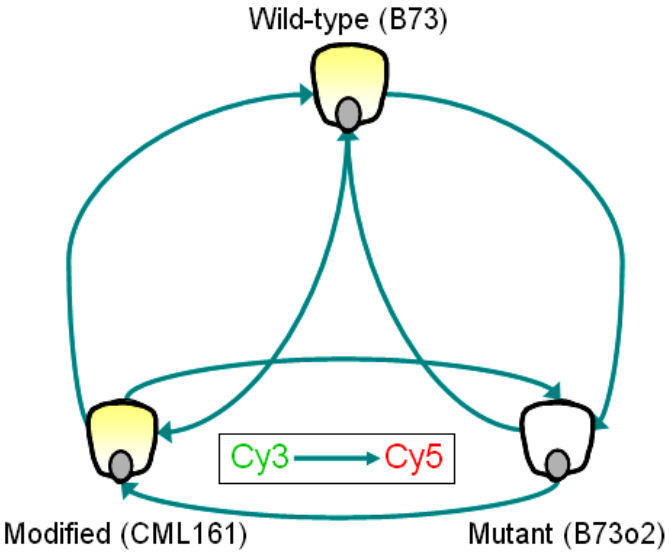

Figure 4.1. Loop design for endosperm microarray hybridizations. Each solid arrow represents 2 replicates (4 comparisons for each pair). Arrowheads indicate samples labeled with Cy5 and arrow bases samples labeled with Cy3.

(http://www.maizearray.org). The hybridization process followed protocols that have been optimized for use with cRNA targets and 70-mer oligonucleotide features (http://www.maizearray.org).

Arrays hybridized at Texas A&M University were imaged with the Genepix 4100A scanner (Molecular Devices, Union City, CA) and the GenePix Pro 6.0 program (Molecular Devices, Union City, CA). Filtering was performed in Microsoft Excel (Microsoft, Redmond, WA). Background correction, within slides normalization (loess) and between slides normalization (quantile normalization) was done with the LIMMA package (Smyth, 2004) of R 2.3.1 (R Development Core Team, 2006). Analysis of variation and contrast analysis was performed with the LIMMA package that uses linear

models and empirical Bayes analysis (Smyth, 2004). To correct for multiple testing, p values were adjusted using a false discovery rate (Benjamini and Hochberg, 1995), at a q value cutoff < 0.05.

Quantitative Real Time PCR

The mRNA abundance in 18 dap endosperm of selected genes was estimated using real-time qRT-PCR. Gene-specific primers were designed using Primer Express 3.0 software (Applied Biosystems, Foster City, CA, USA) with melting temperatures (Tm) of 58–60°C, GC content of 30-80%, primer lengths of 16–26 bp, and amplicon lengths of 90–150 bp (Table 4.1). Primers were evaluated for specificity by analyzing the dissociation curves, and by loading PCR reactions in agarose gels. Primer efficiency was estimated using serial dilutions of B73 cDNA. Forward and reverse primer concentrations were adjusted until primer efficiency was 90-100 %. Primers were redesigned if concentration adjustments did not increase amplification efficiency.

First-strand cDNA was synthesized from 1µg of total RNA in 96 well-plates using the iScript cDNA synthesis kit (Bio-Rad, Hercules, CA). The obtained cDNA was diluted 100-fold. Quantitative PCR reactions were performed in 384-well plates in an ABI PRISM 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA, USA) using SYBR Green to detect dsDNA synthesis. Reactions were carried out in a 10µl volume containing 100-900 nM of each primer, 2 µl of diluted cDNA, and 5 uL 2× Power SYBR Green Master Mix Reagent (Applied Biosystems). For each selected gene, three independent samples were processed, and parallel reactions were set up with

Table 4.1. Gene transcripts and primers selected for quantitative real-time PCR.

Name	Transcript	Accession	Primer Sequences	Category	References
αGL18	α-globulin 18kD	AF371278	TGCAAGACGCAGCATGTGTA TGTAACGAACACGCGACAATATT	Storage proteins	Woo et al., 2001
αZN19B1	α-zein 19 kDa B1	AF371269	AGCTGGTGTGAGCCCTGCTA TGTAAGAGGGTGCCAGCGTTA	Storage proteins	Woo et al., 2001; Song and Messing, 2002
αZN19B3	α-zein 19 kDa B3	AF371271	ACCCCGCAACCCTCTTACA GGGCACCACCAATGATGTG	Storage proteins	Woo et al., 2001; Song and Messing, 2002
αZN19B4	α-zein 19 kDa B4	AF371273	CCCCGCAACCCTCTTACAA AGGGCACCGCCAATGAT	Storage proteins	Woo et al., 2001; Song and Messing, 2002
αZN22Z1	α-zein 22 kDa Z1	AF371274	TCAACCAACTGGCTGTGTCAA GGTTAGCCACTGCCAATGGAT	Storage proteins	Woo et al., 2001; Song et al., 2001
αZN22Z3	α-zein 22 kDa Z3	AF371275	GAGGCAGCAACCCATCGTT GGAAACACATGCCGGAATG	Storage proteins	Woo et al., 2001; Song et al., 2001
αZN22Z4	α-zein 22 kDa Z4	AF371276	TTCATCTAGCCCGCTCGCTAT AGCTGAGTAAGAGCTGGTACAATTTG	Storage proteins	Woo et al., 2001; Song et al., 2001
αZN22Z5	α-zein 22 kDa Z5	AF371277	TCTTTGATGAACCTTGCCTTGTC GGTATGAGGGCAATGTTATCAAGTAC	Storage proteins	Woo et al., 2001; Song et al., 2001
βZN15	β-zein 15 kDa BZ15	AF371264	GCCGAGCGCCATGTTC GGCACGCTCAAGCTGCAT	Storage proteins	Woo et al., 2001; Cord-Neto et al., 1995
PPDK	cytosolic pyruvate orthophosphate dikinase1	AI795369	GCTGGTGGATGTGGTGATCA GGAGATGGGATTGTAGCTTTTATGAG	Carbon Partition/ Amino acid synthesis	Maddalonni et al., 1996; Mechin et al., 2007
δZN10	δ-zein 10 kDa DZ10	AF371266	TCTTTGTGCAAGCGCCACTA TACTGCATGCATGGGTTCATG	Storage proteins	Woo et al., 2001
δZN18	δ-zein 18 kDa DZ18	AF371265	TACTCTGATTCCATCTCGCACATT TGCTGTAAAAACATGGGTGAGATC	Storage proteins	Woo et al., 2001

Table 4.1. Continued.

Name	Transcript	Accession	Primer Sequences	Category	References
EF1Ac	Elongation factor 1-alpha c	AF136825	ACTGGAGCCAAGGTGACGAA CCACGAGGTAACCAACAAGCA	Protein Synthesis	Berberich et al., 1995; Abe et al., 2003
EF1Ab	Elongation factor 1-alpha d	AF136826	GCCTGGATAACTTGCCACTAAAG CTGCGCTACCCAAACAAACA	Protein Synthesis	Carneiro et al., 1999
γZN16	γ-zein 16 kDa GZ16	AF371262	GGCGGTGTCTACTACTGAGGAAAC ACGTTACTCATCGCCGCTAAA	Storage proteins	Woo et al., 2001
γZN27	γ-zein 27 kDa GZ27	AF371261	GGTGTCCCCCACTGAAGAAAC CAATCATCGCCGCTAAATGAC	Storage proteins	Woo et al., 2001
γZN50	γ-zein 50 kDa GZ50	AF371263	TGGTAATGCCTTTTCTCCAATCA GGCTCAATCTGCCTAAGATCATG	Storage proteins	Woo et al., 2001
idbeI	Isoamylase debranching enzyme (Sugary1)	CO454239	CGTGCTGTCACCGTCTACCA GCGCAATACAAGGATGATGGA	Starch synthesis	James et al., 1995
Leg50k	Legumin 50 kDa	AF371279	CCGTCCTCAACCTCGTACAGA TGTGGGCGTTGAAGTTCCA	Storage proteins	Woo et al., 2001; Yamagata et al 2003
LKR/SDH	Lysine-ketoglutarate reductase/ saccharopine dehydrogenase	AF003551	TGGCCACAATGAGCAAGACA GCGTCGCTTGGTGCTTTT	amino acid synthesis	Arruda et al., 2000; Gaziola et al., 1999
O2	Opaque2	X16618	GACGCTAACGTCGACAACA CATCTCTATCACCCGCTTCAG	Transcription regulation	Hartings et al., 1989
b32/RiP	Ribosome Inactivating Protein/ b-32 (32 kDa albumin; RIP3:1)	CD448545	GCGTACGCCGGGTGTATTTA ACGAGGCCTGATCAATGCA	Translation Inhibition/ Defense reactions	Lohmer et al.,1991; Bass et al.,1992
sbeI	Starch branching enzyme I	CA401712	CGGTCCTTTCCTGTAGCTTG CATCCATCCATCCCTGATTA	Starch synthesis	Kim et al., 1998
ssIIa	starch synthase IIa (Sugary2)	AF019296	TGACGGTCTCGGTGAAGAACT TGACACAGCTCAATGCAATCC	Starch synthesis	Harn et al., 1998, Zhang et al., 2004

Table 4.1. Continued.

Name	Transcript	Accession	Primer Sequences	Category	References
b70/BiP	heat shock b70/biP	M59449	CGACCACGACGAGCTGTATATG GCCGCTAAACACCGTTACAAA	Protein folding/ Protein assembly	Fontes et al., 1991
PFP	Pyrophosphate:fructose 6-phosphate 1-phosphotransferase (a subunit)	MZ00015645	GGTAAGCGCCATGGCTTCT CAGCTTGGCCTTCAGATGGT	Starch synthesis/ Sucrose breakdown	Mahajan and Singh, 1989
RAPB	RAPB protein (Oryza sativa)	MZ00041923	CGGCTCTCGCTCGTCTGA GTAGAACAGAACGACCGCCAAT	Transcription regulation	This work
WRKY	WRKY1	MZ00015228	GCGGAGGGTCAAGATTGGA ATAGACAGCCCGAAGCGAAA	Transcription regulation	This work
MZ41626	Hypothetical Protein (Oryza sativa)	MZ00041626	TGGATGAGTAAGCTGCTGAGATAGAC ATCTCAACGGAAGCCAACAAA	unknown	This work

primers for the 18s ribosomal gene for normalization. Reaction conditions were 95°C for 10 min, 47 cycles of 95°C for 15 s, and 60°C for 1 min. Data were analyzed using the SDS 2.2 software (Applied Biosystems). Results from the SDS software were exported to Microsoft Excel (Microsoft, Redmond, WA) for further analyses. The cycle number at which the abundance of accumulated PCR product crosses a specific threshold was registered as the threshold cycle (Ct) for each reaction. The difference in Ct values (ΔCt) between a specific mRNA and the normalizer 18s ribosomal gene was calculated individually for each sample of each RIL. The ΔCT values, which are comparable to the log-transformed, normalized mRNA abundance, were used as the phenotype of the RILs in subsequent analysis (Lan et al., 2003).

Analysis of Variance and Cluster Analysis

Analysis of variance (ANOVA) was conducted separately for each expression trait (as ΔCt values), considering all effects in the statistical model as random, in PROC GLM of SAS 9.1 (SAS Institute, 2003). Genotype least square means were estimated for each gene transcript. Broad sense repeatabilities were calculated using variance components from the ANOVA.

The ΔCt values for each gene and each RIL were standardized to mean 0 and variance 1. The relationship between genes and RILs was assessed by Single Value Decomposition (SVD) of standardized variables using Biplot v1.1 (Lipkovich and Smith, 2002). Data were presented as a biplot with two principal components. Hierarchical clustering with oblique principal components was performed using PROC

VARCLUS (SAS Institute, 2003; Lan et al., 2003). Final clusters were arbitrarily designated by a proportion of the variance explained > 0.8.

Expression QTL Mapping

A linkage map constructed previously for the population of RILs was used for the eQTL analysis (Chapter III). The map consisted of 180 Simple Sequence Repeats (SSRs) well distributed across the genome. The composite interval mapping (CIM) application of Windows QTL cartographer (Wang et al., 2006) was used to map eQTLs for individual traits. Multiple trait composite interval mapping (MT-CIM) analysis was used to map eQTLs for groups of variables clustering together (Wang et al., 2006). Both CIM and MT-CIM were implemented using the forward and backward regression method (probability in = 0.1, probability out = 0.1), 5 cM window size, and 1 cM walk speed. A LOD score threshold of 2.5 was applied to declare the presence of an eQTL.

Results

Previously, we had developed a population of RILs derived from a cross *o2* x QPM (B73o2 xCML161). This population segregates for both endosperm modification and essential amino acid contents, and QTLs have been mapped for both traits (Chapter II; Chapter III).

Endosperm Proteins Fractions

Initially, we analyzed storage proteins in mature endosperm in the parental lines B73, B73*o2* and CML161. Protein extracts from equal amounts of mature endosperm flour of the parental inbred lines were compared following the method developed by Landry and Moreaux (LM) (1970). The LM method produces five soluble fractions: F-I (saline-soluble; albumin-globulin), F-II (alcohol-soluble; α-zeins), F-III (mercaptoethanol, alcohol-soluble; β-, γ-, δ-zeins), F-IV (mercaptoethanol, pH 10-soluble; glutelin-like), and F-V (mercaptoethanol, pH 10-soluble, detergent-soluble; glutelins) (Misra et al., 1976) . The LM fractions were separated and visualized in Coomassie Blue-stained SDS-PAGE gels. As expected, the α-zeins migrated as two major bands, corresponding to the 19 and 22 kDa subfamilies. The identity of the 19 and 22 kDa α-zein bands was confirmed by peptide mass fingerprinting (PMF) with statistically significant Mowse scores ($P < 0.05$). B73*o2* showed a drastic reduction in the relative amount of both 22- and 19 kDa α-zein subfamilies (Fig. 4.2A). Similar results have been reported for other normal vs. o*2* comparisons (Consoli and Damerval, 2001; Gibbon et al., 2003). The concentration of proteins in F-II and F-III combined (i.e. zeins) was 55.4 μg mg^{-1} flour in B73, 40.2 μg mg^{-1} flour in CML161, and 8.1 μg mg^{-1} flour in B73*o2* (Fig. 4.2B), suggesting intermediate abundance of zeins in CML161 when compared to high-zein B73 and low zein B73*o2* . These observations agree with previous reports in which the *o2* modified materials (i.e. QPMs) have shown intermediate levels of several zein proteins (Robutti et al., 1974a), including the levels of

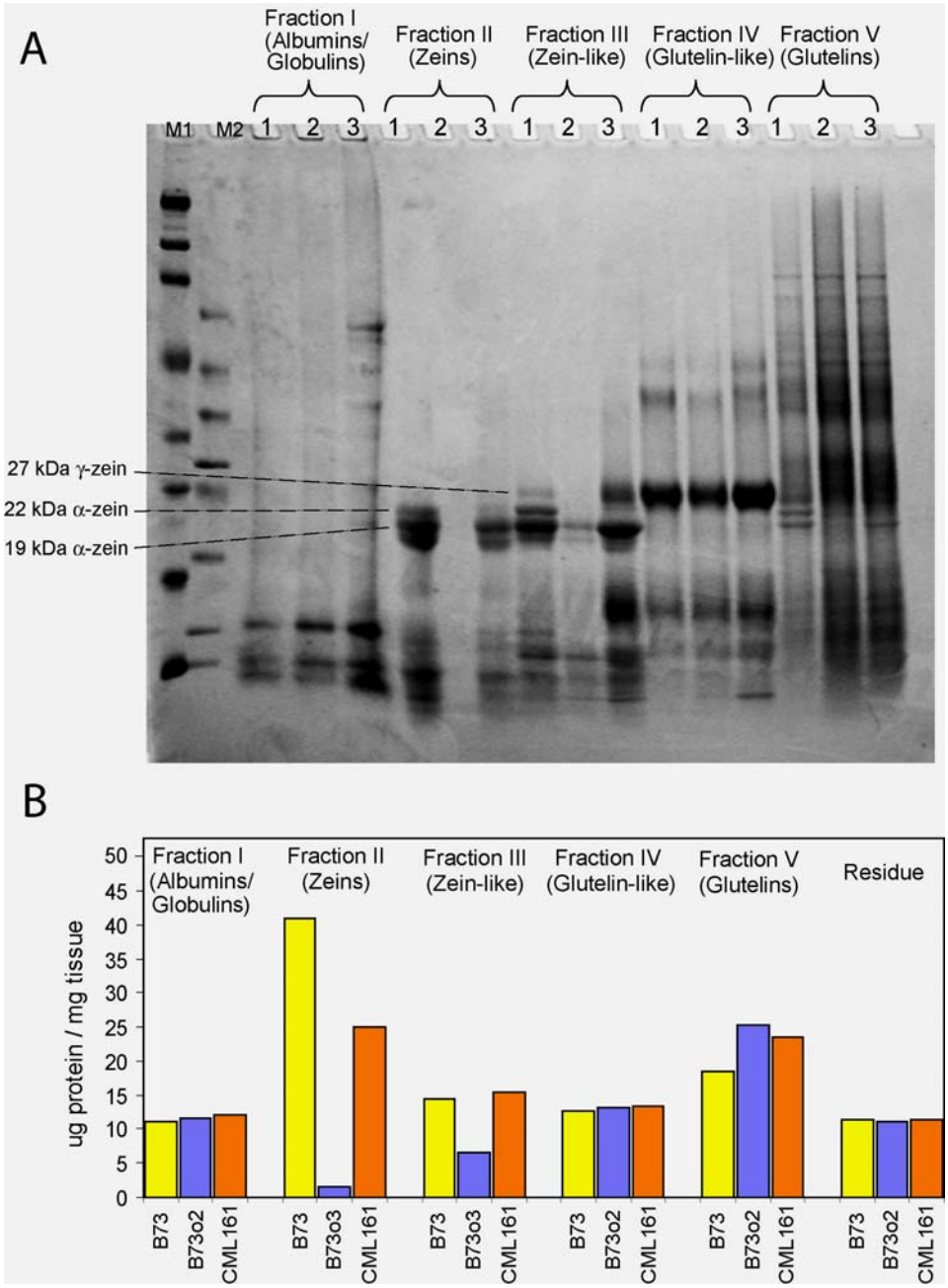

Figure 4.2. Landry-Moreaux (LM) fractions of mature endosperm protein in the inbreds B73, B73o2, and CML161. (A) SDS-PAGE analysis of LM fractions. Different amounts of protein sample were loaded onto the gel (F-I, 3.5% of the total extract from 120 µg; F-II, 1.0%; F-III, 2.0%; F-IV, 2.5%; F-V, 1.0%). Bands corresponding to 19 kDa α-zeins, 22 kDa α-zeins, and 27 kDa 27 γ-zeins are shown. (B) Protein concentrations of LM fractions as estimated by BCA protein assay (F-II and F-III) and by Bradford protein assay (F-I, F-IV, and F-V) using BSA as standard.

the 22kDa α-zein subfamily that is under control of the O2 transcription factor (Schmidt et al., 1992).

Microarray Analysis

Differences in the abundance of mRNA transcripts 18 days after pollination (dap) in the endosperm of the maize lines B73, B73*o2* and CML161 was assessed using microarray analysis. This developmental stage has repeatedly been chosen in maize endosperm research because at this point most endosperm cells are already differentiated and show high levels of gene expression (Demason, 1997; Lopez-Valenzuela et al., 2004; Hunter et al., 2002; Song and Messing, 2003). After data processing and quality control 32,186 features were included in the analysis. This number of features corresponds to a much lower number of genes, since high redundancy is observed in the oligonucleotide array composition. A total of 4,043 features (12.5%) in the arrays that showed significant difference between the wild-type line B73 and the *o2* line B73*o2* (Fig. 4.3A). These mRNA abundance differences at 18 dap may reflect differences in expression of genes that are either directly or indirectly associated with the *O2* gene. We expect B73 and B73*o2* to be near isogenic lines after several cycles of backcrossing and selfing. However, any remnant genome from the *o2* allele donor may contribute to observed differences. Besides, differences may arise due to micro-environmental effects of the greenhouse and developmental progress differences caused by the *o2* mutation. Our results agree with prior reports of the strong pleiotropic effect of the *o2* allele in the overall level of gene expression in endosperm (Doehlert and Kuo, 1994; Hunter et al.,

2002; Jia et al., 2007). A total of 8,412 features (26.1%) showed significant difference between the *o2* line and the QPM line CML161 (Fig. 4.3C). These results might include differences in expression of genes involved in the process of endosperm modification in QPM lines, among all other differences expected due to genetic background and allelic differences between the two genomes. We anticipated differences caused by allelic preference because a great majority of sequence information (i.e. EST, mRNA and genomic DNA) is derived from the inbred B73. However, we did not observe a clear bias towards over-representation of significant features in B73 endosperm when compared to CML161. A total of 7,945 features (24.7%) were significant different between B73 and CML161. Both genotypes show a similar number of up- and down-regulated transcripts (Fig. 4.3B). We were especially interested in those transcripts downregulated in B73*o2* when compared to B73, and upregulated in CML161 when compared to B73*o2*, since we hypothesized that the activity of the modifier genes replaces or overcomes some of the effects of the *o2* mutation. A total of 911 features fell in this category with q values < 0.05 (Appendix). Two of the internal controls of the array were overrepresented in this list; one of them was an EST (gi|4776087), with high homology to an α-zein 19kDa precursor, that appeared 154 times. A second EST (gi|6695907), with homology to an unknown protein of rice, appeared 25 times. A total of 128 features lacked putative annotation, whereas 126 features are homologous to uncharacterized proteins from rice, *Arabidopsis*, wheat and other species. Among the remaining 478 significant features, 105 corresponded to different features associated by putative annotation to zeins (47 related to α-zein 19kDa, 22 to α-zein 22kDa, two to β-

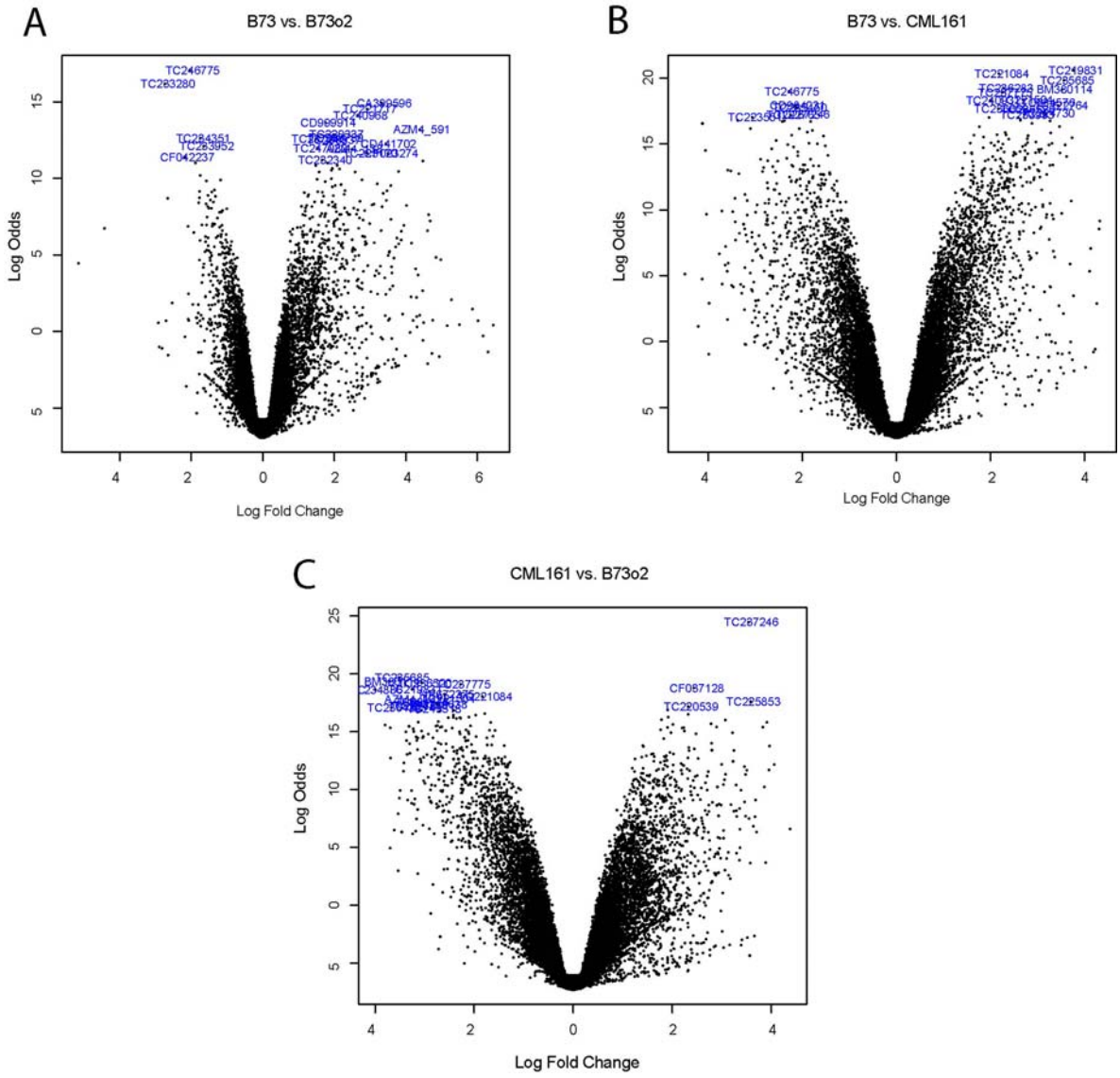

Figure 4.3. Volcano plots of evidence for differential expression in 18 dap endosperms of the inbred lines B73, B73*o2*, and CML161 comparisons. Plot shows log-fold changes (x-axis) versus log-odds of differential expression (y-axis).

zein 15kda, one to γ-zeins and 33 unidentified zeins). A group of 11 features showed putative annotation related to different ribosomal proteins, approximately 10 to transcription and transcription factors, and at least 12 to amino acid metabolism. Other transcripts found on this list were a putative pyrophosphate-fructose 6-phosphate 1-phosphotransferase alpha subunit (PFP), a putative inorganic pyrophosphatase (PIP), and members of the glutathione S-transferase (GST) family. A more complete and straightforward analysis of the array results (e.g. gene ontology analysis) is hindered by incomplete annotation of the publicly available maize DNA and RNA sequences that were used to design the array. The ongoing efforts aimed to produce a complete sequence of the maize genome and to characterize the genes that control endosperm development will facilitate such analysis in the future (http://www.endosperm.info).

Quantitative RT-PCR, Means, ANOVA, and Repeatabilities

In order to corroborate and complement the results of the array analysis, we selected a group of 29 genes based on the preliminary information obtained from the oligoarray analysis, and from previous reports (Table 4.1). Zein nomenclature and gene information followed the work of Woo et al. (2001). Using specific primers for each of the 29 transcripts, cDNA from endosperm 18 dap of the parental inbreds B73, B73*o2* and CML161 was amplified by qRT-PCR and ΔCt values were calculated. Least significant difference (LSD) values were calculated from the ANOVA outputs to determine statistical differences ($n=3$, $p < 0.05$) between mean values (Table 4.2). Twenty transcripts showed significant differences between at least on pair of means. The

Table 4.2. Analysis of variance of the 29 transcripts evaluated by quantitative RT-PCR in endosperm 18 days after pollination of the inbred lines B73, B73*o2* and CML161.

Transcript	Inbred	ΔCt mean	ΔCt SE†	LSD‡	Transcript	Inbred	ΔCt mean	ΔCt SE	LSD
αGL18	B73	10.36	1.08	B	γZN50	B73	8.20	0.67	A
	B73*o2*	10.24	0.64	B		B73*o2*	9.84	0.66	A
	CML161	28.13	3.34	A		CML161	8.44	0.32	A
αZN19B3	B73	8.60	0.52	B	γZN16	B73	6.72	1.09	A
	B73*o2*	14.92	0.45	A		B73*o2*	7.45	0.48	A
	CML161	9.68	0.11	B		CML161	6.50	0.15	A
αZN22Z1	B73	7.61	0.68	C	γZN27	B73	4.81	0.53	B
	B73*o2*	19.68	1.12	A		B73*o2*	6.75	0.30	A
	CML161	13.86	0.10	B		CML161	4.09	0.07	B
αZN19B1	B73	6.22	0.47	C	idbeI	B73	14.32	0.41	A
	B73*o2*	15.65	0.93	A		B73*o2*	14.80	0.65	A
	CML161	10.09	0.17	B		CML161	14.96	0.45	A
αZN19B4	B73	9.35	0.32	B	Leg50k	B73	10.04	0.43	A
	B73*o2*	15.63	0.57	A		B73*o2*	10.52	0.20	A
	CML161	10.30	0.25	B		CML161	9.58	0.52	A
αZN22Z3	B73	5.89	0.47	C	LKR/SDH	B73	14.10	0.34	B
	B73*o2*	18.72	0.55	A		B73*o2*	22.24	0.67	A
	CML161	9.67	0.50	B		CML161	15.94	0.57	B
αZN22Z4	B73	7.56	0.73	C	MZ41626	B73	16.03	0.26	B
	B73*o2*	19.82	1.10	A		B73*o2*	38.28	0.35	A
	CML161	10.58	0.54	B		CML161	14.63	0.52	C
αZN22Z5	B73	9.92	0.49	C	O2	B73	12.22	0.49	B
	B73*o2*	23.96	1.85	A		B73*o2*	21.83	0.95	BA
	CML161	18.26	0.58	B		CML161	30.42	4.78	A
b32/RiP	B73	11.05	0.54	B	PFP	B73	14.97	0.66	B
	B73*o2*	19.42	0.93	A		B73*o2*	18.72	0.39	A
	CML161	13.29	0.53	B		CML161	12.88	0.34	C
b70/BiP	B73	13.09	0.45	A	PPDK	B73	11.08	0.47	B
	B73*o2*	14.26	0.28	A		B73*o2*	13.44	0.50	A
	CML161	13.69	0.66	A		CML161	13.37	0.18	A
βZN15	B73	18.49	0.87	B	RAPB	B73	16.07	0.26	B
	B73*o2*	28.95	1.75	A		B73*o2*	17.25	0.33	A
	CML161	29.73	1.08	A		CML161	15.25	0.38	B
δZN10	B73	7.39	0.64	B	sbeI	B73	12.45	1.79	A
	B73*o2*	11.05	0.57	A		B73*o2*	11.98	1.76	A
	CML161	9.73	1.25	AB		CML161	11.79	1.59	A
δZN18	B73	9.59	0.92	C	ssIIa	B73	13.55	0.35	A
	B73*o2*	16.73	0.55	B		B73*o2*	13.06	0.26	A
	CML161	21.38	0.72	A		CML161	13.29	0.37	A

Table 4.2. Continued.

Transcript	Inbred	ΔCt mean	ΔCt SE†	LSD‡	Transcript	Inbred	ΔCt mean	ΔCt SE	LSD
eEF1Ac	B73	13.12	0.63	A	WRKY	B73	16.31	1.27	A
	B73*o2*	13.45	0.60	A		B73*o2*	13.14	0.29	B
	CML161	13.30	0.57	A		CML161	17.50	0.63	A
eEF1Ab	B73	12.02344	0.38	A					
	B73*o2*	12.6333	0.26	A					
	CML161	12.42269	0.41	A					

† Standard error of the ΔCt mean for three biological replicates.
‡ ΔCt means of inbred lines with the same letter are not significantly different as $P < 0.05$ (LSD).

29 transcripts were also evaluated by qRT-PCR in endosperm 18 dap of a set of 93 RILs planted in the field. The frequency distribution of the ΔCt values in the RILs showed a normal distribution for most of the transcripts except for aGL18, bZN15, dZN18, MZ41626, PFP and WRKY that showed a bimodal distribution (Fig. 4.4). We observed transgressive segregation for the majority of the transcripts. Similar results have been reported in Arabidopsis (West et al., 2007) and yeast (Brem and Kruglyak, 2005). The ANOVA analysis showed significant effects of the RILs ($p < 0.0001$) in the observed variation of all the transcripts analyzed, except by the transcript corresponding to the *O2* gene (Table 4.2). Broad-sense repeatability was estimated for each of the 28 significant transcripts. The range of repeatabilities for the relative abundance of mRNA was 0.65-0.92. It has been proposed that transcriptional heritabilities would tend to be higher than phenotypic heritabilities because large differences in transcript abundance are buffered later during development or by post-transcriptional modifications (Gibson and Weir, 2005).

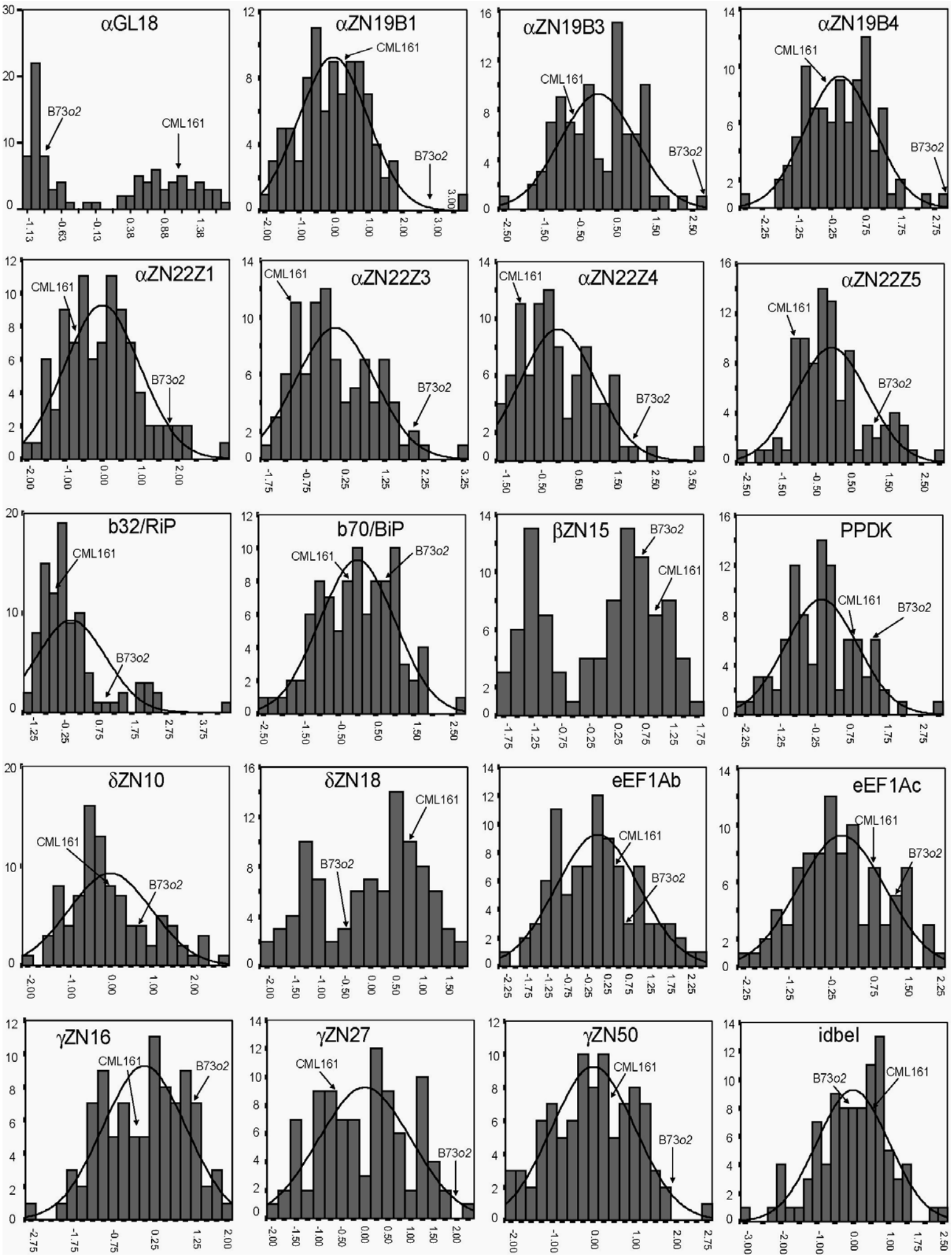

Figure 4.4. Frequency distribution of standardized ΔCt values (normalized to 18s ribosomal gene levels) for each of the 29 transcripts analyzed in 18 dap endosperm in the population of RILs derived from B73*o2* x CML161.

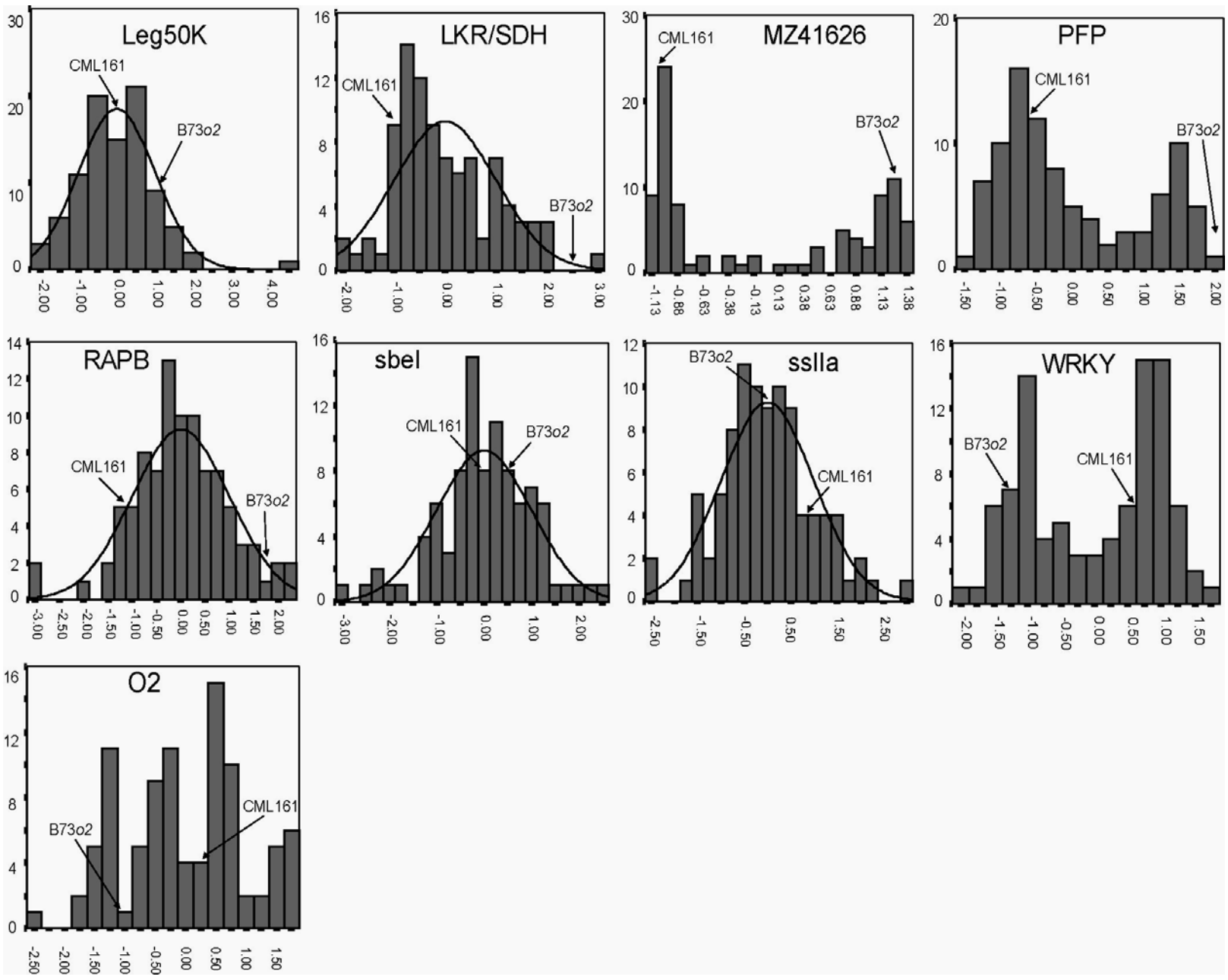

Figure 4.4. Continued.

Individual QTL Analysis

CIM analysis was conducted for relative mRNA levels (as ΔCt values) of each of the 29 transcripts analyzed by qRT-PCR. Most of the eQTLs explained a small proportion of the observed variation (R^2 < 25%) (Fig. 4.5). A smaller proportion of eQTL explained 25-90 % of the variation. The CIM analysis suggested that major eQTLs were identified for those transcripts with bimodal distribution and corresponded to *cis*-acting loci (i.e. coincident with the physical position of encoding genes). The 15 kDa β-zein transcript (βZN15) mapped to the bin 6.01, corroborating the physical location determined by *in silico* mapping (Lawrence et al., 2004; Lawrence et al., 2007). The 18 kDa α-globulin transcript (αGL18) showed a high eQTL peak and a shoulder peak in bins 6.05-6.06. This location corresponded to location of the gene determined in a population derived from B73 x Mo17 (Woo et al., 2001). The 18 kDa δ-zein transcript (δZN18) showed a major eQTL in bin 6.04, which agrees with the location estimated by Swarup et al. (1995). Likewise, the putative PFP was mapped to bin 9.03, the WRKY1 was mapped to bin 1.07, and the transcript MZ41626 was mapped to the bin 10.07 (Table 4.3). Additional eQTLs were located for the transcripts δZN18, MZ416626, PFP, and WRKY1 and the remaining transcripts that explained variation not associated with the encoding loci (i.e. *trans*-acting loci) (Table 4.3).

Individual CIM analysis detected several significant eQTLs for the α-zein family transcripts (Table 4.3). The majority of the eQTLs that increased the relative abundance

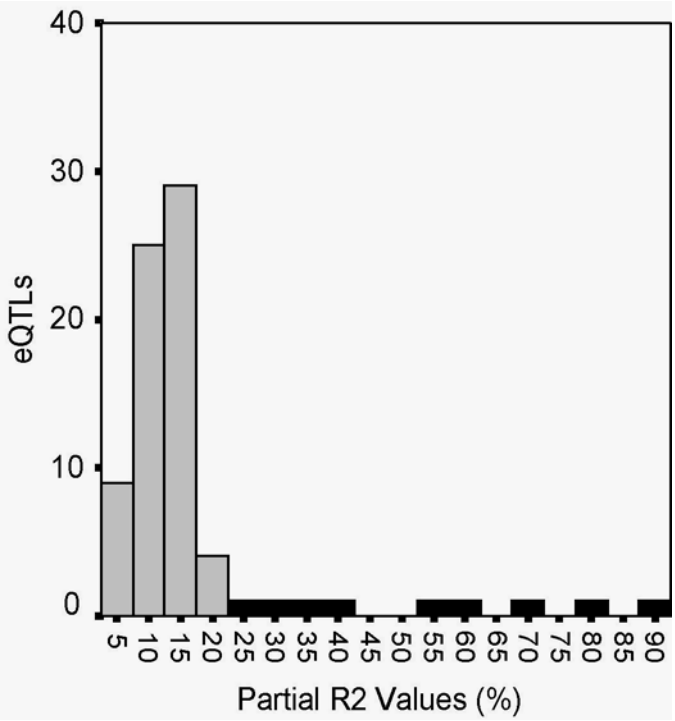

Figure 4.5. Distribution of the percentage of phenotypic variation (R^2) explained by all eQTLs located for the 29 transcripts. The gray bars show eQTLs detected for trans-acting eQTL loci and the black bars show cis-acting eQTLs.

of α-zein mRNA (i.e. decreased the ΔCt value) derived from the inbred CML161. However, there were eQTLs derived from the inbred B73*o2* that increased the abundance of α-zein mRNA. Because of the extensive gene amplification in the 19 kDa α-zeins it is possible that qRT-PCR amplicons correspond to transcripts from more than one loci. For instance, the cDNAs αZN19B1 and αZN19B3 (Woo et al, 2001) were found to be a mixture of different very homologous genes (Song and Messing, 2002). Very often the different copies of the genes are tandemly arranged within a short physical distance (Song and Messing, 2002). The cDNAs αZN19B1, and αZN19B4 has been mapped to chromosome 7 (bin 7.02) inside the locus *zein polypeptidesL2b* (zpl2b) (Lawrence et al., 2004; Lawrence et al., 2007). None of the individual eQTLs detected by CIM mapped to this region of chromosome 7, where the *O2* (bin 7.01) locus also resides. Five eQTLs for αZN19B1 were located on chromosomes 1S (bin 1.01), 3L (bin 3.05), 4S (bin 4.02), 4L (bin 4.10), and 9S (bin 9.03). Two eQTLs for αZN19B4 were detected on chromosomes 3L (bin 3.09), and 9S (bin 9.02). The cDNA αZN19B3 belongs to the zein subfamily z1A (Woo et al., 2001), that has been mapped in two locations of chromosome 4S (bins 4.02 and 4.04) (Song and Messing, 2002). CIM analysis located eQTLs for αZN19B3 on chromosomes 6S (bin 6.00), and 9S (bin 9.03).

The cDNAs αZN22Z1, αZN22Z3, αZN22Z4, and αZN22Z5 belong to the 22 kDa α-zein subfamily z1C, and their genes have been mapped to two chromosome locations in 4S (bins 4.02 and 4.03) (Song et al., 2001). CIM analysis detected eQTLs

Table 4.3. Composite interval mapping analysis for 29 transcripts analyzed by quantitative real-time in endosperm 18 dap from a set of 93 recombinant inbred lines derived from B73*o2* x CML161.

Transcript	Bin	eQTL peak	Marker	eQTL interval[†]	LOD	eQTL Effect[‡]	Partial R2
αGL18	6.05	91.83	umc1805	89.8-93.8	10.99	-8.05	37.75
	6.06	102.12	umc1912	100.1-104.1	20.85	-8.20	53.97
αZN19B1	1.01	17.31	umc2012	9.3-24.4	2.73	-0.64	8.28
	3.05	82.94	umc1307	76.9-87.3	4.19	0.78	12.37
	4.02	44.46	umc1294	40.9-49.5	9.33	-1.14	28.83
	4.10	170.11	umc1180	164.5-170.1	3.08	0.57	7.10
	9.03	86.88	phi065	77.9-94.6	3.51	0.72	11.17
αZN19B3	6.00	16.01	umc1143	5.0-28.5	2.27§	0.46	12.81
	9.03	79.88	phi065	59.1-88.9	2.80	0.50	15.33
αZN19B4	3.09	168.99	umc2152	162.0-170.0	3.71	0.58	14.92
	9.02	65.09	bnlg244	58.1-79.9	3.98	0.56	17.33
αZN22Z1	3.05	83.28	umc1973	75.9-87.3	3.66	1.05	13.86
	9.04	90.56	bnlg1714	82.9-93.6	4.16	1.07	16.46
αZN22Z3	3.05	80.94	umc1307	71.9-87.3	3.06	1.19	11.19
	4.10	168.11	umc1180	158.5-170.1	2.45§	0.90	7.34
	9.02	60.09	bnlg244	45.8-67.1	4.04	1.29	15.35
αZN22Z4	4.06	109.1	bnlg2291	105.4-113.1	3.28	-1.10	9.92
	4.10	170.11	umc1180	158.5-170.1	2.47§	0.95	7.71
αZN22Z5	1.04	94.99	umc2112	89.4-98.0	3.14	0.98	11.31
	3.05	84.28	umc1973	74.9-89.3	2.85	1.01	11.25
	4.06	105.94	bnlg1137	99.4-113.1	2.48§	-0.90	9.32
b32/RiP	6.07	141.35	umc2165	136.6-149.4	2.37§	1.13	7.49
	9.04	90.56	bnlg1714	83.9-94.6	2.78	1.28	8.70
b70/BiP	1.03	79.91	umc2096	72.9-88.4	2.40§	-0.27	11.48
	10.00	9.01	phi041	0.0-18.1	2.76	0.28	12.01
βZN15	6.01	24.52	y1ssr	22.5-26.5	23.97	-3.27	71.28
δZN10	1.02	50.87	bnlg1953	42.9-66.9	3.09	0.85	17.37
	2.03	73.97	bnlg2248	62.5-77.6	2.55	0.67	9.95
	9.03	76.88	phi065	61.1-83.9	2.36§	-0.73	12.57
	10.03	32.72	umc1863	25.1-36.7	3.24	0.76	13.82
δZN18	2.03	77.57	umc1555	71.0-89.5	3.23	1.01	5.75
	3.08	129.05	phi046	126.3-138.4	4.32	-1.26	8.36
	6.04	50.7	umc1014	48.7-56.7	21.82	-3.24	58.61
	8.03	46.5	phi10017	39.5-57.6	3.88	-1.09	6.98
	10.03	32.72	umc1863	27.1-36.7	3.52	1.09	6.42

Table 4.3. Continued.

Transcript	Bin	eQTL peak	Marker	eQTL interval†	LOD	eQTL Effect‡	Partial R2
eEF1Ab	6.01	26.52	y1ssr	21.5-37.3	5.07	0.38	20.75
eEF1Ac	1.11	222.92	umc1553	209.2-236.9	2.98	0.28	15.25
	3.05	107.28	umc1539	98.3-114.6	2.88	-0.27	14.35
	9.03	80.88	phi065	72.7-87.9	3.70	0.30	18.17
γZN16	6.01	27.52	y1ssr	15.0-35.3	2.81	0.38	11.61
	6.06	110.03	umc2170	104.1-120.0	3.15	-0.46	17.45
γZN27	6.02	29.31	bnlg2191	22.5-34.3	3.98	0.50	16.79
	6.05	82.83	umc1805	77.4-91.8	2.62	-0.39	11.31
γZN50	6.00	6.01	umc1143	0.0-20.0	2.15§	0.31	9.24
	6.06	108.12	umc1912	104.1-116.0	5.16	-0.46	21.36
	7.02	72.2	umc2142	67.2-78.2	3.62	-0.35	12.86
idbeI	2.08	169.69	umc1516	157.8-176.7	3.86	0.39	16.7
	3.06	116.64	umc1644	111.1-121.4	2.79	-0.32	10.55
	5.05	136.42	umc2111	132.4-136.4	4.12	0.41	15.32
Leg50	2.07	149.84	umc1042	136.2-171.7	2.85	0.38	14.75
LKR/SDH	1.11	258.38	phi064	254.4-261.4	2.60	-0.53	8.96
	2.06	116.69	nc003	112.5-124.7	4.14	-0.67	14.5
	3.05	84.28	umc1973	76.9-88.3	4.56	0.72	15.36
	10.01	13.08	umc2018	8.0-19.1	3.76	0.58	10.87
MZ41626	10.03	35.72	umc1863	14.1-46.1	2.81	2.35	5.23
	10.07	100.61	umc1084	97.6-103.6	36.08	9.81	89.43
O2	4.06	109.1	bnlg2291	102.4-117.1	2.82	-2.01	14.23
	5.04	121.96	umc1966	99.2-129.0	2.52	2.00	13.62
PFP	9.03	75.88	phi065	74.9-78.9	28.11	2.32	79.05
PPDK	3.06	115.64	umc1644	114.6-117.6	2.68	-0.33	12.52
	3.07	122.25	Txp196L	121.4-125.3	4.24	-0.38	16.05
	4.04	69.89	bnlg490	64.3-74.9	3.11	-0.30	10.2
	6.00	5.01	umc1143	0.0-13.0	2.69	0.35	13.74
RAPB	1.02	47.87	bnlg1953	29.3-60.9	3.41	0.38	20.64
sbeI	3.05	108.28	umc1539	100.3-117.6	3.68	-0.44	15.49
	6.07	134.57	umc1490	131.2-140.6	4.16	-0.42	14.02
	8.06	102.92	umc1724	90.9-110.5	3.22	0.38	12.40
	10.03	46.05	umc2180	43.1-50.6	4.80	0.71	16.82

Table 4.3. Continued.

ssIIa	2.07	149.84	umc1042	136.2-163.8	2.23§	0.26	11.00
	3.05	109.28	umc1539	101.3-111.1	4.71	-0.40	24.26
	6.00	4.01	umc1143	0.0-17.0	2.74	0.31	14.39
	6.07	134.57	umc1490	131.2-140.6	2.91	-0.28	12.39
WRKY1	1.07	175.11	umc1147	171.1-182.3	14.15	-1.70	35.11
	1.10	189.98	umc2149	188.0-193.0	3.56	-0.95	8.83
	2.00	0.01	umc2246	0.0-3.0	3.90	-0.60	6.13
	9.02	56.09	bnlg244	48.8-66.1	2.76	-0.49	4.94

† The QTL support interval corresponded to an interval on either side of the QTL peak that corresponded to a decrease of 1 LOD from the maximum LOD score.
‡ The positive or negative eQTL effects indicate the direction of effect of the B73o2 allele. Positive effects increase ΔCt values (i.e. decrease transcript levels).
§ LOD scores above 2.0 are shown for some transcripts when the R2 value was > 7%.

for aZN22Z1 on chromosomes 3L (bin 3.05), and 9L (bin 9.04); for αZN22Z3 on chromosomes 3L (bin 3.05), 4L (bin 4.10), and 9S (bin 9.02); for αZN22Z4 on chromosome 4L (bins 4.06 and 4.10); and for αZN22Z5 on chromosomes 1S (bin 1.04), 3L (bin 3.05), and 4L (bin 4.06) (Table 4.3).

The 10 kDa δ-zein has been mapped to chromosome 9L (bin 9.03) (Lawrence et al., 2004; Lawrence et al., 2007). CIM analysis for the transcript δZN10 detected a eQTL with LOD 2.36 in the same bin, and additional eQTLs on chromosomes 1S (bin 1.02), 2S (bin 2.03), and 10S near the centromere (bin 10.03) (Table 4.3). The gene encoding the 16 kDa γ-zein was mapped to chromosome 2L (bin 2.07) (Woo et al., 2001). CIM analysis detected eQTLs for the cDNA γZN16 on chromosomes 6S near the centromere (bin 6.01), and 6L (bin 6.06). The gene for the 50 kDa γ-zein has been located on chromosome 7L (bin 7.05) (Woo et al., 2001), and CIM analysis detected eQTLs for the transcript γZN50 on chromosomes 6S (bin 6.00), and 6L (bin 6.06). The

27 kDa γ-zein locus has been mapped on chromosome 7 (Das and Messing, 1987; Geetha et al., 1991). Two eQTLs for the transcript γZN27 were located on chromosome 6L (bins 6.02 and 6.05) (Table 4.3).

Besides some of the zeins, other proteins have been suggested playing roles in the modification of *o2* endosperm. Gibbon et al. (2003) found differences in the starch characteristics of modified *o2* genotypes. Increased concentrations of extractable granule-bound starch synthase I (GBSS I), reduction in levels of intermediate-length chains of amylopectin, and cohesive contacts between starch granules, suggest a role for enzymes of the starch synthesis pathway in the modification of *o2* endosperm (Gibbon et al., 2003). In our eQTL mapping analysis, we included the transcripts of isoamylase debranching enzyme (idbeI), starch branching enzyme I (sbeI), and starch synthase IIa (ssIIa). In the population of RILs, alleles that increased the relative abundance of starch-related transcripts derived from both inbred parents. The idbeI or Sugary1 (Su1) locus has been mapped on chromosome 4S (bin 4.05) (James et al., 1995). CIM analysis detected three eQTLs for idbeI on chromosomes 2L (bin 2.08), 3L (bin 3.06), and 5L (bin 5.05) (Table 4.3). Two copies of the sbeI gene have been mapped on chromosomes 6S (bin 6.01) and 10L (bin 10.04) (Kim et al., 1998). CIM analysis located eQTLs for sbeI on chromosomes 3L (bin 3.05), 6L (bin 6.07), 8L (bin 8.06), and 10L (bin 10.03) (Table 4.3). The ssIIa locus or Sugary2 (Su2) has been mapped to the chromosome 6L (bin 6.04) (Harn et al., 1998; Zhang et al., 2004). CIM analysis detected eQTLs for ssIIa on chromosomes 2L (bin 2.07), 3L (bin 3.05), 6S (bin 6.00), and 6L (bin 6.07) (Table 4.3).

The O2 protein has been found to affect the expression of *PPDK1* by binding to two sites in its promoter (Maddaloni et al. 1996). The functions of *PPDK1* in endosperm seem to be associated with the production of precursors for amino acid synthesis, and the determination of the balance of starch versus protein synthesis (Chastain et al., 2006; Mechin et al., 2007). The locus PPDK1 has been mapped to the chromosome 6L (bin 6.05) (Matsuoka, 1990). We detected eQTLs for PPDK1 using CIM on chromosomes 3L (bins 3.06 and 3.07), 4S (bin 4.04), and 6S (bin 6.00). Three major eQTLs showed favorable alleles derived from B73o2 and one from CML161 (Table 4.3).

We identified eQTLs for two members of the elongation factor 1A (eEF1A) gene family (Carneiro et al., 1999). The expression of eEF1A has been reported upregulated in *o2* mutants (Habben et al., 1995), and strongly correlated with Lys content (Moro et al., 1996). The maize eEF1A genes have been mapped on chromosomes 6, 7 and 8 (Carneiro et al., 1999). CIM analysis detected eQTLs for eEF1Ab on chromosome 6S near the centromere (bin 6.01), and for eEF1Ac on chromosomes 1L (bin 1.11), 3L (bin 3.05), and 9L near the centromere (bin 9.03) (Table 4.3). Only the eQTL on chromosome 6S seem to colocalize with one of the 4 QTLs reported for eEF1A protein content in mature endosperm (Wu et al., 2002)

The expression of the b32 ribosomal inactivating protein (b32/RIP) is influenced by the activity of the O2 transcription factor (Lohmer et al., 1991). The b32/RIP locus has been mapped to the chromosome 8L (bin 8.05) (Soave et al., 1981). CIM analysis detected eQTLs for b32/RIP on chromosomes 6L (bin 6.07), and 9L (bin 9.04). Favorable alleles for both eQTLs derived from CML161 (Table 4.3). The b70/BiP

protein has homology to a chaperon-like heat shock protein (HSP), and is affected by the endosperm mutants *floury2* (*fl2*), *Defective endosperm-B30 (De-b30*), and *Mucronate* (*Mc*). The b70/BiP protein seems to have a role in zein folding and assembly (Boston et al., 1991). CIM analysis detected eQTLs for b70/BiP on chromosomes 1S (bin 1.03), and 10S (bin 10.00) (Table 4.3). O2 has been found affecting the regulation of the lysine-ketoglutarate reductase/ saccharopine dehydrogenase (LKR/SDH) gene in maize endosperm (Brochetto-Braga 1992). The higher concentration of lysine found in *o2* genotypes may be associated to the decrease in lysine reductase activity. The LKR/SDH locus has been mapped in silico to chromosome 4L (bin 4.07) (Lawrence et al., 2004; Lawrence et al., 2007). CIM analysis located eQTLs for LKR/SDH on chromosomes 1L (bin 1.11), 2L (bin 2.06), 3L (bin 3.05), and 10S (bin 10.01). Favorable alleles derived from both parental inbreds (Table 4.3).

The 50 Kda Legumin protein is a non-zein endosperm storage protein (Woo et al., 2001). The locus Leg50k has been mapped to chromosome 6S (bin 6.01) (Woo et al., 2001). CIM analysis identified one eQTL on chromosome 2L (bin 2.07), with the favorable allele derived from CML161 (Table 4.3). Finally, we identified eQTLs for a putative maize RAPB transcription factor that was differentially expressed in the parental inbreds. CIM located one eQTL for the RAPB transcript on chromosome 1S (bin 1.02) (Table 4.3).

Although the ANOVA analysis did not show significant effects of the RILs in the variation observed for the transcript corresponding to the *O2* gene, CIM analysis located eQTLs for the *O2* transcript on chromosomes 4L (bin 4.06), and 5L (bin 5.04). The

alleles that increase the level of *O2* transcript derived from B73*o2* and CML161 respectively.

Variable Clustering

We applied a variable reduction scheme that allowed simplifying the analysis of the genetic control of the 29 expression traits. Substantial benefits can be obtained by variable reduction when statistical analysis is performed on multiple variables (Lan et al., 2003; Upadyayula et al., 2006). Hierarchical analysis and PCA can be used as methods to eliminate redundant dimensions in a data set. Groups of variables are formed with variables that are as correlated as possible among themselves and as uncorrelated as possible with variables in other clusters. The relationship between the 29 transcripts was depicted as a biplot with two principal components (Fig. 4.6). In SVD biplots, an angle of 90° between any two trait vectors suggest that the traits are independent, an angle of >90° indicates negative correlation, and an angle of <90° suggests a positive correlation. SVD suggested a strong correlation between groups of transcripts. Subsequently, we performed hierarchical clustering on the standardized ΔCt

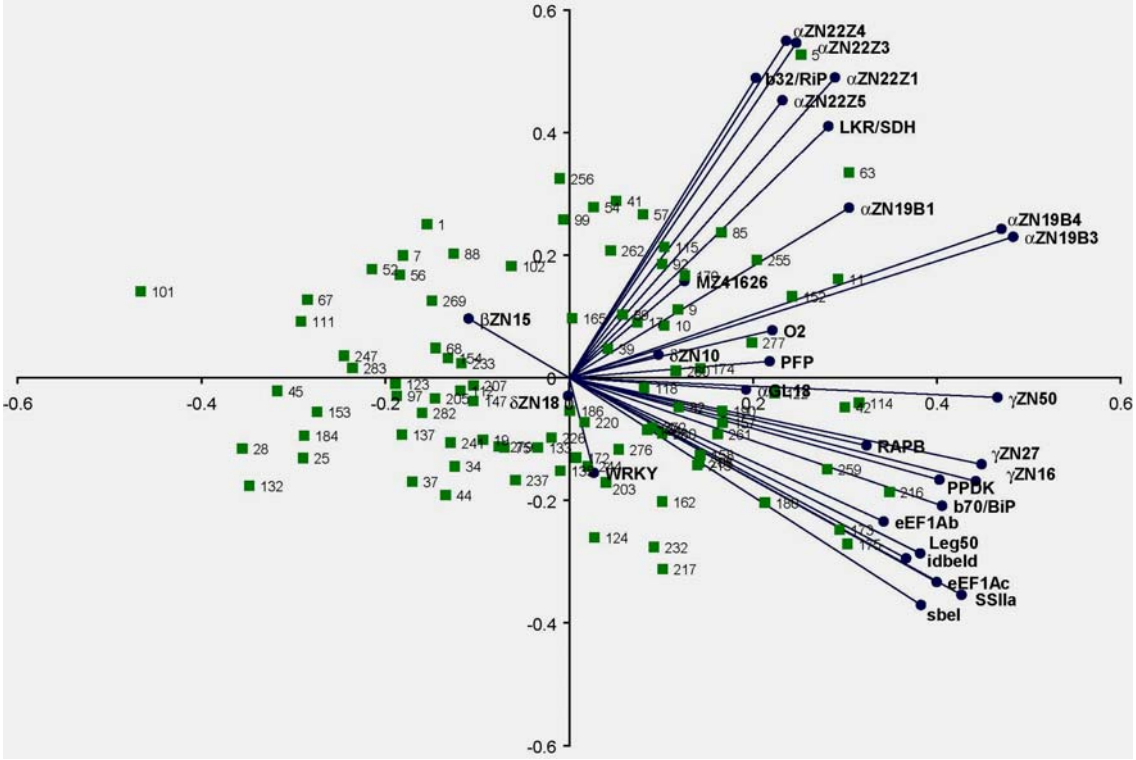

Figure 4.6. Single value descomposition biplot of first two principal components for 29 transcripts analyzed in 18 dap endosperm in the population of RILs. Expression traits are represented as circles, RILs are represented as squares.

values of the 29 transcripts included in this study. Six groups of genes were formed at the clustering level at which the proportion of the variance explained by clustering was more than an arbitrarily chosen threshold of 0.8 (Fig. 4.7). We named the clusters accordingly to their most representative gene transcripts. One group was composed by the three 19 kDa α-zeins included in the study (cluster "αZN19kDa"). Another group (cluster "αZN22kDa") contained four 22 kDa α-zeins, the b32/RiP and the LKR/SDH transcripts. The cluster "δZNs" was composed by the 18 kDa δ-zein and the 10 kDa δ-

zein transcripts. The 16 kDa, 27 kDa, and 50 kDa γ-zeins clustered together (cluster "γZNs"). The cluster "starch" included three starch-related transcripts (idbeI, sbeI and ssIIa) and the 50kDa

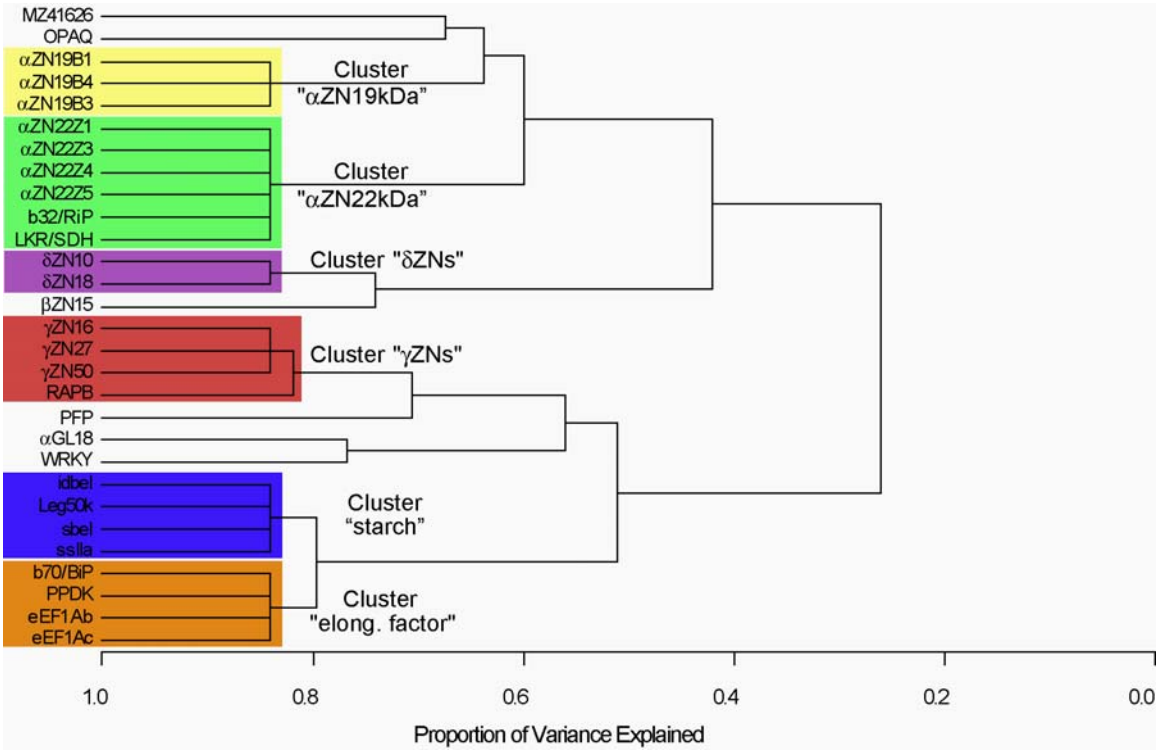

Figure 4.7. Hierarchical clustering of relative abundance of mRNA levels of 29 gene transcripts in maize endosperm of recombinant inbred lines (B73*o2* x CML161) 18 days after pollination. For each gene transcript the ΔCt values (normalized to 18s ribosomal gene levels) were standardized to mean 0 and variance 1.

legumin. Finally, the two elongation factor transcripts, the cytosolic PPDK1, and the b70/BiP transcripts formed another group (cluster "elongation factor"). The clusters "starch" and "elongation factor" are related by a common clustering branch that explained 80% of the variation. The six remaining transcripts did not form groups under the selected criteria.

Multiple-Trait QTL Analysis

We used MT-CIM to identify eQTLs explaining genetic variation in the RILs for the relative levels of mRNA of the selected genes that formed the clusters defined above (Fig. 4.8). We compared the location of these eQTLs with the map positions of QTLs detected by MT-CIM for four measurements of endosperm texture modification and for the contents of the essential amino acids lysine, methionine and trypthophan (Chapter III).

Cluster "αZN19kDa". Seven eQTL peaks were identified with LOD scores above 2.5. These eQTLs were located on chromosomes 1S (bin 1.01), 3L (bin 3.05), 4S (bins 4.01 and 4.02), 5L (bin 5.05), 6L (bin 6.06), and 9S (bin 9.03). The eQTLs mapped on chromosomes 3L, 5L, and 6L by MT-CIM colocalized with QTL hot spots detected for endosperm texture modification (Chapter III). In addition, MT-CIM suggested another eQTL/QTL colocalization on chromosome 9L near the centromere (Fig. 4.8).

Cluster "αZN22kDa". A total of 17 eQTL peaks were identified with LOD scores above 2.5. These eQTLs were located on chromosomes 1S (bin 1.02), 1L (bin 1.04), 2S (bins 2.01, 2.02, and 2.04), 2L (bins 2.06 and 2.08), 3L (bin 3.05), 4S (bins

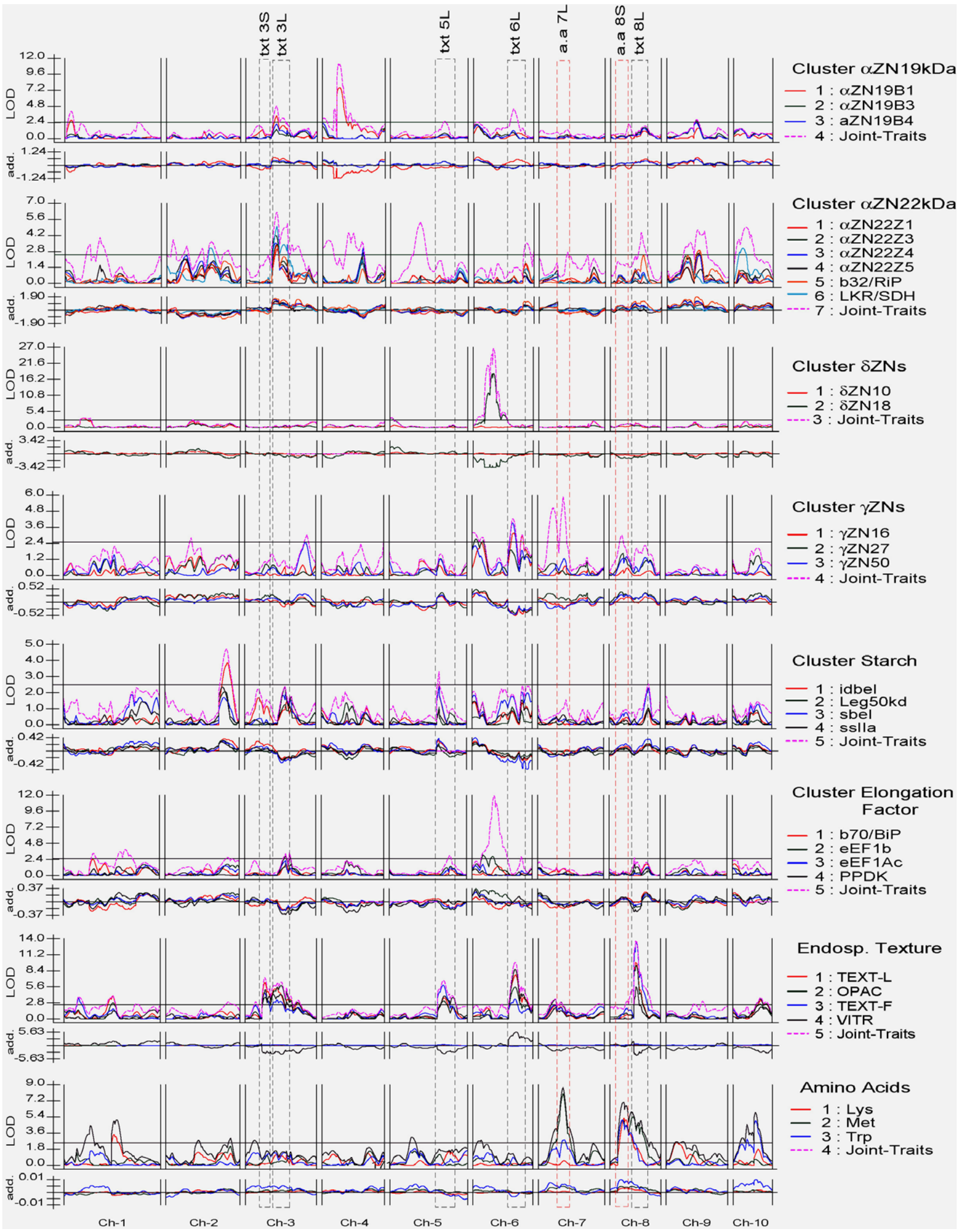

Figure 4.8. Multiple-trait composite interval mapping for mRNA relative abundance of the clusters of transcripts: "αZN22kDa", "αZN19kDa", "δZNs", "γZNs", "Starch", and "Elongation Factor"; and the phenotypic traits endosperm texture and amino acid contents. LOD scores for each transcript and for joint analysis inside each cluster are depicted by solid lines and dashed lines, respectively. The horizontal axis corresponds to the genetic maps in cM. A horizontal line is shown corresponding to the LOD score threshold of 2.5 used to declare the presence of a QTL. The QTL effects of the allele B73*o2* are depicted in the bottom plot for each cluster. Dashed boxes enclose QTL hot spots for endosperm texture and amino acid contents according to Chapter III.

4.00 and 4.04), 5L (bin 5.04), 6L (bin 6.07), 7L (bin 7.02), 8L (bin 8.06), 9S (bin 9.02), 9L (bin 9.03), and 10S (bin 10.02). The eQTL mapped by MT-CIM on chromosomes 3L, 6L, and 8L colocalized with QTL hot spots detected for endosperm texture modification (Chapter III). Another eQTL colocalized with an endosperm texture QTL on chromosome 9L. MT-CIM detected eQTLs for this cluster that colocalized with QTLs for amino acid contents (Chapter III) on chromosomes 1S, 2S, 2L, 5L, 7L, 8L, and 10 near the centromere (Fig. 4.8)

Cluster "δZNs". Six eQTLs peaks with LOD scores above 2.5 were detected. The eQTLs were located on chromosomes 1S (bin 1.02), 2S (bin 2.03), 5S (bin 5.00), and 6L (bins 6.02, 6.05, and 6.06). None of the eQTLs mapped by MT-CIM for cluster "dZNs" colocalize with the 5 QTL hot spots for texture traits (Chapter III). However, MT-CIM suggested an eQTL/QTL colocalization on chromosome 1S. One eQTL on chromosome 1S colocalized with a QTL detected by MT-CIM for amino acid contents (Fig. 4.8)

Cluster "γZNs". Nine eQTLs were identified with LOD scores above 2.5. The eQTLs were located on chromosomes 2S (bin 2.03), 3L (bin 3.09), 4L (bin 4.10), 6S (bin 6.00), 6L (bins 6.06 and 6.07), 7S (bin 7.00), 7L (bin 7.02), and 8S (bin 8.02). One eQTL mapped by MT-CIM for the cluster "γZNs" colocalized with the QTL hot spot identified for texture traits on chromosome 6L (Chapter III). In addition MT-CIM detected eQTLs for this cluster that colocalized with endosperm texture QTLs on both sides of the centromere on chromosome 7. MT-CIM detected eQTLs for Cluster "γZNs" that corresponded to amino acid contents QTLs on chromosomes 6S, 7L, and 8S. (Fig. 4.8)

Cluster “starch”. Three eQTLs were identified with LOD scores above 2.5. These eQTLs were located on chromosomes 2L (bin 2.08), 5L (bin 5.05), and 8L (bin 8.07). One eQTLs mapped by MT-CIM on chromosome 5L colocalized with a QTL hot spot detected for texture traits (Chapter III). MT-CIM suggested another eQTL/QTL colocalization in 2L. One eQTLs for the cluster “starch” colocalized with a QTL for amino acid contents on chromosome 2L. (Fig. 4.8)

Cluster “elongation factor”. Eight eQTLs were detected with LOD scores above 2.5. These eQTLs were located on chromosomes 1S (bin 1.03), 1L (bin 1.07), 2L (bin 2.08), 3L (bins 3.06 and 3.07), 5L (bin 5.06), 6L (bins 6.05 and 6.07). The eQTLs mapped by MT-CIM on chromosomes 3L, 5L, and 8L colocalized with QTL hot spots for texture traits (Chapter III). Besides, MT-CIM suggested another eQTL/QTL colocalization in 2L. MT-CIM detected eQTLs for the cluster “elongation factor” that colocalized with QTLs for amino acid contents on chromosomes 1S, 1L, 2L, and 8L. (Fig. 4.8)

Discussion

The modification of the *o2* phenotype has proven to be a complex genetic mechanism involving additive, dominant and recessive gene actions, as well as paternal (i.e., xenia) and maternal cytoplasm effects (Lopes and Larkins, 1995; Lou et al., 2005; Vasal et al., 1980; Wessel-Beaver and Lambert, 1982; Wessel-Beaver et al., 1985). There is not a single genetic model that fits the genetic mechanism of modification in all the genetic backgrounds (Lopes et al., 1995; Lou et al., 2005).

The inbred line B73 is very susceptible to the effects of the *o2* allele. The kernel of its *o2* conversion, B73*o2*, is completely opaque, more opaque than kernels of other o2 genotypes such as W22*o2* and W64*o2* that we have examined from different sources. Moreover, the germination rate, plant vigor, susceptibility to insects, and tolerance to aflatoxin accumulation are compromised in B73*o2*. On the other hand, CML161, a QPM line, is also homozygous for an *o2* but is the product of recurrent selection for modified endosperm texture, enhanced amino acid content, vigor, yield, and disease resistance. Protein analysis demonstrated that mature endosperm of B73*o2* and CML161 contain patterns of storage proteins typical of *o2* and QPM lines, respectively. The proportion of α-zeins in mature endosperm is drastically reduced in the *o2* inbred. Zeins and several other storage proteins seemed to be in intermediate levels in the QPM line, as has been found in most modified materials (Robutti et al., 1974a; Gibbon et al., 2003). PAGE-SDS suggested a higher accumulation of 27 kDa γ-zeins in the QPM, in agreement with has been reported for several other modified genotypes (Geetha et al., 1991). These results together, suggested that the population of RILs derived from B73*o2* and CML161 was adequate to map modifier genes for the gene *o2* by both classical and expression QTL analysis.

Global analysis of gene expression in the parental inbreds showed a strong pleiotropic effect of the *o2* mutant allele on mRNA relative levels in developing endosperm. Pleiotropic effects of *o2* alleles have been observed in several genetic backgrounds (Damerval and Guilloux, 1988; Lopez-Valenzuela et al., 2004; Hunter et al., 2002). The general effect of the *o2* mutation in endosperm gene expression might be

a consequence of the role of the *O2* protein as a determinant of progression towards endosperm maturity. Endosperms homozygous for non-functional copies of *O2* would be arrested in an early developmental stage (Landry et al., 2004; Robutti, 1997a; Dombrik-Kurtzman and Bietz, 1993). This marginal effect of the *o2* allele on the expression of several hundred genes makes the identification of modifier genes by simple *o2* vs. QPM comparisons a challenging task. Besides, differences in gene expression observed between B73*o2* and CML161 cannot be attributed only to the activity of modifier genes since both genotypes are genetically heterogeneous. Furthermore, the prevalence of nonadditive genetic variation might cause that transcript variation observed in the parents will not be reflected in the segregating progeny (West et al., 2007). Therefore, we complemented the list of genes provided by microarray analysis of gene expression with other genes associated with the effects of the *o2* allele and potentially involved in the modification that were not found as differentially expressed in the array results. For instance, in the eQTL mapping analysis by qRT-PCR we included genes involved in the biosynthesis of starch, genes under the direct control of the O2 protein, genes involved in the catabolism of lysine, and the *O2* transcript.

The zein transcripts were well represented in the list of differentially expressed genes. Our results showed differences in the abundance of mRNA transcripts of the 19- and 22 kDa α-zeins between the *o2* and the modified line. Some members of the 22kDa α-zeins subfamily are under direct transcription regulation by the Opaque2 transcription factor. The O2 protein binds a promoter element in the 22 kDa α-zein to activate their expression (Schmidt et al., 1992). It has been suggested that in some *o2* lines other

regulatory proteins may overcome the loss of O2 in the activation of α-zeins expression under certain growth conditions (Ciceri et al., 2000; Muller et al., 1997). Jia et al. (2007) used microarrays to compare endosperm gene expression of eight inbred lines and their corresponding *o2* conversions. A group of 37 transcripts were affected by the mutation in all the eight genotypes, but the zein genes were not included. However, the severity of the *o2* phenotype among the eight inbred was also variable and associated with the differences in zein expression. Other reports have demonstrated the genetic background dependency of the expression of the zein genes (Bernard et al., 1994; Ciceri et al., 2000; Lopez-Valenzuela et al., 2004). This could be an indication that the modifier genes responsible for the kernel phenotype of QPM lines affect the penetrance of *o2* in different genetic backgrounds. If the phenotypic differences between B73*o2* and CML161 are associated to the level of gene expression of zeins, then the location of the QTLs in charge of the variation would provide the genomic location of modifier genes.

The mRNAs of the zeins are detected in the endosperm tissue from 8-10 dap and reach their maximum expression at approximately 16-18 dap (Dolfini et al., 1992). The same pattern of gene expression is shared by several zein genes, suggesting a coordinated regulation of transcription (Dolfini et al., 1992). The clustering pattern that we observed for both 19- and 22 kDa α-zein transcripts support the idea of a common regulation. We detected eQTLs for relative levels of α-zein mRNA on chromosomes 1, 3, 4, 6, and 9. The majority of the eQTLs for α-zeins did not correspond to genomic location of their encoding genes. Therefore, the eQTLs are likely to correspond to elements that affect the α-zein gene expression *in trans*. The colocalization of eQTLs

with QTL hot spots identified for endosperm texture on chromosomes 3, 5, 6 and 8 (Chapter III) suggest a role for α-zein expression in the modification of *o2* endosperm in the RIL population.

Besides the α-zeins, other zein proteins are involved in the development and maturation of the protein bodies, and therefore can be important in the modification of the endosperm texture in *o2* genotypes. For instance, the 27 kDa γ-zein has been proposed as a modifier of the gene *o2* because QPM genotypes contain large amounts of this cysteine-rich zein storage protein (Geetha et al. 1991; Ortega and Bates, 1983; Wallace et al. 1990) and the 27 kDa γ-zein locus was linked to a significant RFLP probe used in BSA analysis for endosperm modification (Lopes et al., 1995). The increase in the amount of γ-zeins in QPMs is related with an increase the steady-state level of γ-zein transcripts and higher transcription of the γ-zein genes (Lopes et al., 1995). The γ-zeins and β-zeins seem to be involved in the formation of protein bodies by associating with α-zeins and δ-zeins and promoting their accumulation in the lumen of the rough endoplasmic reticulum (Bagga et al., 1997; Coleman et al., 1996; Geli et al., 1994; Kim et al., 2002; Lee, 1998; Woo et al., 2001). The 50 and 27 kDa γ-zeins remain at the surface of the protein bodies, while the 15 kDa β-zein and the 16-kDa γ-zein would be important in determining the organization of α-zeins and δ-zeins in the core of the protein body (Lending and Larkins, 1989; Woo et al., 2001). We detected both *cis*- and *trans*-acting eQTLs in endosperm 18 dap for relative levels of β-, δ– and γ-zein mRNAs in all maize chromosomes, except on chromosome 4. The colocalization of eQTLs with

QTL identified for endosperm texture supports the suggested role of γ-zein gene expression in the modification of *o2* endosperm. Furthermore, eQTLs mapped for γ-zeins colocalized with QTLs for essential amino acid contents.

It was on interest to us to identify the eQTLs that colocalized with the QTL hot spot on chromosome 6L. This QTL hot spot carry major QTLs for endosperm texture modification and its favorable allele was derived from the *o2* inbred B73*o2* (Chapter III). We found eight eQTLs with peaks in the bins 6.05-6.07 (82-134 cM), explaining variation for the transcripts aGL18, b32/RiP, γZN16, γZN27, γZN50, sbeI and ssIIa. In all cases, except by the eQTL for b32/RiP, the allele that increases the mRNA abundance derived from B73*o2*. Likewise, 11 eQTLs with peaks in bins 3.05-3.07 (81-117 cM) explaining variation for the transcripts αZN19B1, αZN22Z1, αZN22Z3, αZN22Z5, eEF1Ac, LKR/SDH, sbeI, ssIIa, idbeI, and PPDK colocalized with a QTL hot spot for texture modification on chromosome 3L whose favorable allele derived from CML161. The allele derived from CML161 increased the levels of mRNA for the 19- and 22 kDa α-zeins, and for the LKR/SDH; whereas decreased mRNA levels of the starch-related enzymes (sbeI, ssIIa, and idbeI), and the PPDK. The colocalization of QTLs detected for expression traits and classical phenotypic traits, and the agreement in the direction of the effect of the parental alleles of these two QTLs hot spots are good evidence of the effectiveness of our approach to map modifier genes for the *O2* gene.

In addition, both individual CIM and MT-CIM analysis pointed out to other genomic locations that might be associated with endosperm modification. For instance, it is noteworthy to mention the eQTLs explaining variation for several α-zeins and one δ-

zein on chromosome 9L, or the colocalization of amino acid QTLs with eQTLs for starch-related transcripts on chromosome 2L, and with eQTLs for γ-zeins in 7L and 8S.

Although not statistically significant, we observed variation for the abundance of *O2* transcript in the population of RILs. By sequencing both genomic DNA and mRNA, we have confirmed that both parental inbreds (B73*o2* and CML161) share identical copies of a non-functional allele of the *O2* gene (Cruz-Vela et al., 2007), ruling out the possibility that observed differences were the result of allelic differences at the *O2* locus. The observed variation could be related to the observation that both B73*o2* and CML161 carry a Bg-like transposable element in the promoter region of the O2 gene (C. Cruz-Vela, A. Gutierrez-Rojas, J. Betran, L. Caulk, S. Bhatnagar and M. Menz, unpublished results). The eQTL for O2 located on chromosome 4L colocalized with eQTLs for aZN22Z4 and aZN22Z5, but not with QTLs for endosperm texture or amino acid contents. Interestingly, the allele that increased the levels of O2 mRNA and the α-zein transcripts was derived from B73*o2*.

We found evidence of an association of between the expression of the zein genes and other endosperm genes such as the ones for enzymes of the starch biosynthesis and the modification of the texture of the *o2* endosperm. Due to the importance of both starch and zeins for the role of the endosperm as storage tissue, it is not unlikely that the zein and starch biosynthetic pathways are interconnected and coordinately regulated (Barbosa and Glover, 1978 Giroux et al.,1994; Kim and Giultinan, 1999; Tsai et al., 1978). A possibility is that the modification of the opaque endosperm requires bypassing the effects of the *o2* mutation, by an unknown mechanism that increase the levels of the

zeins to a point in which the protein bodies would resemble the ones of the normal endosperm. However, scanning electron microscopy studies have demonstrated that modified endosperm lack normal-sized protein bodies (Gibbon et al., 2003; Robutti et al., 1974b, Wasson and Hoseney, 1973; Wolf et al., 1969). Another option that has been suggested is that endosperm modification is reached by modifying the pattern of starch deposition to obtain tightly packed starch granules (Gibbon et al., 2003; Robutti et al., 1974b). For the later option, the increase of the levels of zeins would not be needed unless the modification of structure and organization of the starch granules needed to be accompanied by special type of protein bodies in which the composition of the zeins (i.e. increased 27 kDa γ-zein) rather than their size would help to define the packing density (Geetha et al., 1991; Gibbon et al., 2003). Recently, the role of the enzyme CyPPDK1 in the endosperm has been revised. The PPDK enzymes catalyze the interconversion of ATP, Pi, and pyruvate to AMP, PPi, and PPE. The expression of *CyPPDK1* is regulated by the O2 protein (Maddaloni et al., 1996). Several hypothesis have been proposed for the function of such enzymatic activities in the endosperm of cereals (Chastain et al., 2006; Mechin et al., 2007; Kang et al., 2005). The functions of *PPDK1* in endosperm would be associated with the production of precursors for amino acid synthesis, and the determination of the balance of starch versus protein synthesis (Chastain et al., 2006; Mechin et al., 2007). Clustering of PPDK transcripts and starch-related enzymes suggested a significant correlation in the genetic control of their expression. Furthermore, one major eQTL/QTL hot spot on chromosome 6L is located in the proximity of the *CyPPDK1* locus. Another eQTL/QTL hot spot on chromosome 3L

affects the levels of PPDK transcripts while having opposite affects on the levels of the transcripts of α-zeins and of enzymes in the starch biosynthesis. Finally, a third region on chromosome 6S carries eQTLs for the PPDK transcript, and for three γ-ZNs, αZN19B3, eEF1Ab, bZN15, and ssIIa. Our results support the hypothesis of a role for the cyPPDK1 in the regulation of the starch-protein balance in the developing endosperm of maize. By affecting this balance, the cyPPDK and other possible regulators of its expression, besides the O2 transcription factor, become candidates for modifiers for the gene *o2*.

Unfortunately, we cannot report conclusive findings of the location and characterization of the actual genes acting as modifiers of the allele *o2* in our population of RILs. Nevertheless, the combination of classical QTL mapping and expression traits mapping provided clues about the genomic regions that are segregating in association with the phenotypic traits, and the association of the same regions in the regulation of the expression of a group of endosperm genes.

CHAPTER V

SUMMARY

In order to increase the current understanding on the genetic mechanism of the modification of the soft *o2* endosperm, a population of 146 recombinant inbred lines derived from the *o2* inbred line B73o2 and the QPM inbred line CML161 was developed to study the segregation of modifier loci and their effect on the phenotype.

In Chapter II, it was described how the population of RILs was evaluated in two Texas locations during 2004-2006. Conventional methodologies to measure endosperm texture and opacity were used in addition to an image-based analysis that attempts to remove any bias added to the scores by the observer. Besides, a microbiological-based assay was used to quantify contents of the essential amino acids lysine, tryptophan, and methionine in mature kernels of the RILs. The analysis of the results suggested significant differences in the parental inbreds for all traits. Furthermore, the population of RILs showed also significant differences for all traits, and was segregating for all traits with approximately normal distributions in most cases. High correlations between the traits were observed and considerably high heritability estimates. A negative correlation was observed between the traits measuring endosperm texture modification and amino acid contents. Expected gains from selection were estimated for all traits, and favorable gain is expected for both endosperm texture and amino acid contents if selection is performed for both traits simultaneously. Otherwise, calculated correlated responses from selection suggested that selecting only for endosperm texture will

substantially reduce the obtained gain in lysine obtained by replacing *O2* alleles with non-functional *o2*.

In Chapter III, the construction of a genetic map for the population of RILs is described. The final map contained 180 SSR markers, represented the 10 chromosomes of maize, and showed an adequate marker density for QTL studies. The phenotypic traits for endosperm texture, and the contents of the amino acids Lys, Trp, and Met were analyzed across locations. Composite interval mapping and multiple interval mapping analysis were applied to identify genomic regions controlling the variation observed in the population of RILs. QTLs were identified for all traits in different maize chromosomes. Regions containing several QTLs for different traits or QTL hotspots were located for endosperm texture traits in chromosomes 3S, 3L, 5L, 6L, and 8L. QTL hotspots for amino acid contents were locate on chromosomes 7L and 8S. There was not evidence of overlapping QTL regions for both classes of traits. The number of QTLs detected suggested that in the population of RILs derived from B73*o2* and CML161, the modification of the *o2* phenotype behaved as a polygenic trait. Favorable effects were observed deriving from both parents. The comparison of the QTL position with consensus genetic maps of maize allowed suggesting potential candidate genes for genes involved in the modification.

In chapter IV, the pattern of protein composition of mature endosperm of the parental inbreds B73o2 and CML161, and the inbred B73 was studied by solubility fractioning. A strong reduction of the proportion of α-zeins was observed in the *o2* inbred, and intermediate levels of this subfamily of zeins were present in the QPM line

when compared to the normal inbred B73. Global analysis of gene expression in endosperm 18 dap showed a prominent effect of the *o2* mutation in the overall gene expression of the inbred B73. Several thousand features were found as differentially expressed between the *o2* line and the QPM line, with features corresponding to the zeins genes being well represented. Based on the global expression results and on previous published reports, 29 genes were selected to be analyzed in developing endosperm of a subset of 93 RILs. Among the 29 genes different members of the zein family were included, together with genes involved in the biosynthesis of starch, genes under the regulation of the O2 protein, and other storage protein genes. The results of the analysis of gene expression in the 29 genes were used as traits for expression QTL mapping. Expression QTLs were located for all the 29 gene transcripts. Several eQTLs were identified as *cis*-acting or located where the encoding locus has been mapped. A bigger number of eQTLs were located acting *in trans*, suggesting a more complex genetic control of the expression of several endosperm genes, including the zeins. Clustering analysis allowed creating groups of genes with similar expression patterns in the population of RILs, and the mapping of QTLs for multiple traits based on cluster grouping. The 19 kDa α-zeins, δ-zeins, and γ-zeins formed their own clusters, and the 22 kda α-zeins formed a cluster with the b32/RiP and the LKR/SDH transcripts. The conformation of groups by gene families and function suggested a close relationship between starch-related enzymes, members of the eEF1A family, and the cyPPDK. The colocalization of QTLs mapped for endosperm texture modification or amino acid

contents, with eQTLs located for some of the 29 genes suggested a correlation in the genetic control of phenotypes and expression traits.

The results presented represent an advance in the study of the genetics of endosperm modification in the kernel of maize. This was a unique study in which a population derived from inbred lines identified for breeding purposes was evaluated in multi-location experiment for both texture and amino acid traits. Moreover, classical and expression QTL mapping analysis allowed the identification of genomic regions controlling variation associated with the modification of the endosperm.

REFERENCES

Abe, S., K. Azama, H. Sugimoto, and E. Davies. 2003. Protein accumulation in the maize endosperm: Role of polyribosomes and the cytoskeleton. Plant Physiol. Biochem. 41:125-131.

Annapurna, S., and G.M. Reddy. 1971. Modified opaque maize for possible applied use. Curr. Sci. 40:581-582.

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

Asins, M.J. 2002. Present and future of quantitative trait locus analysis in plant breeding. Plant Breeding 121:281-291.

Azevedo, R.A., M. Lancien, and P.J. Lea. 2006. The aspartic acid metabolic pathway, an exciting and essential pathway in plants. Amino Acids 30:143-162.

Bagga, S., H.P. Adams, F.D. Rodriguez, J.D. Kemp, and C. Sengupta-Gopalan. 1997. Coexpression of the maize delta-zein and beta-zein genes results in stable accumulation of delta-zein in endoplasmic reticulum-derived protein bodies formed by beta-zein. Plant Cell 9:1683-1696.

Bantte, K., and B.M. Prasanna. 2004. Endosperm protein quality and kernel modification in the quality protein maize inbred lines. J. Plant Biochem. Biotech. 13:57-60.

Barbosa, H.M., and D.V. Glover. 1978. Genes and gene interactions affecting protein and lysine content in the endosperm of maize. Brazilian J. Genet. 1:29-39.

Bass, H.W., C. Webster, G.R. Obrian, J.K.M. Roberts, and R.S. Boston. 1992. A maize ribosome-inactivating protein is controlled by the transcriptional activator Opaque-2. Plant Cell 4:225-234.

Beavis, W.D. 1998. QTL analyses: Power, precision, and accuracy, p. 145-162, *In* A. H. Paterson, ed. Molecular Dissection of Complex Traits. CRC Press, New York.

Belousov, A.A. 1987. Genetic analysis of modified endosperm texture in *opaque-2* maize (In Russian, with English abstract.) Genetika 23:677-685.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate. A practical and powerful approach to multiple testing. J. of the Royal Statistical Soc. Series B-Methodological 57:289-300.

Benner, M.S., R.L. Phillips, J.A. Kirihara, and J.W. Messing. 1989. Genetic analysis of methionine-rich storage protein accumulation in maize. Theor. Appl. Genet. 78:761-767.

Berberich, T., K. Sugawara, M. Harada, and T. Kusano. 1995. Molecular cloning, characterization and expression of an elongation factor 1 alpha gene in maize. Plant Mol. Biol. 29:611-615.

Bernard, L., P. Ciceri, and A. Viotti. 1994. Molecular analysis of wild-type and mutant alleles at the *Opaque-2* regulatory locus of maize reveals different mutations and types of O2 products. Plant Mol. Biol. 24:949-959.

Bernardo, R. 2002. Breeding for quantitative traits in plants. Stemma Press, Woodbury, MN.

Bhatnagar, S., F.J. Betran, and D.K. Transue. 2003. Agronomic performance, aflatoxin accumulation and protein quality of subtropical and tropical QPM hybrids in southern US. Maydica 48:113-124.

Bjarnason, M., and S.K. Vasal. 1992. Breeding of Quality Protein Maize (QPM), p. 181-216, *In* J. Janick, ed. Plant Breeding Reviews, Vol. 9. John Wiley & Sons, New York.

Boston, R.S., E.B.P. Fontes, B.B. Shank, and R.L. Wrobel. 1991. Increased expression of the maize immunoglobulin binding-protein homolog b-70 in three zein regulatory mutants. Plant Cell 3:497-505.

Brem, R.B., and L. Kruglyak. 2005. The landscape of genetic complexity across 5,700 gene expression traits in yeast. Proc. Natl. Acad. Sci. U. S. A. 102:1572-1577.

Brochetto-Braga, M.R., A. Leite, and P. Arruda. 1992. Partial purification and characterization of Lysine-Ketoglutarate Reductase in normal and *opaque-2* maize endosperms. Plant Physiol. 98:1139-1147.

Buckler, E. 2006. Large scale genotyping and association studies. *In* Proc. ASA-CSSA-SSSA International Annual Meetings. 12-16 November 2006, Indianapolis, IN.

Burnett, R.J., and B.A. Larkins. 1999. *Opaque2* modifiers alter transcription of the 27-kDa gamma-zein genes in maize. Mol. Gen. Genet. 261:908-916.

Burr, B., and F.A. Burr. 1991. Recombinant inbreds for molecular mapping in maize - theoretical and practical considerations. Trends Genet. 7:55-60.

Carneiro, N.P., P.A. Hughes, and B.A. Larkins. 1999. The eEFlA gene family is differentially expressed in maize endosperm. Plant Mol. Biol. 41:801-813.

Chandrashekar, A., and H. Mazhar. 1999. The biochemical basis and implications of grain strength in sorghum and maize. J. Cereal Sci. 30:193-207.

Chastain, C.J., J.W. Heck, T.A. Colquhoun, D.G. Voge, and X.Y. Gu. 2006. Posttranslational regulation of pyruvate, orthophosphate dikinase in developing rice (Oryza sativa) seeds. Planta 224:924-934.

Churchill, G.A., and R.W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genetics 138:963-971.

Ciceri, P., S. Castelli, M. Lauria, B. Lazzari, A. Genga, L. Bernard, M. Sturaro, and A. Viotti. 2000. Specific combinations of zein genes and genetic backgrounds influence the transcription of the heavy-chain zein genes in maize *opaque-2* endosperms. Plant Physiol. 124:451-460.

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

Clark, L.G., W.P. Zhang, and J.F. Wendel. 1995. A phylogeny of the grass family (Poaceae) based on ndhF sequence data. Systematic Bot. 20:436-460.

Coleman, C.E., J.M. Dannenhoffer, and B.A. Larkins. 1997. The prolamin proteins of maize, sorghum and coix, p. 257-288, *In* B. A. Larkins and I. K. Vasil, eds. Cellular and Molecular Biology of Plant Seed Development. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Coleman, C.E., E.M. Herman, K. Takasaki, and B.A. Larkins. 1996. The maize gamma-zein sequesters alpha-zein and stabilizes its accumulation in protein bodies of transgenic tobacco endosperm. Plant Cell 8:2335-2345.

Consoli, L., and C. Damerval. 2001. Quantification of individual zein isoforms resolved by two-dimensional electrophoresis: Genetic variability in 45 maize inbred lines. Electrophoresis 22:2983-2989.

Cord-Neto, G., J.A. Yunes, M.J. Dasilva, A.L. Vettore, P. Arruda, and A. Leite. 1995. The involvement of Opaque-2 on beta-prolamin gene regulation in maize and Coix suggests a more general role for this transcriptional activator. Plant Mol. Biol. 27:1015-1029.

Corona, L., F.N. Owens, and R.A. Zinn. 2006. Impact of corn vitreousness and processing on site and extent of digestion by feedlot cattle. J. Anim. Sci. 84:3020-3031.

Costa, L.M., J.F. Gutierrez-Marcos, and H.G. Dickinson. 2004. More than a yolk: The short life and complex times of the plant endosperm. Trends Plant Sci. 9:507-514.

Cruz-Vela, C., A. Gutierrez-Rojas, L. Caulk, J. Betran, and M. Menz. 2007. Molecular characterization of an *opaque-2* mutant allele in maize (*Zea mays* L.). p. 190. *In Proc.* Plant and Animal Genome XV, San Diego, CA.

Damerval, C., and M. Le Guilloux. 1998. Characterization of novel proteins affected by the *o2* mutation and expressed during maize endosperm development. Mol. Gen. Genet. 257:354-361.

Dannenhoffer, J.M., D.E. Bostwick, E. Or, and B.A. Larkins. 1995. *Opaque-15*, a maize mutation with properties of a defective *opaque-2* modifier. Proc. Natl. Acad. Sci. USA 92:1931-1935.

Das, O.P., and J.W. Messing. 1987. Allelic variation and differential expression at the 27-Kilodalton zein locus in maize. Mol. Cell. Biol. 7:4490-4497.

Demason, D.A. 1997. Endosperm structure and development, p. 73–115, *In* B. A. Larkins and I. K. Vasil, eds. Cellular and Molecular Biology of Plant Seed Development. Kluwer Academic Publishers, Dordrecht, The Netherlands.

DeVicente, M.C., and S.D. Tanksley. 1993. QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134:585-596.

Doehlert, D.C., and T.M. Kuo. 1994. Gene expression in developing kernels of some endosperm mutants of maize. Plant Cell Physiol. 35:411-418.

Dolfini, S.F., M. Landoni, C. Tonelli, L. Bernard, and A. Viotti. 1992. Spatial regulation in the expression of structural and regulatory storage protein genes in *Zea mays* endosperm. Dev. Genet. 13:264-276.

Dombrink-Kurtzman, M.A., and J.A. Bietz. 1993. Zein composition in hard and soft endosperm of maize. Cereal Chem. 70:105-108.

Dudley, J.W., D.E. Alexander, and R.J. Lambert. 1973. El mejoramiento genetico del maiz de proteina modificada, p. 127-143 (In Spanish.) In L. F. Bauman, et al., eds. Maiz de Alta Calidad Proteinica. Editorial Limusa, Mexico D.F.

Duvick, D.N. 1961. Protein granules of maize endosperm cells. Cereal Chemistry 38:374.

Emerson, R.A., G.W. Beadle, and A.C. Fraser. 1935. A summary of linkage studies in maize. Cornell Univ. Agric. Exp. Sta. Mem. 180:3-83.

Esen, A. 1987. A Proposed nomenclature for the alcohol-soluble proteins (Zeins) of maize (*Zea mays* L). J. Cereal Sci. 5:117-128.

Falconer, D.S., and T.F.C. Mackay. 1996. Introduction to Quantitative Genetics. Longman Inc., New York.

FAOSTAT. 2007. FAOSTAT Agriculture Data, Food and Agriculture Organisation of the United Nations [online]. Available at http://faostat.fao.org/site/339/default.aspx (Verified 18 Sept 2007).

Fontaine, J., B. Schirmer, and J. Horr. 2002. Near-infrared reflectance spectroscopy (NIRS) enables the fast and accurate prediction of essential amino acid contents. 2. Results for wheat, barley, corn, triticale, wheat bran/middlings, rice bran, and sorghum. J. Agric. Food Chem. 50:3902-3911.

Fontes, E.B.P., B.B. Shank, R.L. Wrobel, S.P. Moose, G.R. Obrian, E.T. Wurtzel, and R.S. Boston. 1991. Characterization of an immunoglobulin binding-protein homolog in the maize *floury-2* endosperm mutant. Plant Cell 3:483-496.

Galili, G. 2002. New insights into the regulation and functional significance of lysine metabolism in plants. Annu. Rev. Plant Biol. 53:27-43.

Gardiner, J., S. Schroeder, M.L. Polacco, H. Sanchez-Villeda, Z.W. Fang, M. Morgante, T. Landewe, K. Fengler, F. Useche, M. Hanafey, S. Tingey, H. Chou, R. Wing, C. Soderlund, and E.H. Coe. 2004. Anchoring 9,371 maize expressed sequence tagged unigenes to the bacterial artificial chromosome contig map by two-dimensional overgo hybridization. Plant Physiol. 134:1317-1326.

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. J. Agric. Food Chem. 47:1268-1275.

Geetha, K.B., C.R. Lending, M.A. Lopes, J.C. Wallace, and B.A. Larkins. 1991. *Opaque-2* modifiers increase gamma-zein synthesis and alter its spatial distribution in maize endosperm. Plant Cell 3:1207-1219.

Geli, M.I., M. Torrent, and D. Ludevid. 1994. Two structural domains mediate two sequential events in gamma-zein targeting protein endoplasmic reticulum retention and protein body formation. Plant Cell 6:1911-1922.

Gibbon, B.C., and B.A. Larkins. 2005. Molecular genetic approaches to developing quality protein maize. Trends Genet. 21:227-233.

Gibbon, B.C., X.L. Wang, and B.A. Larkins. 2003. Altered starch structure is associated with endosperm modification in Quality Protein Maize. Proc. Natl. Acad. Sci. U. S. A. 100:15329-15334.

Gibson, G., and B. Weir. 2005. The quantitative genetics of transcription. Trends Genet. 21:616-623.

Giroux, M.J., C. Boyer, G. Feix, and L.C. Hannah. 1994. Coordinated regulation of storage product genes in the maize endosperm. Plant Physiol. 106:713-722.

Glass, G.V. 1966. Testing homogeneity of variances. Am. Educational Res. J. 3:187-190.

Glover, D.V., and E.T. Mertz. 1987. Corn, p. 183-336, *In* R. A. Olson and K. J. Frey, eds. Nutritional quality of cereal grains: Genetic and agronomic improvement. ASA-CSSA-SSSA, Madison, WI.

Griffiths, J., and J. Bryan. 1987. The Climates of Texas Counties. The University of Texas, Austin.

Guignard, L. 1899. Sur les antheroziodes et la double copulation sexuelle chez les vegetaux angiospermes (In French.) Comptes Rendus de l'Academie des Science de Paris 128:864–871.

Habben, J.E., A.W. Kirleis, and B.A. Larkins. 1993. The origin of lysine-containing proteins in *opaque-2* maize endosperm. Plant Mol. Biol. 23:825-838.

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. Proc. Natl. Acad. Sci. U. S. A. 92:8640-8644.

Haberer, G., S. Young, A.K. Bharti, H. Gundlach, C. Raymond, G. Fuks, E. Butler, R.A. Wing, S. Rounsley, B. Birren, C. Nusbaum, K.F.X. Mayer, and J. Messing. 2005. Structure and architecture of the maize genome. Plant Physiol. 139:1612-1624.

Hadi, G. 2004. Effect of the length of the kernel filling period and the kernel filling rate on the grain yield of maize under different water supply conditions. Cereal Res. Comm. 32:465-470.

Hamilton, T.S., B.C. Hamilton, B.C. Johnson, and H.H. Mitchell. 1951. The dependence of the physical and chemical composition of the corn kernel on soil fertility and cropping system. Cereal Chem. 28:163-176.

Harn, C., M. Knight, A. Ramakrishnan, H.P. Guan, P.L. Keeling, and B.P. Wasserman. 1998. Isolation and characterization of the zSSIIa and zSSIIb starch synthase cDNA clones from maize endosperm. Plant Mol. Biol. 37:639-649.

Hartings, H., M. Maddaloni, N. Lazzaroni, N. Difonzo, M. Motto, F. Salamini, and R. Thompson. 1989. The O2 gene which regulates zein deposition in maize endosperm encodes a protein with structural homologies to transcriptional activators. EMBO J. 8:2795-2801

Henry, A.M., D. Manicacci, M. Falque, and C. Damerval. 2005. Molecular evolution of the *Opaque-2* gene in *Zea mays* L. Journal of Molecular Evolution 61:551-558.

Hilton, H., and B.S. Gaut. 1998. Speciation and domestication in maize and its wild relatives: Evidence from the globulin-1 gene. Genetics 150:863-872.

Holland, J.B. 2006. Estimating genotypic correlations and their standard errors using multivariate restricted maximum likelihood estimation with SAS Proc MIXED. Crop Science 46:642-654.

Holland, J.B., W.E. Nyquist, and C.T. Cervantes-Martinez. 2003. Estimating and interpreting heritability for plant breeding: An update, p. 9-112 Plant Breeding Reviews, Vol. 22. John Wiley & Sons, New York.

Huang, S.S., W.R. Adams, Q. Zhou, K.P. Malloy, D.A. Voyles, J. Anthony, A.L. Kriz, and M.H. Luethy. 2004. Improving nutritional quality of maize proteins by expressing sense and antisense zein genes. J. Agric. Food Chem. 52:1958-1964.

Huang, S.S., D.E. Kruger, A. Frizzi, R.L. D'Ordine, C.A. Florida, W.R. Adams, W.E. Brown, and M.H. Luethy. 2005. High-lysine corn produced by the combination of enhanced lysine biosynthesis and reduced zein accumulation. Plant Biotech. J. 3:555-569.

Hunter, B.G., M.K. Beatty, G.W. Singletary, B.R. Hamaker, B.P. Dilkes, B.A. Larkins, and R. Jung. 2002. Maize opaque endosperm mutations create extensive changes in patterns of gene expression. Plant Cell 14:2591-2612.

James, M.G., D.S. Robertson, and A.M. Myers. 1995. Characterization of the maize gene Sugary1, a determinant of starch composition in kernels. Plant Cell 7:417-429.

Jansen, R.C., and J.P. Nap. 2001. Genetical genomics: The added value from segregation. Trends Genet. 17:388-391.

Jia, Z.Y., and S.Z. Xu. 2007. Mapping quantitative trait loci for expression abundance. Genetics 176:611-623.

Kang, H.G., S. Park, M. Matsuoka, and G.H. An. 2005. White-core endosperm floury endosperm-4 in rice is generated by knockout mutations in the C-4-type pyruvate orthophosphate dikinase gene (OsPPDKB). Plant Journal 42:901-911.

Kao, C.H., Z.B. Zeng, and R.D. Teasdale. 1999. Multiple interval mapping for quantitative trait loci. Genetics 152:1203-1216.

Khoo, U., and M.J. Wolf. 1970. Origin and development of protein granules in maize endosperm. Am. J. Bot. 57:1042-1050.

Kim, C.S., Y.M. Woo, A.M. Clore, R.J. Burnett, N.P. Carneiro, and B.A. Larkins. 2002. Zein protein interactions, rather than the asymmetric distribution of zein mRNAs on endoplasmic reticulum membranes, influence protein body formation in maize endosperm. Plant Cell 14:655-672.

Knapp, S.J., and W.C. Bridges. 1990. Using molecular markers to estimate quantitative trait locus parameters- power and genetic variances for unreplicated and peplicated progeny. Genetics 126:769-777.

Kim, K.N., D.K. Fisher, M. Gao, and M.J. Guiltinan. 1998. Molecular cloning and characterization of the amylose-extender gene encoding starch branching enzyme IIB in maize. Plant Mol. Biol. 38:945-956.

Kim, K.N., and M.J. Guiltinan. 1999. Identification of cis-acting elements important for expression of the starch-branching enzyme I gene in maize endosperm. Plant Physiol. 121:225-236.

Lai, J.S., N. Dey, C.S. Kim, A.K. Bharti, S. Rudd, K.F.X. Mayer, B.A. Larkins, P. Becraft, and J. Messing. 2004. Characterization of the maize endosperm transcriptome and its comparison to the rice genome. Genome Res. 14:1932-1937.

Lambert, R.J., Alexande.De, and J.W. Dudley. 1969. Relative performance of normal and modified protein (*opaque-2*) maize hybrids. Crop Science 9:242-243.

Lan, H., J.P. Stoehr, S.T. Nadler, K.L. Schueler, B.S. Yandell, and A.D. Attie. 2003. Dimension reduction for mapping mRNA abundance as quantitative traits. Genetics 164:1607-1614.

Lander, E., J. Abrahamson, A. Barlow, M. Daly, S. Lincoln, L. Newburg, and P. Green. 1987. Mapmaker a computer package for constructing genetic linkage maps. Cytogenetics Cell Genet. 46:642-642.

Landry, J., and T. Moureaux. 1970. Heterogeneity of glutelins of corn seeds: Selective extraction and composition of amino acids of 3 isolated fractions (In French, with English abstract). Bulletin de la Societe de Chimie Biologique 52:1021-1037.

Landry, J., S. Delhaye, and C. Damerval. 2004. Protein distribution pattern in floury and vitreous endosperm of maize grain. Cereal Chem. 81:153-158.

Lawton, J.W., and C.M. Wilson. 1987. Proteins of the kernel, p. 313-354, *In* P. J. White and L. A. Johnson, eds. Corn Chemistry and technology. American Association of Cereal Chemists, Inc., St. Paul, MN.

Lawrence, C.J., O.F. Dong, M.L. Polacco, T.E. Seigfried, and V. Brendel. 2004. MaizeGDB, the community database for maize genetics and genomics. Nucleic Acids Res. 32:D393-D397.

Lawrence, C.J., M.L. Schaeffer, T.E. Seigfried, D.A. Campbell, and L.C. Harper. 2007. MaizeGDB's new data types, resources and activities. Nucleic Acids Res. 35:D895-D900.

Lazzari, B., C. Cosentino, and A. Viotti. 2002. Gene products and structure analysis of wild-type and mutant alleles at the *Opaque-2* locus of *Zea mays*. Maydica 47:253-265.

Lee, D.H. 1998. Characterization of 27K zein as a transmembrane protein. J. Biochem. Mol. Biol. 31:196-200.

Lee, M., N. Sharopova, W.D. Beavis, D. Grant, M. Katt, D. Blair, and A. Hallauer. 2002. Expanding the genetic map of maize with the intermated B73 x Mo17 (IBM) population. Plant Mol. Biol. 48:453-461.

Lending, C.R., and B.A. Larkins. 1989. Changes in the zein composition of protein bodies during maize endosperm development. Plant Cell 1:1011-1023.

Leyva-Ovalle, O.R., A. Carballo, J.A. Mejia, and M.G. Vazquez. 2002. Digital image processing for endosperm texture estimation in maize lines (In Spanish.) Revista Fitotecnia Mexicana 25:355-365.

Li, Z.W., and H.N. Trick. 2005. Rapid method for high-quality RNA isolation from seed endosperm containing high levels of starch. BioTechniques 38:872-876.

Lipkovich, I., and E.P. Smith. 2002. Biplot and singular value decomposition macros for Excel. J. Statistical Software 7:1-15.

Liu, R., C. Vitte, J. Ma, A.A. Mahama, T. Dhliwayo, M. Lee, and J.L. Bennetzen. 2007. A GeneTrek analysis of the maize genome. Proc. Natl. Acad. Sci. U. S. A. 104:11844-11849.

Lizarraga-Guerra, R., B.C. Gibbon, and B.A. Larkins. 2006. Genetic analysis of *opaque2* modifier genes. p. 44. *In* Proc. 48th Annual Maize Genetics Conference Pacific Grove, California.

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 J. 10:617-624.

Lopes, M.A., and B.A. Larkins. 1995. Genetic analysis of *opaque2* modifier gene activity in maize endosperm. Theor. Appl. Genet. 91:274-281.

Lopes, M.A., K. Takasaki, D.E. Bostwick, T. Helentjaris, and B.A. Larkins. 1995. Identification of 2 *opaque2* modifier loci in quality protein maize. Mol. Gen. Genet. 247:603-613.

Lopez-Valenzuela, J.A., B.C. Gibbon, D.R. Holding, and B.A. Larkins. 2004. Cytoskeletal proteins are coordinately increased in maize genotypes with high levels of eEF1A. Plant Physiol. 135:1784-1797.

Lou, X.Y., J. Zhu, Q.D. Zhang, R.C. Zang, Y.B. Chen, Z.L. Yu, and Y.J. Zhao. 2005. Genetic control of the *opaque-2* gene and background polygenes over some kernel traits in maize (*Zea mays* L.). Genetica 124:291-300.

Maddaloni, M., G. Donini, C. Balconi, E. Rizzi, P. Gallusci, F. Forlani, S. Lohmer, R. Thompson, F. Salamini, and M. Motto. 1996. The transcriptional activator Opaque-2 controls the expression of a cytosolic form of pyruvate orthophosphate dikinase-1 in maize endosperms. Mol. Gen. Genet. 250:647-654.

Mahajan, R., and R. Singh. 1989. Properties of pyrophosphate-fructose-6-phosphate phosphotransferase from endosperm of developing wheat (*Triticum aestivum* L) grains. Plant Physiol. 91:421-426.

Mangelsdorf, P.C., and R.G. Reeves. 1939. The origin of Indian corn and its relatives. Bulletin No. 574, Texas Agric. Exp. Stn., College Station, TX.

Matsuoka, M. 1990. Structure, genetic mapping, and expression of the gene for pyruvate, orthophosphate dikinase from maize. J. Biol. Chem. 265:16772-16777.

Mazur, B., E. Krebbers, and S. Tingey. 1999. Gene discovery and product development for grain quality traits. Science 285:372-375.

Mechin, V., C. Thevenot, M. Le Guilloux, J.L. Prioul, and C. Damerval. 2007. Developmental analysis of maize endosperm proteome suggests a pivotal role for pyruvate orthophosphate dikinase. Plant Physiol. 143:1203-1219.

Melchinger, A.E., H.F. Utz, and C.C. Schon. 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383-403.

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

Mesa, D. 1957. Historia natural del maíz (In Spanish.) Separata de la Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales. 10:13-23.

Messing, J., A.K. Bharti, W.M. Karlowski, H. Gundlach, H.R. Kim, Y. Yu, F.S. Wei, G. Fuks, C.A. Soderlund, K.F.X. Mayer, and R.A. Wing. 2004. Sequence composition and genome organization of maize. Proc. Natl. Acad. Sci. U. S. A. 101:14349-14354.

Misra, P.S., E.T. Mertz, and D.V. Glover. 1976. Studies on corn proteins .9. Comparison of amino acid composition of Landry-Moureaux and Paulis-Wall endosperm fractions. Cereal Chem. 53:699-704.

Monjardino, P., A.G. Smith, and R.J. Jones. 2005. Heat stress effects on protein accumulation of maize endosperm. Crop Science 45:1203-1210.

Montes, J.M., H.F. Utz, W. Schipprack, B. Kusterer, J. Muminovic, C. Paul, and A.E. Melchinger. 2006. Near-infrared spectroscopy on combine harvesters to measure maize grain dry matter content and quality parameters. Plant Breeding 125:591-595.

Morison, S.E. 1963. The Journals and Other Documents of the Life of Christopher Columbus. Heritage Press, New York.

Moro, G.L., J.E. Habben, B.R. Hamaker, and B.A. Larkins. 1996. Characterization of the variability in lysine content for normal and *opaque2* maize endosperm. Crop Science 36:1651-1659.

Moro, G.L., M.A. Lopes, J.E. Habben, B.R. Hamaker, and B.A. Larkins. 1995. Phenotypic effects of *opaque2* modifier genes in normal maize endosperm. Cereal Chem. 72:94-99.

Motto, M., C. Lorenzoni, E. Gentinetta, T. Maggiore, and F. Salamini. 1978. Expected genetic gain for protein related traits and allocation of resources in a modified *opaque-2* synthetic. Maydica 23:35-44.

Muller, M., G. Dues, C. Balconi, F. Salamini, and R.D. Thompson. 1997. Nitrogen and hormonal responsiveness of the 22 kDa alpha-zein and b-32 genes in maize endosperm is displayed in the absence of the transcriptional regulator Opaque-2. Plant J. 12:281-291.

National Research Council. 1988. Quality-Protein Maize. National Academic Press, Washington, DC.

Nawaschin, S.G. 1898. Resultate einer Revision der Befruchtungsvogange bei *Lilium martagon* und *Fritillaria tenulla* (In German.) Bulletin de l'Academie des Sciences de Saint Petersbourg 9:377–382.

Olsen, O.A. 2004. Nuclear endosperm development in cereals and *Arabidopsis thaliana*. Plant Cell 16:S214-S227.

Openshaw S, and E. Frascaroli. 1997. QTL detection and marker assisted selection for complex traits in maize. Proc. Corn Sorghum Ind. Res. Conf. 52:44–53.

Ortega, E.I., and L.S. Bates. 1983. Biochemical and agronomic studies of two modified hard-endosperm *opaque-2* maize (*Zea mays* L) populations. Cereal Chem. 60:107-111.

Osborne, T.B., L.B. Mendel, E.L. Ferry, and A.J. Wakeman. 1914. Nutritive properties of proteins of the maize kernel. J. Biol. Chem. 18:1-16.

Paez, A.V., J.L. Helm, and M.S. Zuber. 1969. Lysine content of *opaque-2* maize kernels having different phenotypes. Crop Science 9:251-252.

Paulsen, M.R., and L.D. Hill. 1985. Corn quality factors affecting dry milling performance. J. Agric. Eng. Res. 31:255-263.

Perkins, D.N., D.J.C. Pappin, D.M. Creasy, and J.S. Cottrell. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551-3567.

Piperno, D.R., and K.V. Flannery. 2001. The earliest archaeological maize (*Zea mays* L.) from highland Mexico: New accelerator mass spectrometry dates and their implications. Proc. Natl. Acad. Sci. U. S. A. 98:2101-2103.

Pixley, K.V., and M.S. Bjarnason. 1993. Combining ability for yield and protein quality among modified endosperm *opaque-2* tropical maize inbreds. Crop Science 33:1229-1234.

Pixley, K.V., and M.S. Bjarnason. 2002. Stability of grain yield, endosperm modification, and protein quality of hybrid and open-pollinated quality protein maize (QPM) cultivars. Crop Science 42:1882-1890.

Pollacsek, M. 1970. Effect of a dominant suppressor gene on expression of *Opaque-2* gene in maize (In French, with English abstract.) Annales De L' Amelioration Des Plantes 20:337-343.

Pollacsek, M., M. Caenen, and M. Rousset. 1972. Demonstration of a second suppressor gene of gene *Opaque-2* in maize (In French, with English abstract). Annales De Genetique 15:173-176.

Pradilla, A., D.D. Harpstead, D. Sarria, F. Linares, and C.A. Francis. 1972. El maiz de alta calidad proteinica y la nutricion humana, p. 29-39, (In Spanish.) *In* L. F. Bauman, et al., eds. Maiz de Alta Calidad Proteinica. Editorial Limusa, Mexico D.F.

Prasanna, B.M., S.K. Vasal, B. Kassahun, and N.N. Singh. 2001. Quality protein maize. Curr. Sci. 81:1308-1319.

Purseglove, J. 1972. Tropical crops. Monocotyledons. Longman Group Limited, London, England.

Radwanski, E.R., and R.L. Last. 1995. Tryptophan biosynthesis and metabolism: Biochemical and molecular genetics. Plant Cell 7:921-934.

R Development Core Team. 2006. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rieseberg, L.H., M.A. Archer, and R.K. Wayne. 1999. Transgressive segregation, adaptation and speciation. Heredity 83:363-372.

Robutti, J.L., F.S. Borras, and G.H. Eyherabide. 1997. Zein compositions of mechanically separated coarse and fine portions of maize kernels. Cereal Chem. 74:75-78.

Robutti, J.L., R.C. Hoseney, and C.W. Deyoe. 1974a. Modified *opaque-2* corn endosperms.1. Protein distribution and amino acid composition. Cereal Chem. 51:163-172.

Robutti, J.L., R.C. Hoseney, and C.E. Wassom. 1974b. Modified *opaque-2* corn endosperms .2. Structure viewed with a scanning electron microscope. Cereal Chem. 51:173-180.

Salamini, F., B. Borghi, and Lorenzon.C. 1970. Effect of *opaque-2* gene on yield in maize. Euphytica 19:531-538.

Salvi, S., and R. Tuberosa. 2005. To clone or not to clone plant QTLs: Present and future challenges. Trends Plant Sci. 10:297-304.

SAS Institute. 2003. SAS/STAT guide for personal computers. Release 9.1.3. SAS Institute. Cary, NC.

Schadt, E.E., S.A. Monks, T.A. Drake, A.J. Lusis, N. Che, V. Colinayo, T.G. Ruff, S.B. Milligan, J.R. Lamb, G. Cavet, P.S. Linsley, M. Mao, R.B. Stoughton, and S.H. Friend. 2003. Genetics of gene expression surveyed in maize, mouse and man. Nature 422:297-302.

Schaeffer, M.L., H. Sanchez-Villeda, M. McMullen, and E.H. Coe. 2006. IBM2 2005 neighbors: 45,000 locus resource for maize. Proc. Plant and Animal Genome Conference XIV. 14-18 January 2006, San Diego, CA.

Scheuch, F., and C.A. Francis. 1973. Sistemas de seleccion para aumentar la capacidad proteinica del maiz harinoso con el gene opaco-2, p. 109-26, (In Spanish.) In L. F. Bauman, et al., eds. Maiz de Alta Calidad Proteinica. Editorial Limusa, Mexico D.F.

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 DNA. Proc. Natl. Acad. Sci. U. S. A. 87:46-50.

Schmidt, R.J., M. Ketudat, M.J. Aukerman, and G. Hoschek. 1992. Opaque-2 is a transcriptional activator that recognizes a specific target site in 22-Kd-zein genes. Plant Cell 4:689-700.

Scott, M.P., S. Bhatnagar, and J. Betran. 2004. Tryptophan and methionine levels in quality protein maize breeding germplasm. Maydica 49:303-311.

Sene, M., C. Thevenot, D. Hoffmann, F. Benetrix, M. Causse, and J.L. Prioul. 2001. QTLs for grain dry milling properties, composition and vitreousness in maize recombinant inbred lines. Theor. Appl. Genet. 102:591-599.

Shimada, A., and T.R. Cline. 1974. Limiting amino acids of Triticale for growing rat and pig. J. Anim. Sci. 38:941-946.

Smyth, G.K. 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Statistical Applications Genet. Mol. Biol. 3:1-25.

Soave, C., and F. Salamini. 1984. Organization and regulation of zein genes in maize endosperm. Philos. Trans. R. Soc. London, Ser. B 304:341-347.

Soave, C., L. Tardani, N. Difonzo, and F. Salamini. 1981. Zein level in maize endosperm depends on a protein under control of the *Opaque-2* and *Opaque-6* loci. Cell 27:403-410.

Song, R.T., V. Llaca, E. Linton, and J. Messing. 2001. Sequence, regulation, and evolution of the maize 22-kD alpha zein in gene family. Genome Res. 11:1817-1825.

Song, R.T., and J. Messing. 2002. Contiguous genomic DNA sequence comprising the 19-kD zein gene family from maize. Plant Physiol. 130:1626-1635.

Song, R.T., and J. Messing. 2003. Gene expression of a gene family in maize based on noncollinear haplotypes. Proc. Natl. Acad. Sci. U. S. A. 100:9055-9060.

Sturtevant, A.J. 1899. Varieties of corn. USDA Office of Exp. Stn. Bull. 57. U.S. Gov. Print. Office, Washington, DC.

Swarup, S., M.C.P. Timmermans, S. Chaudhuri, and J. Messing. 1995. Determinants of the high-methionine trait in wild and exotic germplasm may have escaped selection during early cultivation of maize. Plant J. 8:359-368.

Thompson, G.A., and B.A. Larkins. 1989. Structural elements regulating zein gene expression. BioEssays 10:108-113.

Timmermans, M.C.P., T.P. Brutnell, and P.W. Becraft. 2004. The 46th annual maize genetics conference, unlocking the secrets of the maize genome. Plant Physiol. 136:2633-2640.

Tsai, C.Y., B.A. Larkins, and D.V. Glover. 1978. Interaction of *Opaque-2* Gene with starch-forming mutant genes on synthesis of zein in maize endosperm. Biochem. Genet. 16:883-896.

Upadyayula, N., J. Wassom, M.O. Bohn, and T.R. Rocheford. 2006. Quantitative trait loci analysis of phenotypic traits and principal components of maize tassel inflorescence architecture. Theor. Appl. Genet. 113:1395-1407.

Vasal, S.K. 1971. CIMMYT's quality protein program, p. 83-95. *In* Proc. International Maize Workshop, CIMMYT, El Batan, Mexico.

Vasal, S.K. 2000. The quality protein maize story. Food Nutrition Bulletin 21:445-450.

Vasal, S.K., E. Villegas, M. Bjarnason, B. Gelaw, and P. Goertz. 1980. Genetic modifiers and breeding strategies in developing hard endosperm *opaque-2* materials., p. 37-71, *In* W. G. Pollmer and R. H. Phipps, eds. Improvement of quality traits of maize for grain and silage use. Nijhoff, The Hague.

Verza, N.C., T.R.E. Silva, G.C. Neto, F.T.S. Nogueira, P.H. Fisch, V.E. de Rosa, M.M. Rebello, A.L. Vettore, F.R. da Silva, and P. Arruda. 2005. Endosperm-preferred expression of maize genes as revealed by transcriptome-wide analysis of expressed sequence tags. Plant Mol. Biol. 59:363-374.

Vicente-Carbajosa, J., S.P. Moose, R.L. Parsons, and R.J. 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. Proc. Natl. Acad. Sci. U. S. A. 94:7685-7690.

Vietmeyer, N.D. 2000. A drama in three long acts: The story behind the development of quality protein maize. DIversity 16:29-32.

Viotti, A., G. Cairo, A. Vitale, and E. Sala. 1985. Each zein gene class can produce polypeptides of different sizes. EMBO J. 4:1103-1110.

Vyn, T.J., and M. Tollenaar. 1998. Changes in chemical and physical quality parameters of maize grain during three decades of yield improvement. Field Crops Res. 59:135-140.

Wallace, J.C., M.A. Lopes, E. Paiva, and B.A. Larkins. 1990. New methods for extraction and quantitation of zeins reveal a high content of gamma-zein in modified *opaque-2* maize. Plant Physiol. 92:191-196.

Wang, S., C.J. Basten, and Z. Zeng. 2006. Windows QTL Cartographer 2.5. [Online]. Available by Department of Statistics, North Carolina State University, Raleigh, NC. http://statgen.ncsu.edu/qtlcart/WQTLCart.htm.

Wang, X.L., and B.A. Larkins. 2001. Genetic analysis of amino acid accumulation in opaque-2 maize endosperm. Plant Physiol 125:1766-1777.

Wang, X.L., 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 Physiol. 125:1778-1787.

Wassom, C.E., and R.C. Hoseney. 1973. Selection of high lysine corn using a scanning electron microscope. Crop Science 13:462-463.

Watson, S.A. 2003. Description, development, structure, and composition of the corn kernel, p. 69-106, *In* P. J. White and L. A. Johnson, eds. Corn Chemistry and Technology, second ed. American Association of Cereal Chemist, Inc., St. Paul, MN.

Weerakoon, L.K., D.F. Weber, and M.C. Schneerman. 1993. Estimating genetic relatedness of maize to 23 plant species by RFLP analysis. Maydica 38:231-237.

Wei, F., E. Coe, W. Nelson, A.K. Bharti, F. Engler, E. Butler, H. Kim, J.L. Goicoechea, M. Chen, S. Lee, G. Fuks, H. Sanchez-Villeda, S. Schroeder, Z. Fang, M. McMullen, G. Davis, J.E. Bowers, A.H. Paterson, M. Schaeffer, J. Gardiner, K. Cone, J. Messing, C. Soderlund, and R.A. Wing. 2007. Physical and genetic structure of the maize genome reflects its complex evolutionary history. PLoS Genet. 3:e123.

Wessel-Beaver, L., and R.J. Lambert. 1982. Genetic control of modified endosperm texture in *opaque-2* maize. Crop Science 22:1095-1098.

Wessel-Beaver, L., R.J. Lambert, and J.W. Dudley. 1985. Genetic variability and correlations in a modified endosperm texture *opaque-2* maize population. Crop Science 25:129-132.

West, M.A.L., K. Kim, D.J. Kliebenstein, H. van Leeuwen, R.W. Michelmore, R.W. Doerge, and D.A.S. Clair. 2007. Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis. Genetics 175:1441-1450.

Wilcox, D.S., B.W. Dove, D. McDavid, and D.B. Greer. 2002. UTHSCSA ImageTool. Release 3.0. Available at http://ddsdx.uthscsa.edu/dig/itdesc.html (posted 22 Feb. 2002; verified 18 Sept. 2007).

Wolf, M.J., U. Khoo, and Seckinge.Hl. 1969. Distribution and subcellular structure of endosperm protein in varieties of ordinary and high-lysine maize. Cereal Chem. 46:253-263.

Woo, Y.M., D.W.N. Hu, B.A. Larkins, and R. Jung. 2001. Genomics analysis of genes expressed in maize endosperm identifies novel seed proteins and clarifies patterns of zein gene expression. Plant Cell 13:2297-2317.

Wu, R.L., X.Y. Lou, C.X. Ma, X.L. Wang, B.A. Larkins, and G. Casella. 2002. An improved genetic model generates high-resolution mapping of QTL for protein quality in maize endosperm. Proc. Natl. Acad. Sci. U. S. A. 99:11281-11286.

Xingming, F., T. Jing, H. Bihua, and L. Feng. 2001. Analyses of combining ability and heterotic groups of yellow grain quality protein maize inbreds. Proc. 7th Eastern and Southern Africa Regional Maize Conference. Nairobi, Kenya. 11-15 February 2001.

Yamagata, T., H. Kato, S. Kuroda, S. Abe, and E. Davies. 2003. Uncleaved legumin in developing maize endosperm: Identification, accumulation and putative subcellular localization. J. Exp. Bot. 54:913-922.

Zeng, Z.B., J.J. Liu, L.F. Stam, C.H. Kao, J.M. Mercer, and C.C. Laurie. 2000. Genetic architecture of a morphological shape difference between two drosophila species. Genetics 154:299-310.

Zhang, X.L., C. Colleoni, V. Ratushna, M. Sirghle-Colleoni, M.G. James, and A.M. Myers. 2004. Molecular characterization demonstrates that the Zea mays gene sugary2 codes for the starch synthase isoform SSIIa. Plant Mol. Biol. 54:865-879.

APPENDIX

The appendix is a Microsoft Excel file accompanying the pdf file. This file contains a table with the differentially expressed features (q <0.05) downregulated in B73*o2* in comparison to B73 and upregulated in CML161 in comparison to B73*o2*.

VITA

Libardo Andres Gutierrez-Rojas received his Bachelor of Science degree in biology from the National University of Colombia in 1999. Right after graduating, he started working at the Colombian Corporation of Agricultural Research-CORPOICA, where he was involved in several scientific projects concerning applied biotechnology for the improvement of tropical crops such as banana, plantain, guava and maize. While working for CORPOICA, he completed in 2003 his Master of Science degree in biochemistry from the National University of Colombia. In the Fall of 2003, he entered the Molecular and Environmental Plant Sciences (MEPS) program at Texas A&M University to pursue his Ph.D. degree.

He is married to Cristina Cruz-Vela and has a daughter, Paula Andrea. He can be contacted at Institute for Plant Genomics and Biotechnology, 2123 TAMU-Borlaug Center, College Station, TX 77843. His email is andresguti@yahoo.com.