2003

Molecular and genetic characterization of the inheritance and expression of the synthetic porcine alpha-Lactalbumin transgene in maize (Zea mays L)

Earl H. Bicar
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Agricultural Science Commons, Agriculture Commons, Agronomy and Crop Sciences Commons, Genetics Commons, and the Molecular Biology Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Molecular and genetic characterization of the inheritance and expression of the synthetic porcine alpha-Lactalbumin transgene in maize (Zea mays L.)

by

Earl H. Bicar

A dissertation submitted to the graduate faculty in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Plant Breeding

Program of Study Committee:
Michael Lee, Major Professor
Paul M. Scott
Kendall R. Lamkey
Phillip Becraft
Kan Wang

Iowa State University

Ames, Iowa

2003
Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Earl H. Bicar

has met the doctoral requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program
### TABLE OF CONTENTS

**Acknowledgements** iv

**Chapter 1. General Introduction** 1
- Introduction 1
- Dissertation Organization 4
- Review of Literature 4
- References 22

**Chapter 2. Inheritance and expression patterns of a synthetic porcine α-Lactalbumin transgene in maize *Zea mays* L.** 37
- Abstract 37
- Introduction 38
- Materials and Methods 41
- Results 45
- Discussion 51
- References 57

**Chapter 3. Expression of the synthetic porcine α-Lactalbumin gene results in increased Lysine levels in transgenic maize kernels** 70
- Abstract 70
- Introduction 72
- Materials and Methods 74
- Results 80
- Discussion 87
- References 96

**Chapter 4. The inheritance and expression patterns of the selectable marker gene BAR in maize *Zea mays* L.** 110
- Abstract 110
- Introduction 112
- Materials and Methods 113
- Results and Discussion 117
- References 127

**Chapter 5. General Conclusions** 135

**Appendix** 140
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Michael Lee, my major professor and adviser, for his guidance and support throughout my graduate program, for providing an environment conducive to learning and for instilling in me skills for critical thinking. I also would like to thank Dr. Paul Scott for his advice and invaluable help in the conduct of my research. To the members of my POS committee, Drs. Kendall Lamkey, Kan Wang and Phillip Becraft, I am grateful for their contributions and help.

I am also grateful to Pioneer Hi-bred International for awarding me a fellowship and to the Institute of Plant Breeding, University of the Philippines for granting me a study leave to pursue graduate studies here in ISU. In the course of my stay here in Iowa, I have made friends with several people who have inspired and given me encouragement. I am grateful for their company and friendship. I also appreciate the help and support provided by members of the Lee Lab and Scott Lab.

I am most grateful to my family for their understanding and patience.
To my wife Cleofe and sons Vanko and Isaiah
Chapter 1. General Introduction

Introduction

Maize is not only an important source of food but is also a major feed crop in many parts of the world. Here in the US, about 60-80% of maize production is utilized as a component of livestock feed (USDA, 2002). Though it supplies more energy than any other grain, maize does not provide a well balanced diet because it is deficient in the essential amino acids Lysine and Tryptophan (National Research Council, 1988). To maximize the amount of protein and achieve a more balanced feed formulation, farmers and feed producers supplement maize feed with synthetic Lysine and Tryptophan thereby increasing expense.

Recent advances in genetic engineering and the improvements in transformation technology have provided new opportunities to improve the nutritional composition of the maize grain through the use of transgenes. Transgenes that express plant proteins have been successfully used to improve the nutritional quality of several crops (Chakraborty et al., 2000; Drakakaki et al., 2000; Lai and Messing, 2002). One drawback however of plant proteins is that they can be poorly digested in animal diets or cause allergenic reactions (Nordlee et al., 1996; Scott and Bicar, 2003). Milk proteins have been proposed as ideal proteins for expression in plants for nutritional improvement (Yang et al., 2002; Chong et al., 1997) not only because they are a natural component of animal diets but also milk proteins have excellent nutritional parameters (Matthews and Hughes, 1993; Renner, 1988). In an attempt to improve the nutritional quality of the maize grain, a codon adjusted version of a gene encoding the milk protein porcine α-lactalbumin was synthesized and different versions of this gene have been transformed into maize (Yang et al., 2002). The porcine α-
lactalbumin protein is a major component of sow's milk (Permyakov and Berliner, 2000). It has a well-balanced amino acid composition with high levels of Lysine and Tryptophan complementing the amino acid balance of maize. By expressing the porcine α-lactalbumin protein to high levels in the maize endosperm, the levels of Lysine in the grain is expected to improve as the plant responds to an increasing demand for this amino acid.

In a preliminary study, two versions of the synthetic porcine α-La transgene that differed in their subcellular targeting domains have been characterized (Yang et al., 2002). The porcine α-La protein accumulated in the endosperm of plants transformed with the version of the gene (P42) that contained a 27 kDa gamma-zein signal sequence from the 27 kDa gamma-zein gene in maize while no porcine α-La protein was detected in the endosperm of plants transformed with a version of the gene (P40) that did not contain this signal peptide sequence (Yang et al., 2002). Although the porcine α-La protein was detected, changes in the amino acid composition as a result of porcine α-La expression were not examined. It was postulated that the levels of porcine α-La that accumulated in the endosperm were not sufficient to impact the amino acid balance of the maize kernels. In addition, it was observed that the α-La transgene was transmitted from the T0 to the F1 generation, however, it was not established if the transgene can be inherited and expressed in succeeding generations.

In order to obtain optimal information on the expression levels of synthetic porcine α-La protein, three other constructs that differ in their promoters and subcellular targeting domains were made and transformed into maize callus by particle bombardment. These constructs were called pUbi/zsig/lac/KDEL, pZein/zsig/lac/KDEL and pZein/zsig/lac. T0
plants were obtained, however, the expression and inheritance patterns of the \(\alpha\)-La transgenes in succeeding generations were unknown. It is important to establish the expression and inheritance patterns of the porcine \(\alpha\)-La transgenes as segregation of transgene phenotype often exhibits distortion resulting from the abnormal expression of the transgene (Pawlowski and Somers, 1996). Transgenes can only be of practical use when they are consistently expressed and inherited in a predictable manner. Also, it was not established if the modified porcine \(\alpha\)-La transgenes in these constructs are functional. Moreover, the levels of porcine \(\alpha\)-La protein as well as the changes in amino acid composition as the result of the expression of the \(\alpha\)-La protein are not known. It is important to obtain this information because to appreciably alter the amino acid composition of proteins in the maize endosperm, the porcine \(\alpha\)-La must be expressed to high levels in the kernel endosperm.

In addition to the synthetic porcine \(\alpha\)-La transgene, the gene BAR, which confers herbicide resistance, was also transformed into maize as a selectable marker. In many transformation and genetic studies of transgenes, emphasis is focused usually on the expression of the transgene of interest while the expression and inheritance patterns of selectable marker genes are mostly ignored despite the potential of these genes to contribute to the expression patterns of the transgene of interest or endogenous genes (Matzke and Matzke, 1990; Matzke et al., 1993). Also, reduction in agronomic performance as result of the expression of selectable marker genes has been reported (Hucl et al., 1998a).

The objectives therefore of this study were to 1) assess the inheritance and expression patterns of the porcine \(\alpha\)-La transgenes for 3 generations; 2) quantify the levels of porcine
α-La protein that accumulate in the kernel endosperm; 3) determine the levels of Lysine and evaluate the associated changes in amino acid composition of proteins in the endosperm of kernels that express the porcine α-La protein and; 4) establish the inheritance and expression patterns of the gene BAR and its co-segregation with porcine α-La transgene.

**Dissertation Organization**

This dissertation is composed of 5 chapters. Chapter 1 includes the general introduction and review of literature. Chapters 2 to 4 are the “experiment” chapters which cover the different studies conducted. Chapter 2 focuses on the inheritance and expression patterns of the modified porcine α-La transgenes while Chapter 3 focuses on expression levels of the porcine α-La protein and the improvements in amino acid composition. Chapter 4 reports on the inheritance and expression patterns of the selectable marker gene BAR. Information from chapters 2-4 will comprise 2-3 papers for publication in refereed journals. Each of these chapters has its own introduction, materials and methods, results and discussion sections including separate references at the end of each chapter. Finally, chapter 5 presents the general conclusions of the whole study.

**Review of Literature**

This review of literature describes the current literature on the use of transgenes to improve the nutritional composition of crops. As a background, it will briefly cover the early attempts to improve maize grain quality. A portion of this review evaluates the milk protein α-La and the characteristics that make it an ideal protein to use to improve maize quality.
Information will also be presented on the genetic factors that may affect the successful expression of transgenes in a plant system including general observations on transgene inheritance and integration patterns. Some information that is not mentioned in this section are reviewed and presented in the introduction and discussion sections in the relevant chapters.

1. The amino acid limitations of maize and the early breeding efforts

Maize endosperm proteins are comprised of different fractions. Based on their stability, these can be classified into albumins (water soluble), globulins (soluble in saline solution), zeins or prolamines (soluble in alcohol) and glutelins (alkali soluble) (Osborne, 1924). The proportions of these various fractions are zeins 60%, glutelins 34%, albumins 3% and globulins 3%. Although zeins constitute 60% of the total endosperm proteins, zeins are deficient in two essential amino acids, Lysine and Tryptophan (Osborne and Clapp, 1908) and this deficiency limited the nutritional value of the maize grain (National Research Council, 1988).

Early efforts to improve maize grain quality utilized the opaque 2 (o2) mutation which dramatically increased the percent of Lysine in the grain (Mertz et al., 1964) but the soft starchy endosperm of the mutant kernel causes it to be susceptible to pests and mechanical damage (Ortega and Bates, 1983). These problems as well as the reduced yield and protein content of the seed, were largely responsible for preventing the so called “high Lysine” maize from gaining wide acceptance. However, o2 gene modifiers were identified that improved the phenotype of the mutant, giving it a normal appearance (Paez et al., 1969a). These were used to develop maize varieties with normal kernel hardness and protein
content as well as enhanced percentage of Lysine. These modified $\text{o}_2$ genotypes were designated as quality protein maize or QPM (Vasal et al., 1980; Villegas et al., 1992).

Despite the success of the QPM genotypes, problems remain that discourage breeders to use the modified $\text{o}_2$ system to improve or develop highly nutritious cultivars. The $\text{o}_2$ system requires conversion of inbred lines or cultivars into the modified $\text{o}_2$ genetic background (Vasal, 2001). The $\text{o}_2$ modifiers however, are a genetically complex system (Vasal et al., 1980). In addition to the triploid nature of the endosperm, $\text{o}_2$ modifier genes have incomplete and unstable expression (Or et al., 1993) and have been found to show variable penetrance in different genetic backgrounds (Belousov, 1987). This could hinder breeding efforts in converting inbred lines or cultivars into the modified $\text{o}_2$ background. Despite this complexity, improvement in maize grain quality continued using the $\text{o}_2$ mutation and its gene modifiers. Alternative procedures, such as recurrent selection, did not gain support despite the existence of considerable variation in Lysine content among different races and strains of maize (Aguirre et al., 1953; Paez et al., 1969b; Tello et al., 1965) which could allow recurrent selection as a procedure to improve Lysine content in normal endosperm maize. Zuber (1975) reported an increase in Lysine content in two populations after 3 cycles of S1 family recurrent selection. One of the populations, Logan County Composite, obtained Lysine levels equal to opaque-2 (Zuber, 1975).

2. Transgenic Approaches

2.1 Modifying the biosynthetic pathway

The advances in genetic engineering and the improvements in transformation technology have provided new opportunities to modify the nutritional composition of the
maize grain through the use of transgenes. One strategy to improve maize grain quality is to modify the amino acid biosynthetic pathway of the amino acid of interest by introducing feedback insensitive versions of rate limiting enzymes (Lundquist et al., 2001). This strategy has been used to improve the Lysine content of maize (Falco et al., 1993; Jones et al., 1998). The first step that is committed to Lysine biosynthesis is catalyzed by the enzyme dihydrodipicolinate synthase (dhps). This enzyme is feedback inhibited by cellular Lysine concentrations of less than 100 mM (Frisch et al., 1991) effectively controlling the accumulation of Lysine. By bypassing this cellular feedback regulation, the levels of Lysine can be potentially increased. A gene called \textit{dapA} which encodes a feedback insensitive form of dhps was isolated from the prokaryote \textit{Corynebacterium glutamicum} (Falco et al., 1993) and from maize (Bittel et al., 1996; Frisch et al., 1991). Maize transformed with \textit{dapA} contained high levels of free Lysine in the kernel while the fixed Lysine content remain unchanged (Jones et al., 1998). The free Lysine leached out from the plant tissue and were lost during boiling and other processing (Falco et al., 1995; Shaul and Galili, 1992). A similar increase in free Lysine was also detected in canola, soybeans and rice transformed with genes encoding the feedback insensitive version of dhps (Falco et al., 1995; Lee et al., 2001b). Another drawback to this approach is that since Lysine is not the only product of the biosynthetic pathway, intermediates can accumulate that can be potentially toxic to the plant or to the animal consuming the plant (Jones et al, 1995).

\textit{2.2 Expression of proteins rich in the target amino acid: The sink principle.}

Cells normally maintain the levels of free Lysine and other amino acids in a carefully regulated biosynthetic pathway. Rather than working against the natural regulation of amino
acid biosynthesis, an alternative approach to improve the amino acid composition is to take advantage of this regulation by creating a sink of the amino acid of interest by over-expressing a foreign protein that contains high levels of the target amino acid. As the foreign protein is expressed to high levels in the plant, the amount of the target amino acid is also expected to increase as the plant responds to an increased demand of this amino acid. This approach has been used successfully in several crops. For example, legumes are usually deficient in sulfur-containing amino acids such as Methionine and Cysteine. Expression of a Methionine-rich 2S albumin from Brazil nut (Bertholletia excelsa) at levels of 4% of the total seed protein resulted to a 33% increase in Methionine content in the seeds of canola (Altenbach et al., 1992a) while in lupins, expression of a gene from sunflower encoding seed albumin rich in Methionine and Cysteine, resulted to a 94% increase in Methionine content of lupin seeds. However, a 12% reduction in Cysteine was also observed while no significant changes were detected in the other amino acids or in the total nitrogen and sulfur of lupin seeds (Molvig et al., 1997). Additionally, the seed albumin gene AmAl from Amaranthus hypochondriacus when expressed in potato resulted to a 2-4 fold increase in most essential amino acids including Lysine, Methionine, Cysteine and Tyrosine, which are otherwise limited in potato. The increase in the levels of these amino acids was also accompanied by an increase in total protein content. An increase in the size and number of potato tubers was also observed in transgenic populations (Chakraborty et al., 2000). In maize, transformation with the Dzs10 gene which encodes a seed-specific high-methionine storage protein resulted to not only a significant increase in Methionine but also to a 23 % increase in Glycine content (Lai and Messing, 2002). Transgenes that express plant proteins have been used for nutritional improvement of several crops with varying degrees of success (Altenbach et al.,
1992a; Drakakaki et al., 2000; Hagan et al., 2003; Hood et al., 1997; Molvig et al., 1997; Sharma et al., 1998; Sindhu et al., 1997; Zheng et al., 1995).

One drawback of over-expressing plant proteins for nutritional improvement is that plant proteins can be poorly digested in animal diets or cause allergenic reactions such as the case of soybean transformed with the methionine-rich 2S albumin from Brazil nut (Bertholletia excelsa) protein, where commercialization did not occur because of the protein’s allergenic properties (Nordlee et al., 1996). Alternatively, genes encoding milk proteins have been proposed as an attractive target for introduction into plants for nutritional improvement because they are a natural component of animal diets and have excellent nutritional parameters (Gasser and Fraley, 1989; Matthews and Hughes, 1993). Among milk proteins, $\alpha$-La was found ideal for expression in plants because it has a well-balanced amino acid composition that complements not only the nutritional requirements of animal feed but also of the human diet (Renner, 1988; Robinson, 1986). The expression of $\alpha$-La in plants can supplement the deficiencies of Lysine in cereals and Methionine in legumes and since it is a calcium-binding protein, it may also serve to supply calcium as a nutrient (Matthews and Hughes, 1993; McKenzie and White, 1991). Additionally, $\alpha$-La has a digestibility of 99%, biological value of 94% and a protein efficiency ratio of 3.2 (Robinson, 1986) which are important parameters for maize intended as a major feed component. Also, $\alpha$-La is small in size which is conducive to the construction of a synthetic coding region using plant preferred codons (Permyakov and Berliner, 2000; Yang et al., 2002). A codon adjusted version of porcine $\alpha$-La gene has been transformed into maize cells in an attempt to improve the nutritional quality of maize. Plants that express the porcine
$\alpha$-La protein in the endosperm were obtained. However, the changes in amino acid composition in the endosperm of transformed kernels were not determined (Yang et al., 2002). In addition to porcine $\alpha$-lactalbumin, other milk proteins that have been transformed into plants for nutritional improvement include the human $\beta$-casein in potato (Chong et al., 1997b), human lactoferrin in tobacco (Mitra and Zhang, 1994) and human $\alpha$-lactalbumin in tobacco (Takase and Hagiwara, 1998).

3. Genetic considerations for expression of transgenes in plants

3.1 Codon usage

In expressing foreign proteins in plants, it is desirable to achieve the maximum level of expression that does not interfere with plant development. Several factors must be considered in order to maximize the expression of a non-plant protein. One important consideration is codon usage. In general, genes within a taxonomic group exhibit a similar codon usage pattern regardless of the function of these genes (Ikemura, 1985). Bias in codon usage within a gene in a single species influences the level of expression of the protein encoded by that gene. As the pattern of codon usage in plants is different from that of animals, expression of animal or a non-plant protein in plants requires most often a complete or partial modification of the coding sequences in order to achieve strong expression. For example, the *Bacillus thuringiensis cryIA(b)* toxin gene has high A+U bias in the corresponding mRNA. Complete modification of the coding sequence resulted in strong expression in transgenic tomato and tobacco (Perlak et al., 1991). Also, the gene encoding green fluorescent protein (GFP) was resynthesized to adapt its codon usage for expression in plants by increasing the frequency of codons with a C or G in the third position from 32 to
60%. Such alterations in the gfp coding sequence enhanced the expression of GFP in tobacco (Chiu et al., 1996; Rouwendal et al., 1997). Increased levels of protein expression after codon usage adjustment have been reported in other plants species (Adang et al., 1993; Hood et al., 1997; Horvath et al., 2000; Koziel et al., 1993; Sardana et al., 1996; Stewart et al., 1996).

3.2 Promoters

Promoters determine the tissue specificity and strength of gene transcription. Promoters are available for tissue/organ specific, temporal and/or induced expression. Promoters can either be constitutive or tissue specific, allowing developmental control of expression of novel proteins. Some of the promoters that have been used for foreign gene expression in maize include the ubiquitin promoter which is reported to confer high level of expression in all tissues including the endosperm (Christensen and Quail, 1996; Cornejo et al., 1993; Stoger et al., 1999). Seed specific promoters, derived from genes that encode proteins that are abundant in the seed, have also been reported to direct high levels of foreign protein expression in the seed. For example, the zein promoters from the 19 and 27 kDa gamma-zein genes in maize have been reported to confer strong endosperm-specific expression of transgenes (Torrent et al., 1997), although poor expression in other tissues apart from the endosperm has also been reported (Russell and Fromm, 1997b). On the other hand, the CaMV 35S promoter, often considered to be constitutively expressed, is reported to direct high levels of transgene transcription in maize (Hood et al., 1997). However, because of its viral origin, many transgenes controlled by 35S promoter became downregulated following CaMV infection leading to loss of transgene expression (Al-Kaff et al., 2000; Covey et al., 1997; Ratcliff et al., 1997).
3.3 Targeting and retention signals

The translation of all cellular proteins (except for a small set of polypeptides synthesized in chloroplast and mitochondria) starts in the cytosol. At some point, either concurrent with translation or after the polypeptide has been released in the cytosol, a protein is targeted to one of the several cellular compartments. All proteins, except those that remain in the compartment where they are translated, include one or more targeting domains that act as an address label for each protein to find their way to the correct subcellular location (Buchanan et al., 2000).

Targeting domains are usually short peptides or amino acid motifs (but can also be glycans). They are often located at the amino (N)-terminal end of a protein but may be present in the carboxyl (C)-terminus or elsewhere in the amino acid sequence (Buchanan et al., 2000). Different names apply to targeting domains, depending on the organelle to which a protein is being targeted. For example, proteins destined to the secretory pathway have at their N-terminus a signal peptide of 16-30 amino acids that directs the polypeptide chain to the ER membrane and translocates the entire protein into the lumen of the ER. Such signal peptides are found in nearly all secreted proteins as well as proteins that reside in the ER (Buchanan et al., 2000). Signal peptides of proteins that enter the secretory pathway differ from one another but share important structural features. A signal peptide typically consists of one or more positively charged amino acids followed by a structure of 6-12 hydrophobic amino acids and then several additional amino acids. The signal peptides of different secretory proteins are interchangeable, not only among plant proteins but also among plant and animal proteins. Furthermore, the presence of a signal peptide is sufficient to direct the secretion of a stable and correctly folded protein. Although the targeting domain is essential
for protein transport, it may not be a part of the active protein and proteases in the target location often remove the targeting domain to create a functional, mature polypeptide (Buchanan et al., 2000).

Target sequences have been found to play a vital role in the accumulation of foreign proteins. In plants, the default pathway for proteins transported through the ER is secretion (Denecke et al., 1990; Neuhaus and Rogers, 1998). Expression of foreign proteins in plants has been found to increase significantly when targeted to the secretory pathway than when targeted to the cytosol (Conrad and Fiedler, 1998). For example, antibodies lacking a signal peptide were expressed in the cytosol at very low levels while a significant improvement in expression levels were detected when a signal peptide was included (Conrad and Fiedler, 1998). Also, the synthetic porcine α-lactalbumin protein was not detected in maize endosperms transformed with a construct that did not encode a signal peptide from 27 kDa maize gamma-zein while the α-La protein accumulated in the endosperm transformed by α-La construct containing the 27 kDa gamma-zein signal sequence (Yang et al., 2002). In some cases, foreign proteins targeted to the cytosol were presumed to be toxic in transformed maize cells but when the protein was sequestered into a specific compartment or organelle or transporting it to extracellular matrix, higher expression levels were obtained (Hood et al., 1997; Mason et al., 1998).

Proteins can also be retained in the ER. ER-resident proteins in mammalian and yeast cells have a carboxyterminal domain consisting of four amino acids KDEL (Lys-Asp-Glu-Leu) or HDEL (His-Asp-Glu-Leu) (Pelham, 1989). It has been shown that these tetrapeptides are ER-retention signals and both are necessary and sufficient for the retention of ER-resident proteins in the lumen of the ER (Munro and Pelham, 1987). Inclusion of ER-
retention signals has been found to lead to higher accumulation of foreign proteins in plants. For example, C-terminal fusion of a signal KDEL (Lys-Asp-Glu-Leu) resulted to a 100-fold increase in expression levels of an scFv antibodies in plants (Schouten et al., 1996; Torres et al., 2001). Also, high expression levels of δ-zeins were detected in tobacco for δ-zein constructs that contain KDEL compared to those constructs that did not contain KDEL (Belluci et al., 2000). Other amino acid motifs such as HDEL (Pueyo et al., 1995) and SEKDEL (Chikwamba et al., 2002) have been reported to increase the levels of recombinant proteins in plants. Extended retention in the ER is thought to promote correct protein folding leading to higher stability and accumulation of proteins (Schouten et al., 1998).

4. Genetic factors affecting transgene expression

4.1 Transgene integration patterns

Transgenes introduced by particle bombardment generally exhibit: 1) the full length introduced transgene/s (Pawlowski and Somers, 1996); 2) transgene rearrangements that differ in size from the full-length introduced gene (Cooley et al., 1995; Fromm et al., 1990b; Gordon-Kamm et al., 1990); 3) occasional evidence for concatenation of introduced plasmids carrying the transgene (Finer et al., 1999); and 4) variation in copy number (hybridization intensity) among the full-length transgene and transgene-hybridizing fragments (Cooley et al., 1995; Wan and Lemaux, 1994).

Transgene rearrangements are detected by the appearance of additional hybridizing restriction fragments of altered molecular weights in Southern blot analyses (Pawlowski and Somers, 1996). These rearrangements of transgenic sequences can either be deletions of diagnostic restriction sites (Kartzke et al., 1990) or the ligation of the introduced DNA
causing insertions of DNA between restriction sites (Riggs and Bates, 1986). Concatenation of introduced plasmid DNA is a frequently observed form of transgene rearrangement (Riggs and Bates, 1986). Concatenation of plasmid DNA results in complexes of linearized copies, linked in head-to-head or head-to-tail fashion. This occurs prior to integration in the genome so that multiple copies of transgenic DNA sequences are co-integrated at a single genomic location (Finer et al., 1999; Kholi et al., 1999). Often these copies are separated by plant DNA (Kholi et al., 1998).

The number of transgene copies that are integrated in the genome varies widely (Cooley et al., 1995; Spencer et al., 1992; Wan and Lemaux, 1994). Commonly reported numbers range from 1-20 transgene copies. Though multiple copies are integrated, they usually co-segregate as a transgenic locus. This indicates that the transgene sequences integrate into the genome at either tightly linked loci or at a single locus rather than multiple randomly distributed integration sites throughout all of the chromosomes (Pawlowski and Somers, 1996). The transgenes however are not contiguous at the insertion locus as the transgene DNA is interspersed with sequences of plant genomic DNA (Cooley et al., 1995; Kholi et al., 1998; Pawlowski and Somers, 1996; Pawlowski et al., 1998; Sawasaki et al., 2001). Transgene integration patterns usually remain stable across generations (Wan and Lemaux, 1994; Weeks et al., 1993). However, recent reports indicate transgene instability at the DNA level brought about by recombination resulting in novel patterns (Choffness et al., 2001; Svitashev et al., 2002).

In co-transformation experiments, in which different transgenes carried on two separate plasmids were cointroduced into plant cells (one of the plasmids usually carries the selectable marker gene, i.e. BAR, while the other the plasmid carries the transgene of
interest), genes from both plasmids are found co-integrated into the same genomic transgene locus (Christou and Swain, 1990; Cooley et al., 1995). The frequencies of co-transformation when transgenes come from separate plasmids are similar to co-transformation frequencies when two transgenes are introduced on the same plasmid. It is not known whether there is any sequence or genomic region preference in genomic integrations of transgenes delivered by particle bombardment (Timland, 1996). Although in Agrobacterium-mediated T-DNA insertions, transgenes have a tendency to integrate at transcribed regions (Iglesias et al., 1997; Wang et al., 1995). Sequence analysis of transgene-plant genomic DNA border junctions in Arabidopsis (Sawasaki et al., 1998), rice (Kohli et al., 1999) and tobacco (Shimizu et al., 2001) provide strong evidence that in plants genetically engineered by direct DNA delivery, transgene integration and rearrangement result from illegitimate recombination or involve DNA synthesis dependent mechanisms (i.e DNA-break-repair mechanisms) (Svistashev et al., 2002).

4.2 Transgene inheritance and expression.

Transgene inheritance usually is scored by observing the segregation of the transgene phenotype. However, the transgene phenotype is variable because transgene expression level is unpredictable. Several factors may influence the expression of transgenes and factors modifying transgene expression will be reflected in distorted transgene segregation ratios. Transgene DNA or transgene phenotype introduced by particle bombardment, in most cases, is inherited as a single Mendelian factor (Armstrong et al., 1995a; Bregitzer and Tonks, 2003; Fromm et al., 1990b; Gordon-Kamm et al., 1990; Pawlowski and Somers, 1996). This implies the presence of only one integration locus or the presence of closely linked loci
segregating together (Pawlowski and Somers, 1996). Because the transgenic locus is, in most cases, hemizygous and most transgenes provide plants with a gain of function, they behave as dominant genes and a 3:1 segregation ratio for transgenic to wild-type progeny is normally expected in diploids or diploidized polyploids when a hemizygous plant is selfed. Also, the inheritance of transgene integration has been investigated (Cooley et al., 1995; Register et al., 1994; Spencer et al., 1992). Consistent with the transgenic phenotype segregation pattern, the multiple hybridizing transgene restriction fragments observed in transgenic plants cosegregate in most cases. These observations indicate that even complex integration patterns involving full-length transgene sequences, rearrangements, concatemers, and transgene sequences interspersed with plant genomic DNA are tightly linked loci or are integrated into a single site.

Levels of transgene expression in plants are generally unpredictable and vary between independent transformants. Several factors may influence the expression of transgenes. Differences in the number of transgene copies integrated in the genome are known to contribute to transgene expression levels. Although it has been reported that higher number of transgene copies were associated with low levels of expression (Flavell, 1994), the literature is conflicting on the relationship between transgene copy number and expression in plants. A positive correlation between high transgene copy number and expression has been reported (Gendloff et al., 1990; Leisy et al., 1989; Stockhaus et al., 1987; Stoger et al., 1999; van der Hoeven et al., 1994). In contrast, no strict correlation was found (Hobbs et al., 1993; McCabe et al., 1999; Shirstat et al., 1989; van der Krol et al., 1990), whereas a negative correlation was reported in several transgenic crops including maize (Allen et al., 1993; Cooley et al., 1995; Linn et al., 1990). Such discrepancies have been attributed, at least in
part to construct-specific effects (Cooley et al., 1995). The transgene integration sites and patterns (position effects) have been found also to have impact on expression of a transgene. When tranGenes are integrated into euchromatin regions, their expression may be influenced by regulatory sequences of nearby host genes. If they integrate into or near heterochromatin regions, they can be methylated or silenced. This position effect of expression of transgenes is often complicated by the integration pattern of a transgene at a given transgene locus. It has been found that multiple copies of the transgenes especially when they are inverted, may lead to significant reduction of expression, or gene inactivation (Finnegan and McElroy, 1994; Flavell, 1994; Matzke and Matzke, 1995).

4.3 Transgene inactivation

Several mechanisms mediate transgene inactivation (gene silencing). It will not be the purpose of this literature review to give a comprehensive review of the mechanisms that contribute to transgene inactivation as more detailed reviews on the subject have been reported (for a more current and detailed review refer to Fagard and Vaucheret, 2000; Chandler and Vaucheret, 2001). However, relevant information about homologous sequences will be briefly mentioned. Presence of sequences homologous to either transgene sequences or endogenous gene sequences have been reported to be associated with transgene inactivation. Homologous sequences can lead to methylation of the promoter region blocking transcription (Davies et al., 1997; Kumpatla et al., 1997b; Mittelsen et al., 1998). Alternatively, one model postulates that such methylated homologous sequences may form or recruit repressive chromatin complexes. The repressive chromatin then spreads to adjacent
homologous sequences or interact in trans with other homologous sequences altering the chromatin structure (Chandler and Vaucheret, 2001; Fagard and Vaucheret, 2000).

4.4 Transgene transmission

In addition to factors that affect transgene expression, segregation distortion of transgene phenotypes may also arise due to reduced transgene transmission or loss (Register et al., 1994). Distorted transgene segregation at the early generation could be due to effects of chimera within transgenic callus and subsequent plants, tissue-culture induced genetic variations, weak plant vigor and biased selection on the basis of selectable markers (Zhong, 1998). Persistent distortion of transgene segregation may occur when the transgene is linked to a deleterious locus or when the transgene disrupts an endogenous locus which is critical to plant development or gamete transmission (Pawlowski and Somers, 1996). Also, segregation distortion may occur if the transgene produces a product that interferes with the endogenous biochemical pathways for gamete development and transmission (Hood et al., 1997). In general, the male gamete is more prone to deleterious effects than the female gamete (Zhong, 1998). The reduced gamete transmission may result in the absence of, or fewer homozygous transgenic progeny than expected (Sangtong et al., 2002).

5. α-lactalbumin (α-La)

α-Lactalbumin (α-La) is a small (Mr 14, 200), acidic (pI 4-5), Ca$^{2+}$ binding milk protein (Permyakov and Berliner, 2000). It is synthesized only in the mammary gland (McKenzie et al., 1991). Most α-lactalbumins including human and porcine consist of 123 amino acid residues and are homologous in sequence to the lysozyme family such as hen egg-white lysozyme (Permyakov and Berliner, 2000; McKenzie et al., 1991). It performs
several important functions: First, in mammary secretory cells, it is one of the two components of lactose synthase, which catalyzes the final step in lactose biosynthesis in the lactating mammary gland (Permyakov and Berliner, 2000). The other component of this system is galactosyl transferase (GT), which is involved in the processing of proteins in various secretory cells by transferring galactosyl groups from UGP-galactose to glycoproteins containing N-acetylglucosamine. In the lactating mammary gland, the specificity of GT is modulated by interaction with α-La, which increases its affinity and specificity for glucose: $\text{UDP-Gal + glucose} \xrightarrow{\text{GT/α-La}} \text{lactose + UDP}$ (Permyakov and Berliner, 2000). Second, α-La strongly binds calcium. It is frequently used as a simple model of calcium binding effects on protein interactions with protein, peptides, membranes and low molecular weight organic compounds (Permyakov et al., 2000). Third, it has been found recently that some forms of α-La have bactericidal activity (Hakansson et al., 2000) and can induce apoptosis in tumor cells (Hakansson et al., 1995; Svensson et al., 1999). A cDNA clone encoding porcine α-La has been isolated and sequenced (Gupta et al., 1992). The cDNA encoded a preprotein of 141 amino acids, including a leader peptide of 19 amino acids. The codon usage of the porcine α-La was markedly different from the codon usage typically found in maize genes. It was thought that this difference could limit the accumulation of porcine α-La in plants. So the amino acid sequence derived from this clone was redesigned to create a porcine α-La gene that could be more efficiently translated and thus lead to high levels of expression in maize cells (Yang et al., 2002). The amino acid sequence derived from this cDNA clone was reverse-translated into a nucleotide sequence using maize preferred codons—those codons that appeared most frequently in GenBank maize
accessions for a given amino acid (GCG, assembled by Mike Cherry, Stanford University; Yang et al., 2002). A comparison of the codon usage of the native porcine α-La coding sequence and codon adjusted version was reported and different constructs that contain this modified porcine α-La coding sequence were made and transformed into maize (Yang et al., 2002).
References


Osborne, T.B. 1924. The Vegetable Proteins Longmans, Green, London.


Chapter 2. Inheritance and Expression Patterns of a Synthetic Porcine 
\(\alpha\)-Lactalbumin Transgene in Maize

Earl H. Bicar\(^1\), Wendy L. Woodman-Clikeman\(^1\), Paul M. Scott\(^2\) and Michael Lee\(^1\)*

Key words: transgene, porcine \(\alpha\)-lactalbumin, gene expression, \textit{Zea mays}

Abstract

A codon-adjusted version of a gene encoding the porcine milk protein \(\alpha\)-lactalbumin was synthesized to create maize with improved nutritional quality. Three expression vectors containing the synthetic gene were constructed and transformed into maize callus by particle bombardment. The constructs differ in their promoters and subcellular targeting domains. The objective of this study was to characterize the expression and inheritance patterns of these transgenes. \(\alpha\)-La DNA sequences were detected by PCR analysis of F1 plants in 6 of 8 events of P64, 6 of 6 events of P45 and 4 of 5 events of P57, indicating that the \(\alpha\)-La gene was integrated into the maize genome and transmitted to the F1 generation. Western blot analysis indicated that the \(\alpha\)-La protein accumulated in the endosperm, showing that the

\(^1\)Department of Agronomy, Iowa State University, Ames, IA 50011

\(^2\)USDA-ARS Corn Insects and Crop Genetics Research Unit

*Author for correspondence
transgenes were functional. Genomic Southern blot analysis showed that each progenitor F1 plant had a unique pattern of bands, indicating that each event had a unique integration pattern. The integration patterns in the F2 and F3 progenies were identical to their corresponding F1s, indicating that new rearrangements were not detected and demonstrating stable integration.

The expression and inheritance patterns of the α-La transgenes were assessed. In 5 of 6 events of P64, the α-La transgene was inherited as a single locus through the F3 generation. In P45, irregular expression and inheritance patterns were observed. In 3 events, the α-La transgene was inherited as a single locus in the BC1F1 and F2 generations. In the F3 generation, segregation distortion manifested by an increase in frequency of kernels with no detectable α-La protein was observed. PCR analysis of these negative kernels confirmed presence of the α-La transgene, indicating that the transgene was not functional. In 2 events that contained α-La in the F1, the α-La protein was not detected in the BC1F1 and F2 generations. However, in F3 kernels derived from self-pollination of these F2 null plants, α-La was detected, suggesting that the transgene was reactivated. This putative reactivation was also observed in the F3 generation in 2 events of P57 that manifested transgene inactivation in the BC1F1 and F2 generations. Possible causes of these irregular inheritance patterns are discussed.

Introduction

Cereal grains are major sources of nutrition for humans and other animals. The major cereal crops however are deficient in one or more of the essential amino acids and therefore
do not provide a well-balanced diet. Maize, for example, accumulates storage proteins that are low in Lysine, Methionine and Tryptophan (Osborne and Clapp, 1908). These deficiencies limit the nutritional value of the grain. Recent advances in genetic engineering and improvements in transformation technology have provided new opportunities to modify the nutritional composition of the maize grain through the use of transgenes.

One strategy has been to modify an amino acid biosynthetic pathway by introducing feedback insensitive versions of rate limiting enzymes (Lundquist et al., 2001). This approach has been used to increase the free Lysine content of rice (Lee et al., 2001a) and maize (Jones et al., 1998). Modifying the biosynthetic pathways however can lead to accumulation of intermediates that might be toxic to the plant (Falco et al., 1995). A second approach has been to overproduce a protein rich in the amino acids that are naturally deficient in seeds (Lai and Messing, 2002). In some cases, the use of plant proteins to achieve this goal, such as the high-methionine 2S albumin Brazil nut protein, can be problematic in animal diets because they can be poorly digested or cause allergenic reactions (Nordlee et al., 1996). Milk proteins are ideal for production in plants because they have excellent nutritional parameters (Chong et al., 1997a; Mitra and Zhang, 1994; Takase and Hagiwara, 1998; Yang et al., 2002). Among milk proteins, porcine $\alpha$-lactalbumin is an attractive target for expression to improve the nutritional value of maize because it is highly digestible, non-antigenic and has a high content of Lysine, thereby complementing the amino acid balance of maize. Also, several potentially beneficial biological properties of $\alpha$-La have been identified recently. A folding variant of $\alpha$-La has been shown to have bactericidal activity against streptococci (Hakansson et al., 2000), and to induce apoptosis in tumor cells (Hakansson et al., 1999; Svensson et al., 2000).
In an attempt to improve the nutritional value of maize, a codon-adjusted porcine \( \alpha \)-lactalbumin coding sequence was synthesized (Yang et al., 2002). Constructs that differ in subcellular targeting domains were made using this coding sequence, and were introduced into maize callus by particle bombardment. Maize plants were obtained from regeneration. Yang et al. (2002) tested one construct that has no targeting information and one with a maize signal sequence. Accumulation of \( \alpha \)-La was not detected in the kernels of plants transformed with a construct with no targeting information while \( \alpha \)-La accumulated in the grain when a maize signal sequence was added. However, the \( \alpha \)-La expression was frequently unstable and not seen in subsequent generations.

In order to obtain more information on the expression of the synthetic porcine \( \alpha \)-La transgene in maize, three other constructs were made. These constructs differ in their promoters and in their subcellular targeting domains (Figure 2.1). These were introduced into maize cells by particle bombardment and plants were obtained from regeneration. However, it was not established whether the transgenes were integrated into the maize genome and functional. Additionally, the expression and inheritance patterns of the porcine \( \alpha \)-lactalbumin transgenes in succeeding generations from these constructs are unknown.

The progeny of transgenic plants produced by particle bombardment often exhibit segregation distortion resulting from the abnormal expression of the transgene (Pawlowski and Somers, 1996). Abnormal expression of transgenes can in many instances be detected in the first generation of transgenic plants. However, in some experiments abnormal transgene expression is not observed until large-scale field trials (Brandle et al., 1995). Therefore, long term studies are needed that characterize the expression and inheritance of transgenes.
through several generations. Transgenes can only be of practical use when they are inherited and expressed in a predictable manner. The objectives of this study were to 1) confirm the integration of three synthetic porcine \(\alpha\)-lactalbumin transgenes into the maize genome; 2) determine whether the transgenes are functional; 3) assess the stability of the transgene in the genome and; 4) evaluate the inheritance and expression patterns of the transgenes for three generations.

**Materials and Methods**

**Gene constructs**

A modified porcine alpha-lactalbumin (\(\alpha\)-La) coding sequence (Yang et al., 2002) was used to produce 3 different constructs (Figure 2.1). In construct pUbi/z-sig/lactal/KDEL, the synthetic \(\alpha\)-La coding sequence is transcriptionally regulated by the maize Ubi-1 promoter and the nos 3' untranslated region. In this construct, the coding region encodes mature porcine \(\alpha\)-lactalbumin and is translationally fused to the maize 27-kDa gamma-zein signal sequence at the N-terminus and an endoplasmic reticulum (ER) retention sequence (KDEL) at the C-terminus of the protein (Figure 2.1A). Transgenic maize calli and plants transformed by this construct were designated as P45. Construct pZein/z-sig/lactal/KDEL is identical to the first construct except that the maize Ubi-1 promoter was replaced by the maize 27-kDa gamma-zein promoter (Figure 2.1B). Transgenic maize calli and plants transformed by this construct were designated as P64. Construct pZein/z-sig/lactal is identical to the second construct except that the coding sequence does not encode the ER retention sequence (KDEL) at the C-terminus of the protein (Figure 2.1C). Transgenic maize
calli and plants transformed by this construct were designated as P57. The selectable marker gene BAR, driven by the maize Ubi-1 promoter, is contained on the same plasmid as the α-La gene in pUbi/z-sig/lactal/KDEL. In pZein/z-sig/lactal/KDEL and pZein/z-sig/lactal, the BAR gene was on a separate plasmid (pBAR184) and the two plasmids were co-bombarded.

Plant transformation and tissue culture

Plant transformation and tissue culture was done at the Plant Transformation Facility at Iowa State University as described (Frame et al., 2000). Highly embryogenic, Type II callus was initiated from embryos of the Hi II genotype (Armstrong et al., 1991). Established Type II embryogenic callus lines were bombarded with pUbi/zsig/lactal/KDEL, pZein/zsig/Lactal/KDEL or pZein/zsig/Lactal using particle bombardment–mediated transformation with a PDS-1000/HE Biolistic Particle Delivery System (Bio-Rad) following the manufacturer’s recommendations. One week after bombardment, calli were moved to selective media containing Bialaphos. After 8 weeks, bialaphos resistant calli were moved to regeneration media supplemented with bialaphos. The plants obtained from regeneration were designated as T0 plants and were then transplanted into pots.

Development of segregating populations

The F1 generation was produced by matings between a T0 plant and inbred line B73 using the T0 plants as females. Five F1 kernels from each event were planted and the F1 plants obtained were crossed as males to B73 to obtain BC1F1 kernels. In addition, the F1 plants were self-pollinated to obtain the F2 kernels. From each F2 ear, 40 to 65 F2 kernels were planted and the resulting plants were self-pollinated to obtain F3 kernels. From each event, 2 to 20 F3 ears were obtained and from these, 2 to 5 F3 ears were selected. Seventeen
F3 kernels from each selected F3 ear were planted and self-pollinated to obtain the F4 generation.

_Evaluation of protein production by Western blot analysis of kernels_

A small portion of the endosperm (~ 10 mg) was removed from the kernels using a hand-held drill. Protein was extracted from this tissue with 100 µl of SDS–PAGE sample buffer (0.5M Tris-HCl pH 6.8; 10% SDS; 10% glycerol; 5% BME) per 10 mg of endosperm powder. The samples were then placed in a vortex shaker for 30 minutes. The insoluble material was removed by centrifugation at 13,000 rpm for 5 minutes in a micro-centrifuge. The supernatant was boiled for 5 minutes before loading on a 15 % SDS-PAGE gel (Laemmli, 1970). Gels were blotted onto a nylon–backed nitrocellulose membrane (0.45 µm) using a mini-transblot apparatus (Bio-Rad). The membrane was treated with 1 % ovalbumin in PBS-Tween buffer (8.0 g NaCl; 0.2 g KCl; 1.15 g Na₂HPO₄; 0.24 g KH₂PO₄; 1 ml Tween 20) for 1 hour, allowed to react for 8-10 hours with a polyclonal antibody against human α-La raised in rabbits and visualized according to the manufacturer’s protocol for colorimetric visualization of an alkaline phosphatase-conjugated anti-rabbit IgG (Bio-Rad).

_PCR analysis_

F1 kernels that were sampled for Western blot analysis were planted in the field. Three weeks after germination, leaf tissue was collected from the plants. DNA was extracted from the fresh leaf tissues using the Puregene DNA Isolation Kit (Gentra Systems) following the manufacturer’s recommendations. The DNA pellet was resuspended in 50 µl of 50 mM Tris, 10 mM EDTA, pH 8.0. PCR was carried out using 1 µl of isolated genomic DNA (~ 50 ng) in a 20 µl total reaction volume containing 100 µM of each dNTP, 2 µl of 10X PCR
buffer (Gibco BRL), 1 unit Platinum Taq DNA Polymerase (Gibco BRL), 2.85 mM of MgCl₂ and 0.2 μM of each primer. The forward primer (LA-F = AAGCAGTTCAACCAAGTGCGAGC) and the reverse primer (LA-R = TCTTCTTGCGCACATCATGTC) corresponded to the α-La coding sequence and amplified a fragment of about 300 bp. PCR conditions were 35 cycles of 30 seconds at 94 °C, 30 seconds at 58 °C and 2 minutes at 72 °C, with 5 minutes at 94 °C prior to the reaction and 5 minutes of final extension at 72 °C in a Rapidcycler (Idaho Technologies, Inc). PCR products were analyzed on 1.5 % (w/v) agarose gels stained with ethidium bromide and photographed.

Southern blot analysis

DNA was prepared from leaf tissue by the CTAB method (Saghai-Maroof et al., 1984). Genomic DNA was digested with enzymes that cut the plasmid once outside the α-La transgene. P64 and P57 genomic DNA was digested with Sea I and P45 DNA was digested with Ssp I, and then fractionated by electrophoresis on an 0.8 % agarose gel. The α-La-containing plasmids used to produce P64 and P57 are about 4 kb while the plasmid used in P45 is about 8 kb in size. Transfer to nylon membranes (Hybond, Amersham) was performed as described (Southern, 1975). Approximately 5 μg of genomic DNA was used per digest. Hybridization was carried out using a ³²P-labelled probe at 65 °C for 12-16 hrs. The probe was obtained by amplifying the α-La coding region from the transforming plasmid YPS 112 in a PCR. The amplified band was resolved in a low melting point agarose gel, cut out and then labeled with (α-³²P) dCTP according to Feinberg and Vogelstein (1983).
Results

Integration of the porcine α-lactalbumin transgene into the maize genome

A codon-adjusted version of a gene encoding the porcine milk protein α-lactalbumin was synthesized to create maize with improved nutritional quality (Yang et al., 2002). Three constructs were made that differ in their promoters and subcellular targeting domains. The features of these constructs are described in Figure 2.1. These constructs were introduced into maize callus cells by the Iowa State University Plant Transformation Facility using their standard protocol for particle bombardment (Frame et al., 2000). Transformed maize calli were transferred to regeneration media to obtain T0 plants. The T0 plants were then crossed as females to non-transformed B73 plants to obtain the F1 kernels. F1 kernels were obtained from 8 different P64 T0 plants, 7 different P45 T0 plants and 5 different P57 T0 plants. PCR analysis was conducted to confirm the integration of the transgene into the genome of the F1 plants. The F1 kernels were planted and DNA was obtained from leaf tissue of F1 plants for PCR analysis. Using primers corresponding to the synthetic porcine α-lactalbumin coding sequence, a PCR product was detected in 6 of 8 different P64 F1 plants, in 6 of 6 different P45 F1 plants and in 4 of 5 different P57 F1 plants (Figure 2.2 and data not shown for P57). Since the F1 plants were obtained from a cross between T0 and non-transformed B73 plants, the presence of α-lactalbumin DNA sequences in these F1 plants as determined by PCR analysis indicates that in each event, the porcine α-La transgene was integrated into the maize genome and transmitted to the F1 generation.
Accumulation of porcine α-lactalbumin in maize kernels.

Having established that the transgene was successfully integrated in the maize genome, accumulation of the porcine α-La protein was assessed to determine whether the synthetic porcine α-La transgenes function in the plant system. Western blot analysis was performed to detect porcine α-lactalbumin protein in the endosperm of transgenic maize kernels. Five F1 kernels from each event were analyzed using protein extracts from endosperm tissues. Accumulation of porcine α-lactalbumin was detected by Western blot in 6 events of P64, 6 events of P45 and in 4 events of P57 (Table 2.1). As shown in Figure 2.3, novel protein bands that reacted with the polyclonal antibody raised against human α-lactalbumin in rabbits were detected in some transgenic kernels. A prominent band co-migrated with the human α-lactalbumin positive control, suggesting that porcine α-lactalbumin was synthesized and accumulated in the endosperm tissue. When protein extracts from the nontransgenic (B73) endosperm were analyzed, there was no detectable reaction with the polyclonal antibody. Interestingly, immunoreactive bands of larger molecular masses (~ 31 kDa and 45 kDa) were also detected in positive kernels and these bands were not observed in the human α-lactalbumin or in the negative control. These bands may be aggregated porcine α-lactalbumin. These results demonstrated that not only was the porcine α-lactalbumin transgene transmitted to the F1 generation but that the transgene was also functional in P64, P45 and P57 F1 kernels.
Stable integration and transmission of porcine α-La transgenes

Southern blot analyses were performed on several F1 plants and their F2 and F3 progenies from each event in order to characterize the transgene integrations and determine transgene stability through meiosis. Blots were hybridized to a labeled fragment of the porcine α-La coding sequence. Each progenitor F1 plant had a unique pattern of bands, indicating that each event had a unique integration pattern (Figure 2.4 A and B). The integration patterns in the F2 and F3 progenies were identical to their corresponding progenitor F1s (Figure 2.4 B), indicating that rearrangements of the transgene insertions were not detected. Additionally, since the restriction enzymes used to digest the genomic DNA cut only once in the plasmid, the number of transgene copies can be estimated based on the number of the bands. Between 1 and 12 copies of the transgene were detected in the events examined.

Stable inheritance and consistent expression patterns in the BC1F1, F2 and F3 generations among events in P64

The expression and inheritance of transgenes in the first generation of transgenic plants is not always a good predictor of transgene stability in subsequent generations. Studies have reported that problems with transmission and expression of a transgene may arise in advanced generations (Register et al., 1994). To characterize the expression and inheritance of the synthetic porcine α-La transgenes, transgene expression was monitored by western blot analysis for up to three generations. Among the P64 BC1F1 kernels analyzed (10 kernels from each of 8 events), porcine α-La protein was detected by Western blot in at least one kernel in 6 events. Of the 6 events, 4 events showed a phenotypic segregation ratio
for the presence and absence of the $\alpha$-lactalbumin that was not significantly different from 1:1 ratio ($P = 0.01$) as would be expected with a single dominant locus in independent assortment in the BC1F1 generation (Table 2.2).

Western blot analysis was performed on F2 kernels representing 8 P64 events. Sixty five kernels from one plant of each of 8 events were analyzed. Porcine $\alpha$-La was detected in some kernels of 6 events (Table 2.2). Four of the 6 events (P64-18, P64-14, P64-15, P64-1) showed a phenotypic segregation ratio that is not significantly different from the 3:1 phenotypic ratio expected for a single dominant locus segregating in the F2. These events are the same events that showed a single locus inheritance pattern in the BC1F1 generation. In the analysis of between 2 and 15 F3 lines obtained by self-pollination of positive F2 plants in an event, the phenotypic ratios obtained were also not significantly different from a single locus inheritance pattern in most cases, indicating a consistent expression of the synthetic porcine $\alpha$-La transgene for three successive generations in these events. One event (P64-27) displayed a phenotypic segregation ratio that deviated from the single locus inheritance ratio both in the F2 and F3 generations.

*Inheritance and expression in the BC1F1, F2 and F3 generations in P45 and P57*

Among 6 events of P45, porcine $\alpha$-lactalbumin was detected by Western blot in 4 events in the BC1F1 and F2 generations (Table 2.3). Three of the 4 events (P45-3, P45-22, P45-16) showed a segregation ratio that is not significantly different from the expected 1:1 phenotypic ratio for a single dominant locus segregating in BC1F1 generation or the 3:1 phenotypic ratio in the F2 generation (Table 2.3). In the F3 generation however, Western blot analysis of kernels from many F3 lines obtained by self-pollination of positive F2 plants
from events P45-3 and P45-22 displayed segregation ratios that deviated from the expected 3:1 phenotypic ratio. This segregation distortion was due to an abundance of kernels with no detectable α-lactalbumin protein (null phenotype). PCR analysis of 10 of these negative F3 kernels showed that α-lactalbumin DNA sequences were present in these kernels (data not shown).

Three other events of P45 displayed irregular inheritance and expression patterns (P45-11, P45-4, P45-19). In P45-11, a high frequency of negative kernels in the BC1F1 and F2 generations was observed. In the F3 generation, several F3 lines showed further increase in the frequency of these negative kernels when compared to the F2 generation. In P45-4 and P45-19, the α-lactalbumin protein was not detected in 10 BC1F1 and 44 F2 kernels analyzed even though the α-lactalbumin protein was detected in the progenitor F1s. PCR analysis of these negative kernels in the BC1F1 and F2 generations revealed the presence of α-La DNA sequences in 4 of 10 BC1F1 and 9 of 16 F2 kernels, indicating that the transgene was transmitted but not active. The absence of detectable α-lactalbumin protein in many BC1F1 and F2 kernels indicates that transgene inactivation occurred as early as the BC1F1 generation.

To investigate whether transgenes that were inactive in the F2 generation could be re-activated in the F3 generation, P45-4 F3 lines derived from the self-pollination of PCR positive but α-La negative F2 plants were analyzed. In 2 of the 5 lines, the α-lactalbumin protein was not detectable. In these lines, some of the negative F3 kernels contain α-La DNA sequences as detected in a PCR analysis, suggesting that the transgene inactivation persisted and was maintained in the F3 generation. Intriguingly, 3 of the 5 F3 lines of P45-4
had a low frequency of kernels containing the α-lactalbumin protein as detected in a Western blot, suggesting that the transgene was reactivated. A similar reactivation of the transgene was observed in 2 F3 lines of P45-19. These results demonstrated that although transgene inactivation may persist for generations, transgenes can be reactivated in the progeny of the plants that exhibited transgene inactivation.

In the 4 events of P57 that expressed the α-La protein in the F1, the α-lactalbumin protein was not detected by Western blot in either the BC1F1 or F2 generations (Table 2.3). PCR analysis of BC1F1 and F2 null kernels revealed presence of α-La DNA sequences in some of these negative kernels, indicating that the transgene was transmitted but not functional. In the F3 generation however, 1 out of 5 F3 lines from event P57-35 showed 1 out of 17 kernels analyzed that is positive for α-La protein, while in the other 4 F3 lines from the same event, kernels with α-La protein were not observed. Similarly in P57-24, two of the 5 F3 lines showed one or two kernels with detectable α-La while in the remaining 3 F3 lines, α-La was not detected. These observations are somewhat similar to the transgene inactivation observed in BC1F1 and F2 generations and the reactivation of the transgene in the F3 generation in some events of P45 construct. Southern blot analysis of one event (P57-24, data not shown) that showed transgene inactivation in the F2 and putative reactivation in the F3 showed that the hybridization profiles of the progenitor F1, F2 plant and the “reactivated” F3 plant were identical indicating that the “reactivated” plant is of the same genetic makeup as its F1 and F2 parents and that this plant was not a result of contamination that may have occurred during self-pollination.
Of the events analyzed, 5 in 6 events of P64 (or 83 %) demonstrated a consistent expression and stable inheritance of the synthetic porcine α-lactalbumin for three successive generations. Conversely, 5 in 6 events of P45 (or 83 %) and all 4 events of P57 displayed irregular patterns of expression and inheritance of the synthetic porcine α-lactalbumin transgene from the BC1F1 to the F3 generations.

Discussion

The inheritance of several events from each of three different constructs encoding the modified porcine alpha lactalbumin was examined. In stably inherited events, segregation analysis of the presence of α-La showed single active transgene locus segregation ratios. This result agrees with several studies that found that transgenes (or transgene phenotypes) introduced by microprojectile bombardment are inherited as a single Mendelian factor (Armstrong et al., 1995b; Cooley et al., 1995; Register et al., 1994). This would imply the presence of only one active locus in the genome or the presence of several closely linked loci segregating together as observed in oat transgenes (Pawlowski and Somers, 1996).

In cases where distortion in the inheritance and expression patterns were detected, transgene inactivation as manifested by kernels that carry the transgene but have no detectable α-La protein was also observed. This transgene inactivation can be explained by gene silencing. Transgene silencing, herein defined as the inactivation of the transgene despite the presence of an unchanged, but possibly methylated transgene sequence in the plant genome (Pawlowski et al., 1998) has been extensively documented in transgenic plants produced by Agrobacterium-mediated transformation and also those transformed by microprojectile bombardment (Chandler and Vaucheret, 2001; Fagard and Vaucheret, 2000;
Finnegan and McElroy, 1994; Matzke and Matzke, 1995). In this study, transgene silencing was observed in 1 of 6 events of P64, in 5 of 6 events of P45 and in all 4 events of P57.

It is interesting to compare the inheritance of the P45 and P64 plants because both contain the zein signal sequence and the ER retention signal KDEL but they differ in their promoters (Figure 2.1). P64 is driven by the seed-specific zein promoter while P45 is driven by the constitutive ubiquitin promoter. In the P45 plants, gene silencing was much more prominent when compared to P64 (86% vs 16% of the events studied). Also, the selectable marker gene BAR used a different promoter than the α-La transgene and was on a separate plasmid that was co-bombarded with the α-La transgene in P64. The selectable marker gene was on the same plasmid as the α-La transgene in P45, and both genes used the \textit{ubi-1} promoter. It is conceivable that the higher frequency of gene silencing in P45 is related to the presence of homologous promoter sequences in the same vector in P45. This homology in the promoter sequences might have provided a target for the inactivation of the α-La transgenes. It has been shown in other studies that the presence of homologous or identical promoter sequences on the same vector can lead to transgene inactivation (Finnegan and McElroy, 1994; Matzke and Matzke, 1995; McCabe et al., 1999). As little as 90 base pairs of promoter homology have been reported to be sufficient for transgene silencing to occur (Fagard and Vaucheret, 2000; Matzke et al., 1994; Vaucheret, 1993). This would have important implications in the future design of α-La transgene constructs. For an efficient and stable expression of the transgene, it may be preferable to avoid linked transgenes with identical promoters or repeated elements. It is possible that homology with endogenous promoter sequences from the zein and ubiquitin genes contributed to the α-La transgene
inactivation. If so, this would indicate that the ubiquitin promoter used in P45 is more prone to silencing than the zein promoter from P64.

Three events of P45 (P45-3, P45-22, P45-11) showed increased frequency of negative phenotypes in the F3 compared to the F2 generation. These events have fewer transgene copies (~1-3 copies) based on the number of hybridizing fragments in a Southern blot (Figure 2.4 A and B). In these events, the frequency of silenced phenotypes increased after selfing. It has been shown that the increase in homozygosity resulting from selfing contributes to gene silencing (Flavell, 1994; Hobbs et al., 1993) and so it is plausible that a gene dosage dependent mechanism of gene silencing may also have contributed to the silencing of α-La in these events. This will have important practical considerations in the handling of these events as continued selfing that results in homozygosity might be avoided to maximize transgene expression.

It is possible that differences in transgene copy number also contributed to gene silencing observed among events. For example, P45-4 contains multiple fragments in a Southern blot (~ 14 bands, Figure 2.4A) and this event displayed gene inactivation in the BC1F1 and F3 generations. It is difficult however to attribute the differences in transgene expression patterns among the different events solely on copy number as some events, such as P64-18, showed consistent expression despite presence of multiple transgene fragments in a Southern blot (~ 12 bands, Figure 2.4A). Although it has been reported that higher number of transgene copies were associated with low levels of expression (Flavell, 1994), the literature is conflicting on the relationship between transgene copy number and expression in plants. A positive correlation between high transgene copy number and expression has been reported (Gendlof et al., 1990; Leisy et al., 1989; Stockhaus et al., 1987; Stoger et al., 1999;
van der Hoeven et al., 1994). In contrast, no strict correlation was found (Hobbs et al., 1993; McCabe et al., 1999; Shirstat et al., 1989; van der Krol et al., 1990), whereas a negative correlation was reported in several transgenic crops including maize (Allen et al., 1993; Cooley et al., 1995; Linn et al., 1990). Such discrepancies have been attributed, at least in part to construct-specific effects (Cooley et al., 1995). Position effects (i.e the effect of the chromosomal locus on transgene expression) have also been reported to explain the differences in transgene expression among different events. Position effects are, in turn, partly explained by the methylation status of the integration site (Prols and Meyer, 1992).

P57 and P64 are both driven by zein promoters and contain zein signal sequences. Unlike P64 however, P57 does not contain the ER retention signal KDEL. The porcine α-La is a secreted protein and the KDEL sequence may cause the protein to be retained in the ER. Without the KDEL, the α-La may be secreted into the intercellular space. Comparison of α-La expression between P64 and P57 shows that α-La was consistently detected for 3 generations in P64 indicating a more stable expression. In contrast, α-La was not detected among events in P57 in the BC1F1, F2 and in the F3 generations. Since the two constructs differ only in the presence or absence of the KDEL sequence, it is reasonable to conclude that the non-detection of α-La in P57 might be due to the absence of KDEL in this construct. It is likely that the α-La is expressed at levels below the detection limit in P57 due to the lack of KDEL sequences. Chikwamba et al., (2003) in a study using constructs driven by zein promoters, detected higher levels of LT-B protein using a construct that contained KDEL than one which did not have KDEL. The non-detection of α-La among events in P57 seemed related also to transgene silencing. In P57, the α-La was detected in the F1 but not in
the F2 generation. PCR analysis of some of these F2 null kernels revealed the presence of α-La DNA sequences indicating that the transgene was transmitted but inactivated. It is not clear whether the KDEL might have contributed to gene silencing in this construct. This only illustrates the complexity of characterizing transgene expression and inheritance patterns based on phenotypes. The inheritance and expression patterns of P57 are in agreement with Yang et al. (2002). Yang et al. (2002) detected the α-La protein among F1 sibling kernels in a construct driven by the ubiquitin promoter (P42) and without the KDEL sequences. However in the F2 generation, no α-La protein was detected among F2 kernels obtained by selfing positive F1 plants.

Transgene silencing and the distortions of transgene inheritance were unstable. Although in most cases the silenced phenotype in the F2 was maintained in the F3 generation, α-La expression was observed in some F3 lines derived from self-pollination of inactivated F2 plants suggesting reactivation of α-La expression. Such “reactivated” kernels were infrequent in some F3 lines and were observed in P45-4, P45-19, P57-24 and P57-35 (Table 2.3). One can argue that since the reactivated kernels occur rarely, it may be possible that these “reactivated” kernels were obtained by contamination during self-pollination from pollen that is positive for the α-La transgene. If this reactivated F3 kernel is a result of contamination, its hybridization profile in a Southern blot should be different from its progenitor F1, F2 plant, or with its uncontaminated F3 siblings. Although reactivation of silenced phenotypes has been reported for transgenes, the putative reactivation observed in this study needs to be further investigated by evaluating the phenotypes of the F4 progenies from these “reactivated” F3 kernels. It would be interesting to determine whether the
progenies of "reactivated" plants will continue to express $\alpha$-La and to establish the segregation and expression patterns of the $\alpha$-La transgene in later generations. The occurrence of putative reactivation here only indicates that the inheritance of transgene expression is highly variable and irregular.
References


Chikwamba, R. 2003. Maize as a production and delivery vehicle of edible vaccines against the enterotoxigenic E. coli and the swine transmissible gastroenteritis.


<table>
<thead>
<tr>
<th>Construct</th>
<th>α-La gene structure</th>
<th>Callus ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. pUbi/zsig/lac/kdel</td>
<td><img src="image" alt="Gene Structure A" /></td>
<td>P45</td>
</tr>
<tr>
<td>B. pZein/zsig/lac/kdel</td>
<td><img src="image" alt="Gene Structure B" /></td>
<td>P64, pBAR184</td>
</tr>
<tr>
<td>C. pZein/zsig/lac</td>
<td><img src="image" alt="Gene Structure C" /></td>
<td>P57, pBAR184</td>
</tr>
</tbody>
</table>

Figure 2.1. Structure of α-lactalbumin genes used in transformation. The boxes in bold indicate the position of the 27 kDa gamma-zein signal sequence. The striped boxes indicate the position of the ER retention signal (KDEL). Ubi and zein are the constitutive and seed specific maize promoters, respectively. α-La indicates the codon adjusted mature porcine alpha-lactalbumin sequence. Nos indicate the position of transcriptional terminator sequence from the nopaline synthase gene from Agrobacterium. BAR is the selectable marker linked on the same plasmid in construct pUbi/zsig/lac/kdel. Calli transformed by this construct were designated as P45. In constructs pZein/zsig/lac/kdel and pZein/zsig/lac, BAR was on a separate plasmid (pBAR184) and co-bombarded. Calli transformed by pZein/zsig/lac/kdel were designated as P64 and calli transformed by pZein/zsig/lac were designated as P57. The values above the boxes are the sizes of the individual cassette in kilobases while the values in bold with arrows are the sizes of the whole transgene cassette.
**Figure 2.2.** PCR analysis to determine the presence of $\alpha$-lactalbumin transgene among progenitor F1s from different events of P64 and P45 constructs. B = nontransgenic B73 DNA used as a template for the negative control, P = plasmid vectors that contain the synthetic $\alpha$-lactalbumin gene used as positive control. Lanes 1-8 contain DNA from F1 plants from different events of P64. Lanes 9-15 contain DNA from F1 plants from different events of P45. The + and - signs are the corresponding Western blot scores of F1 kernels indicating the presence (+) and absence (-) of $\alpha$-lactalbumin protein. All plants that were positive for the presence of porcine $\alpha$-lactalbumin DNA sequences by PCR analysis were also positive for the presence of the protein in the endosperm by Western blot, suggesting that the synthetic porcine $\alpha$-lactalbumin transgene was functional.
Figure 2.3. Western blot analysis to determine the presence of α-lactalbumin protein in the endosperm of transformed maize kernels in P45-16. H is the positive control lane that contains human α-lactalbumin. B is the negative control lane that contains protein extracts from the endosperm of a non-transformed inbred line B73. Lanes 1-11 contain total protein extracts from the endosperm. Kernels with detectable α-lactalbumin protein contain bands of about 14 kDa in size. Positive and negative signs correspond to the presence (+) and absence (-) of α-La DNA sequences as detected by PCR analysis. Kernels containing α-La transgene were also positive for the α-lactalbumin protein.
Figure 2.4. A. Hybridization profiles of progenitor F1s from different events of P45, P64 and P57. Numbers below the constructs are the different events. Each progenitor F1 plant had unique pattern of bands indicating unique integration pattern. B. Southern blot analysis of porcine α-La transgene in progenitor F1 and the corresponding F2 and F3 progenies from 2 events in P45 and P64 constructs. Genomic DNA was digested with *Ssp I* for P45 and *Sca I* for P64 and probed with α-La coding sequence. P1 and P2 are the plasmids used to transform P45 and P64, respectively as positive controls. It contains 5 pg of DNA representing 1 copy per diploid genome. B= untransformed DNA from B73 as negative control. Lanes 1 contain DNA from the progenitor F1 plant. Lanes 2 contain DNA from an F2 plant. Lanes 3 and 4 contain DNA from 2 F3 sibling progenies. M= molecular weight standards in kilobases (kb). P45-11, P45-16, P64-24 and P64-15 are independent events. Plus sign (+) = presence of α-La protein in the endosperm as detected in a Western blot; negative sign (-) = no α-La protein.
Table 2.1 Frequency of events in P64, P45 and P57 constructs accumulating the modified porcine α-Lactalbumin protein in maize F1 kernels.

<table>
<thead>
<tr>
<th>P64 events</th>
<th>F1 kernels (+/-) a</th>
<th>P45 events</th>
<th>F1 kernels (+/-) a</th>
<th>P57 events</th>
<th>F1 kernels (+/-) a</th>
</tr>
</thead>
<tbody>
<tr>
<td>P64-18</td>
<td>4/1</td>
<td>P45-3</td>
<td>2/3</td>
<td>P57-35</td>
<td>3/2</td>
</tr>
<tr>
<td>P64-14</td>
<td>3/2</td>
<td>P45-22</td>
<td>4/1</td>
<td>P57-24</td>
<td>3/2</td>
</tr>
<tr>
<td>P64-15</td>
<td>4/1</td>
<td>P45-16</td>
<td>4/1</td>
<td>P57-27</td>
<td>5/0</td>
</tr>
<tr>
<td>P64-1</td>
<td>5/0</td>
<td>P45-11</td>
<td>3/2</td>
<td>P57-25</td>
<td>3/2</td>
</tr>
<tr>
<td>P64-24</td>
<td>1/0</td>
<td>P45-19</td>
<td>3/2</td>
<td>P57-26</td>
<td>0/5</td>
</tr>
<tr>
<td>P64-27</td>
<td>2/3</td>
<td>P45-4</td>
<td>2/3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P64-23</td>
<td>0/5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P64-13</td>
<td>0/5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Events</td>
<td>8</td>
<td>Total Positive</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>

a At least 5 F1 kernels from each event were analyzed. b An event is considered positive if at least 1 kernel accumulated α-La in the endosperm from among those analyzed by Western blot. (+) = presence and (-) absence of α-La protein in a Western blot.
Table 2.2. Segregation of a modified porcine α-lactalbumin transgene among events in P64 construct based on the presence and absence of the α-lactalbumin protein as detected by Western blot analysis.

<table>
<thead>
<tr>
<th>Transgenic Line</th>
<th>F1 (+/-)</th>
<th>BC1F1 (+/-)</th>
<th>F2 (+/-)</th>
<th>F3 (+/-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P64-18-5-5</td>
<td>4/1</td>
<td>5/5 ns</td>
<td>46/14 ns</td>
<td>14/5 ns; 15/5 ns; 14/6 ns; 15/5 ns; 12/5 ns; 12/4 ns; 11/6 ns; 10/3 ns; 12/5 ns; 9/5 ns</td>
</tr>
<tr>
<td>P64-14-4-3</td>
<td>3/2</td>
<td>5/5 ns</td>
<td>49/18 ns</td>
<td>9/6 ns; 13/4 ns; 14/3 ns; 10/7 ns; 12/5 ns; 10/7 ns; 11/5 ns; 12/5 ns; 10/7 ns; 14/3 ns; 14/3 ns; 13/4 ns; 14/3 ns; 12/4 ns; 17/0 H; 11/7 ns; 11/6 ns;</td>
</tr>
<tr>
<td>P64-15-6-4</td>
<td>4/1</td>
<td>5/5 ns</td>
<td>34/12 ns</td>
<td>11/6 ns; 12/5 ns; 11/6 ns; 13/4 ns; 17/0 H; 13/4 ns; 14/2 ns</td>
</tr>
<tr>
<td>P64-1-4-1</td>
<td>5/0</td>
<td>50/16 ns</td>
<td>14/3 ns; 10/7 ns</td>
<td></td>
</tr>
<tr>
<td>P64-24-4-1</td>
<td>1/0</td>
<td>4/6 ns</td>
<td>39/27 **</td>
<td>17/0 H; 15/2 ns; 16/1 ns; 20/0 H; 17/3 ns</td>
</tr>
<tr>
<td>P64-27-5-2</td>
<td>2/3</td>
<td>23/43 **</td>
<td>0/17 **; 9/8 ns; 0/17 **; 0/17 **; 0/17 **</td>
<td></td>
</tr>
</tbody>
</table>

ns = not significantly different from 3:1 phenotypic ratio for a single dominant locus model in the F2 and F3; 1:1 ratio for BC1F1; ** significantly different; \( \chi^2 \) (0.01, 1 df) = 6.64; H = homozygous for α-lactalbumin protein expression. Numbers in each generation column are the proportion of kernels expressing (+) / not expressing (−) the porcine α-lactalbumin protein in the endosperm.
Table 2.3. Segregation of a modified porcine α-lactalbumin transgene among events in P45 and P57 constructs based on the presence and absence of the α-lactalbumin protein as detected by Western blot analysis.

<table>
<thead>
<tr>
<th>Transgenic Line</th>
<th>Generations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1 (+/−)</td>
<td>BC1F1 (+/−)</td>
</tr>
<tr>
<td>P45-3-3-4</td>
<td>2/3</td>
<td>3/7 ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P45-22-1-1</td>
<td>4/1</td>
<td>4/6 ns</td>
</tr>
<tr>
<td>P45-16-2-4</td>
<td>4/1</td>
<td>3/7 ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P45-11-3-1</td>
<td>3/2</td>
<td>1/9 **</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P45-4-5-3</td>
<td>2/3</td>
<td>0/10 **</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P45-19-7-1</td>
<td>3/2</td>
<td>0/44 **</td>
</tr>
<tr>
<td>P57-35-8-1</td>
<td>3/2</td>
<td>0/10 **</td>
</tr>
<tr>
<td>P57-24-10-1</td>
<td>3/2</td>
<td>0/10 **</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P57-27-3-1</td>
<td>5/0</td>
<td>0/10 **</td>
</tr>
<tr>
<td>P57-25-2-1</td>
<td>3/2</td>
<td>0/10 **</td>
</tr>
</tbody>
</table>

ns = not significantly different from 3:1 phenotypic ratio for a single dominant locus model in the F2 and F3; 1:1 ratio for BC1F1; ** significantly different; \( \chi^2 (0.01, 1 \text{df}) = 6.64 \); H= homozygous for α-lactalbumin protein expression. Numbers in each generation column are the proportion of kernels expressing (+) / not expressing (−) the porcine α-lactalbumin protein in the endosperm.
Chapter 3. Expression of the synthetic porcine α-Lactalbumin gene results in increased Lysine levels in transgenic maize kernels

Earl H. Bicar, Paul M. Scott*, and Michael Lee

Keywords: transgene, porcine α-lactalbumin, targeting signal, Lysine

Abstract

Cereal grains are major sources of nutrition for humans and other animals. The major cereals are deficient in one or more of the essential amino acids such as Lysine and Tryptophan and therefore do not provide a well-balanced diet. These deficiencies limit the nutritional value of the grain. Recent advances in genetic engineering and the improvements in transformation technology have provided new opportunities to modify the nutritional composition of the maize grain. In an effort to elevate the amount of Lysine in maize grain, a codon-adjusted version of a gene encoding the porcine milk protein α-lactalbumin was synthesized and used to transform maize. The objectives of this study were to 1) assess the activity and tissue specificity of the α-La expression; 2) determine the levels of the porcine

1Department of Agronomy, Iowa State University, Ames, IA 50011

2USDA-ARS Corn Insects and Crop Genetics Research Unit

*Author for correspondence
α-La protein produced in the kernel endosperm; 3) quantify the levels of Lysine in the endosperm expressing porcine α-La; 4) evaluate the changes in amino acid composition and 5) determine the changes in Nitrogen, Carbon and Total protein content in α-La producing kernels.

Analysis of the activity and tissue specificity of α-La expression under the control of the 27 kDa gamma-zein and ubiquitin promoters was performed using a combination of techniques including Western blot, dot blot and tissue print analysis. Accumulation of α-La protein was assessed in the root, leaf, pollen and endosperm tissues. As expected, the 27 kDa zein promoter directed α-La expression detected only in the endosperm. Differential expression of α-La was observed for events in a construct driven by the constitutive ubiquitin promoter. Accumulation of the α-La protein was detected in the endosperm but not in the embryo. In other events of this construct, the α-La protein accumulated in the endosperm, roots and pollen but not in the embryo and leaves, indicating that the ubiquitin promoter did not have uniform tissue specificity or worked poorly for the expression of α-La in the whole plant.

The levels of the porcine α-La protein that accumulated in the endosperm ranges from 0.0014- 0.095 g/100g endosperm tissue as measured by ELISA. A significant difference in Lysine levels was detected in kernels expressing α-La when compared to their null siblings. The percent difference in Lysine content ranges from 29 % - 49 % in both F2 and F3 generation kernels. This difference was confirmed when Lysine levels were measured by complete amino acid analysis and by a bacterial assay using E. coli auxotrophs. In addition to Lysine, there were also significant changes in the content of other amino acids such as a
24% difference in Aspartic Acid levels. Expression of the porcine α-La did not result in substantial changes in percent total Carbon, Nitrogen and total protein content. The results demonstrated that a transgenic approach using a gene that encodes a protein such as a modified porcine α-La can be an effective means of improving the nutritional quality of maize grain.

Introduction

Maize is an important component of human diets and a major feed crop in many parts of the world. Maize, however is deficient in Lysine and Tryptophan and so does not provide well balanced protein (Mertz et al., 1964). One approach to improve Lysine content has been to manipulate the regulation of amino acid biosynthetic pathways by introducing feedback insensitive versions of rate limiting enzymes (Lundquist et al., 2001). This strategy increased the free Lysine in canola, soybean and maize but did not lead to an increase in the fixed Lysine content (Falco et al., 1995; Jones et al., 1998). The free Lysine leached out from the plant tissue and was lost during boiling and other processing (Shaul and Galili, 1992; Falco et al., 1995). Also, modifying the amino acid biosynthetic pathways can lead to accumulation of intermediates that might be toxic to the plant (Falco et al., 1995).

An alternative strategy to improve amino acid balance is the insertion and the expression of gene/s encoding a heterologous protein that is rich in the essential amino acid. For example, expression of the Methionine-rich protein from Brazil nut resulted in an increase in Methionine content in canola grain (Altenbach et al., 1992b). Also, the expression of a balanced seed albumin gene from *Amaranthus hypochondriacus* not only resulted in an increase in Lysine content but also in most essential amino acids and total protein content in
potato (Chakraborty et al., 2000). This strategy of expressing heterologous proteins to manipulate amino acid composition has been used in several crops and varying degrees of success have been reported (Altenbach et al., 1992a; Chakraborty et al., 2000; Chong et al., 1997a; Drakakaki et al., 2000; Molvig et al., 1997; Zheng et al., 1995). The milk protein porcine α-lactalbumin is an attractive target for expression to improve the nutritional value of maize for the following reasons: i) it is highly digestible; ii) non-antigenic; iii) it has a high content of Lysine thereby complementing the amino acid balance of maize. In an attempt to improve maize nutritional quality, a modified version of this gene was made and used to transform maize callus. The synthetic porcine α-La transgenes were placed under the control of the maize ubiquitin promoter. Maize plants that accumulated porcine α-lactalbumin in the endosperm were obtained (Yang et al., 2002). In order obtain more information on the expression patterns of the α-La transgene, different versions of the synthetic porcine α-La gene were constructed. The α-La transgenes in these new constructs were placed under the transcriptional control of the ubiquitin and 27 kDa gamma-zein promoters. The constructs also differed in their subcellular targeting domains as described in Figure 2.1. In chapter 2, the transgene integration as well as the inheritance and expression patterns of these transgenes were assessed for three generations.

Promoters determine the tissue specificity and level of gene transcription and promoter activity is often associated with the amount of protein accumulation in the target tissue. The ubiquitin promoter confers high level of expression in all tissues including the endosperm (Christensen and Quail, 1996; Cornejo et al., 1993; Stoger et al., 1999) and has been shown to induce higher levels of foreign protein in the seed than seed-specific promoters. The zein promoter on the other hand has been reported to confer strong
endosperm-specific expression of transgenes (Torrent et al., 1997), although poor expression in tissues apart from the endosperm has also been reported (Russell and Fromm, 1997). Pertinent information on the activity and tissue specificity of the ubiquitin and 27 kDa zein promoters driving α-La transcription are unknown.

In order to appreciably alter the amino acid composition or increase the levels of Lysine in the maize grain, high level expression of porcine α-La in the endosperm is required. In addition, the α-La protein must accumulate to high levels in the endosperm. In chapter 2, the consistent expression of the α-La protein has been demonstrated. However, the levels of the porcine α-La protein that accumulated in the kernel endosperm are not known. Additionally, the changes in amino acid composition as well as improvements in the levels of Lysine or total protein as a result of porcine α-La expression are also unknown.

The objectives of this study were therefore to: 1) assess the activity and tissue specificity of the α-lactalbumin expression; 2) determine the levels of the porcine α-La protein produced in the kernel endosperm; 3) quantify the levels of Lysine in the endosperm expressing porcine α-La; 4) evaluate the changes in amino acid composition and 5) determine the changes in Nitrogen, Carbon and Total protein content in α-La producing kernels.

Materials and Methods

Plant transformation and tissue culture

Plant transformation and tissue culture was done at the Plant Transformation Facility at Iowa State University as described (Frame et al., 2000). Highly embryogenic, Type II
callus was initiated from embryos of the Hi II genotype (Armstrong et al., 1991). Established Type II embryogenic callus lines were bombarded with the constructs pUbi/zsig/lactal/KDEL, pZein/zsig/Lactal/KDEL or pZein/zsig/Lactal (Figure 2.1) using particle bombardment-mediated transformation with a PDS-1000/HE Biolistic Particle Delivery System (Bio-Rad) following the manufacturer’s recommendations. Calli bombarded with these constructs were designated as P45, P64 and P57, respectively. One week after bombardment, calli were moved to selective media containing Bialaphos. After 8 weeks, bialaphos resistant calli were moved to regeneration media supplemented with bialaphos. The plants obtained from regeneration were designated as T0 plants.

*Plant growth and propagation and development of segregating populations*

T0 plants derived from P45 and P64 were crossed as females with non-transformed B73 plants to obtain F1 kernels. The F1 kernels were then sown in jiffy pots filled with soil mix (Miracle Grow). Two weeks after germination, the F1 plants were transplanted in the field at the ISU Agronomy Farm. The F1 plants were then self-pollinated to produce the F2 kernels and the F2 kernels were analyzed for the presence and absence of the α-La protein in the endosperm by Western blot. In the following season, α-La positive and negative F2 kernels were sown in the field at the ISU Agronomy Farm. Standard management practices for maize production at ISU Agronomy Farm were used. The F2 plants were then self-pollinated to obtain the F3 kernels and after self-pollination, the transgenic plants were detasseled.
Protein extraction, Western blot, dot blot and tissue print analyses

Leaf, root and pollen tissues were collected and ground in liquid nitrogen into a fine powder. Endosperm tissue was removed from the kernels using a miniature hand drill. About 10 mg of leaf, root, pollen, embryo and endosperm tissues were obtained and protein was extracted from these tissues with 100 μl of SDS-PAGE sample buffer (0.5M Tris-HCl pH 6.8; 10% SDS; 10% glycerol; 5% BME) per 10 mg of tissue. The samples were then placed in a vortex shaker for 30 minutes. The insoluble material was removed by centrifugation at 13,000 rpm for 5 minutes in a micro-centrifuge and Western blot analysis was performed. The supernatant was boiled for 5 minutes before loading on a 15 % SDS-PAGE gel (Laemmli, 1970). Gels were blotted onto a nylon–backed nitrocellulose membrane (0.45 μm) using a mini-transblot apparatus (Bio-Rad). The membrane was treated with 1 % ovalbumin in PBS-Tween buffer (8.0 g NaCl; 0.2 g KCL; 1.15 g Na₂HPO₄, 0.24 g KH₂PO₄; 1 ml Tween 20) for 1 hour to prevent non-specific antibody binding, then allowed to react for 8-10 hours with a polyclonal antibody against human α-La raised in rabbits and visualized according to the manufacturer’s protocol for colorimetric visualization of an alkaline phosphatase-conjugated anti-rabbit IgG (Bio-Rad). For dot blot analysis, 5 μl of protein extract was blotted onto a nylon–backed nitrocellulose membrane (0.45 μm), treated with 1 % ovalbumin for 1 hour then allowed to react with rabbit polyclonal antiserum against human α-La for 8-12 hours and then visualized according to the manufacturer’s protocol for colorimetric visualization of an alkaline phosphatase-conjugated anti-rabbit IgG (Bio-Rad). Tissue print analysis was performed by cutting a mature maize kernel lengthwise to expose the endosperm and the embryo section of the kernel. About 5 μl of SDS-PAGE sample
buffer were added to the exposed section of the endosperm and the embryo. After 30 minutes, this exposed section was then blotted onto a nylon-backed nitrocellulose membrane (0.45 μm) which was then treated with 1% ovalbumin in PBS-Tween buffer (8.0 g NaCl; 0.2 g KCl; 1.15 g Na₂HPO₄; 0.24 g KH₂PO₄; 1 ml Tween 20) for 1 hr., then allowed to react with a rabbit polyclonal antibody against human α-La for 8-12 hours and then visualized following to the manufacturer’s protocol for colorimetric visualization of an alkaline phosphatase-conjugated anti-rabbit IgG (Bio-Rad).

**Quantification of α-La in maize kernels by ELISA**

Protein from 10 mg of endosperm tissue was extracted with 100 μl of SDS-PAGE sample buffer (0.5M Tris-HCl pH 6.8; 10% SDS; 10% glycerol; 5% BME). α-La from the endosperm extract was captured in microtiter plates (Costar 3590, Fisher Scientific, PA, USA) for 16-18 hours at 37 °C. A 50 μl volume of endosperm extract was loaded into each well. The plates were treated to prevent non-specific antibody binding with 1% ovalbumin for one hour at room temperature. The α-La from endosperm extract was then allowed to react for 4 hours at room temperature with a rabbit polyclonal antibody against human α-lactalbumin. Alkaline phosphatase-conjugated anti-rabbit IgG (Bio-Rad) was then added for colorimetric visualization according to the manufacturer’s protocol. Absorbance at 405 nm was immediately measured spectrophotometrically over a 2-hour period in a Dynatech MRX Plate Reader (Dynex Technologies, VA, USA). OD/min (Vmax) was calculated and the levels of α-La in ng/μl were obtained using a standard with known α-La concentrations (50, 40, 30, 20, 10 ng/μl). The levels of α-La in ng/μl were then expressed in grams of α-La per 100g of tissue dry weight (g/100g). Between each step, wells were washed 3-4 times with 1x
PBS-Tween buffer (8.0 g NaCl; 0.2 g KCl; 1.15 g Na2HPO4; 0.24 g KH2PO4; 1 ml Tween 20).

*Analysis of amino acid composition, total nitrogen, carbon and crude protein content and quantification of Lysine levels using E. coli auxotrophs.*

The amino acid composition of pairs of α-La positive and negative sibling kernels was determined. About 100 mg of ground endosperm tissue from each of 3 α-La positive and negative sibling kernels were submitted to the Experiment Station Chemical Laboratories, University of Missouri-Columbia for complete amino acid analysis using a Beckmann Model 6300 Amino Acid Analyzer (AOAC-Official Methods of Analysis method 982.30E [a,b], 1995). The same amount of endosperm tissue from each of 3 α-La positive and negative sibling kernels was also submitted to the Soil and Plant Analysis Laboratory, Iowa State University to determine the proportion of total Nitrogen and Carbon content by combustion analysis. The levels of Lysine in α-La positive and negative sibling kernels were also determined by a microbial assay using an *E. coli* auxotroph by the following procedure: 10 mg of endosperm tissue were placed in a V-well plate (Nalge Nunc, International). In order to extract and hydrolyze the proteins, to each sample, 200 μl of pepsin buffer (50 ml of 0.2 M KCl; 10.6 ml of 0.2 M HCL and pepsin 0.2 mg/ml) was added. The plates were shaken for 14 -16 hours in 225 rpm at 37 °C. The plates were then centrifuged for 20 minutes at 3000 rpm. Ten μl of the resulting hydrolysate was transferred to a new plate. Four replications of each sample were assigned randomly to wells of a microtiter plate. 200 μl of inoculum [consisting of *E coli* strain KL334 in M9 media (Sambrook et al., 1989) supplemented with 1 mg/ml Lysine] were added to 10 ml of M9 media and 100 μl of this
mixture were added to each well and the samples were shaken for 20 hours. Absorbance at 595 nm was immediately measured spectrophotometrically in a Dynatech MRX Plate Reader (Dynex Technologies, VA, USA). OD/min (Vmax) was calculated and the values were regressed with a standard with known α-La concentrations (2 mM, 1 mM, 0.75 mM, 0.5 mM, 0.25 mM and 0.10 mM) to obtain the levels α-La in the samples. The levels of α-La in mM were then expressed in grams of α-La per 100g of tissue dry weight (g/100g).

Statistical Analysis

Mean levels of α-La in each event were obtained from 5-15 α-La positive sibling kernels. For data obtained by complete amino acid and combustion analysis, mean amino acid values as well as mean Nitrogen, Carbon and total protein values were obtained from 3 positive and 3 negative sibling kernels from each event. Standard error of the mean was calculated from the formula $\sigma / \sqrt{n}$; where $\sigma$ is the standard deviation and $n$ is the number of kernel samples. T-tests were performed to characterize differences between positive and negative sibling kernels in mean α-La and Lysine levels as well as in percent total Nitrogen, Carbon and Crude Protein. Percent crude protein content was obtained from the formula

$\text{Crude protein} = \% \text{ Nitrogen} \times 6.25$. Correlation was also performed to determine the linear relationship between the levels of α-La and Lysine in the endosperm.
Results

*Endosperm specific expression of α-La protein in P64*

Three constructs of a synthetic porcine α-La transgene were used to transform maize calli and plants expressing the α-lactalbumin protein were obtained as described in Chapter 2. These constructs differed in their promoters and the subcellular targeting domains. In selecting a promoter, the objective was to obtain a high level of α-lactalbumin transcription and to be able to analyze several tissues for expression. The 27 kDa gamma-zein promoter was chosen because it is the strongest compared with all the other zein promoters. (Ueda and Messing, 1991). The maize ubiquitin promoter was chosen as a compromise between high expression level and broad tissue specificity. The ubiquitin promoter has been shown to confer high level of expression in all tissues including the endosperm (Cornejo et al., 1993). Though the 27kDa gamma-zein promoter is known to direct endosperm-specific transcription, promoter activity in other tissues has been reported (Russell and Fromm, 1997; Ueing et al., 1988). It is important to establish if the 27 kDa zein promoter is sufficient to direct endosperm-specific transcription of porcine α-La. To analyze the tissue specificity of the 27 kDa zein promoter, accumulation of the α-lactalbumin protein among events in P64 was assessed in the leaf, root, pollen, embryo and endosperm tissues primarily by Western blot analysis. Combinations of techniques including dot blot and tissue print analysis were also performed.

Dot blot analysis of 6 P64 events showed that the α-La protein was present in the endosperm. This tissue specificity was observed in α-La positive kernels analyzed from all the 6 events of P64. P64 is driven by an endosperm-specific zein promoter from the 27 kDa
gamma-zein gene in maize and so it is expected that the $\alpha$-La protein will accumulate only in the endosperm. When these $\alpha$-La positive kernels were planted, $\alpha$-La was not detected in the leaf, pollen and root tissues of plants grown directly from these kernels (Figure 3.1A). Western blot analysis also revealed protein bands that reacted with a polyclonal antibody against human $\alpha$-lactalbumin when protein extracts from the endosperm were analyzed. These bands co-migrated with the human $\alpha$-La positive control indicating that the porcine $\alpha$-La accumulated in the endosperm. In contrast, no bands were observed when total protein extracts from the root, leaf and pollen tissues were allowed to react with the polyclonal antibody against human $\alpha$-lactalbumin in Western blots indicating that the porcine $\alpha$-La did not accumulate to detectable levels in these tissues. Tissue print analysis further showed that $\alpha$-La protein accumulated only in the kernel endosperm and not in the embryo (Figure 3.1B). As expected, the zein promoter directed $\alpha$-La expression only in the endosperm tissue.

**Differential expression of $\alpha$-La among events in P45.**

To analyze the activity and tissue specificity of $\alpha$-lactalbumin expression under the control of the ubiquitin promoter, accumulation of $\alpha$-La protein was also assessed in the root, leaf, pollen, embryo and endosperm tissues among events in P45. Western blot analysis revealed the porcine $\alpha$-La protein in the endosperm. However, as indicated in Figure 3.2A and 3.2B and summarized in Table 3.1, Western and tissue blot analysis showed that the $\alpha$-La protein was not detectable in the embryos of kernels expressing $\alpha$-La protein in the endosperm. This was observed in kernels analyzed from all 6 events of P45. Additionally, Western blot analysis also detected the accumulation of $\alpha$-La protein in the endosperm and
pollen in P45-11, P45-22, P45-3-3 and P45-16. Intriguingly, \( \alpha \)-La protein was neither detected in the embryo and leaf in P45-11 nor in the embryo, root and leaf tissues in P45-3, P45-22 and P45-16 (Table 3.1). P45 is driven by the ubiquitin promoter, a strong constitutive promoter, and so transcription was expected to occur in all plant cells and for the \( \alpha \)-La protein to accumulate in all plant tissues. The accumulation of the \( \alpha \)-La protein in some tissues and not in others may suggest a differential expression of the \( \alpha \)-La among events in P45 indicating that the ubiquitin promoter did not have uniform tissue specificity or worked poorly for the expression of \( \alpha \)-La in the whole plant.

*Levels of porcine \( \alpha \)-la in the endosperm of transformed kernels*

To quantify the level of porcine \( \alpha \)-La protein in the endosperm, ELISA was performed using a polyclonal antibody against human \( \alpha \)-La. Variation in the amounts of \( \alpha \)-La was observed among sibling kernels (i.e. on the same ear) in an event. The \( \alpha \)-La levels in positive kernels (5-15 kernels) in an event were averaged to obtain the mean \( \alpha \)-La value for that event. Among P64 events, the average amount of \( \alpha \)-La in the endosperm on per unit dry weight basis ranged from 0.0014 – 0.039 g/100 g endosperm tissue in F2 kernels and from 0.003 – 0.095 g/100 g endosperm tissue in F3 kernels (Table 3.2). For P45, the \( \alpha \)-La levels in the endosperm ranged from 0.023 – 0.062 g/100 g endosperm tissue in F2 kernels. Many F3 kernels among events in P45 have no detectable \( \alpha \)-La as determined by Western blot, presumably due to transgene inactivation as described in Chapter 2 and consequently, ELISA was not performed. The few positive F3 kernels obtained were planted to advance to the next generation. The levels of \( \alpha \)-La in F2 kernels among events in P64 and P45 are
similar. However, a 2-4 fold difference in α-La in F3 kernels was observed among events in P64 while many F3 kernels in 5 of 6 events in P45 have no detectable α-La protein as determined in a Western blot.

*Amino acid composition in the endosperm of kernels expressing the α-La protein*

To determine if the expression of the porcine α-La protein in the endosperm will lead to changes in amino acid composition, complete amino acid analysis was performed on α-La positive and negative sibling kernels. Sibling kernel pairs from P64-14, in which one expressed α-La and one does not, were initially compared for their relatively higher amounts of α-La protein in the endosperm as measured by ELISA (Table 3.2). As summarized in Table 3.3, the mean Lysine content of α-La positive kernels on a per unit dry weight basis was 0.24 g/100 g endosperm tissue and the mean Lysine content of α-La negative kernels was 0.20 g/100 g endosperm tissue. This is a 20 % difference in the amount of Lysine in α-La positive kernels when compared to their null siblings. However, this difference in Lysine content between α-La positive and negative sibling kernels was not significant in a t-test (P=0.12). Interestingly, the difference in Lysine was accompanied by a 21.3 % difference in the level of Aspartic Acid in α-La positive kernels when compared to their null siblings (from 0.61 g/100 g to 0.75 g/100 g dry weight), while the levels of the other amino acids were not changed (Table 3.3). This increase in Aspartic Acid content in α-La positive kernels when compared to their null siblings was significant (P=0.04). In the modified porcine α-La coding sequence, Aspartic Acid is the most frequent amino acid (15 codons of
141) (Yang et al., 2002) and so it is expected that as the α-La protein is expressed in the endosperm, Aspartic Acid levels would show the greatest change.

The data presented in Table 3.3 were obtained from only one event (P64-14). In order to obtain more information on the correlated changes in amino acid composition, three additional events from P64 and four events from P45 were selected based on their relatively higher α-La levels as measured by ELISA and a microbial assay using *E. coli* auxotrophs was performed to determine the levels of Lysine in these events. As summarized in Table 3.4, α-La positive kernels of P64-14 exhibited a 25% difference in Lysine levels when compared to their null siblings (from 0.12 g/100 g to 0.15 g/100 g dry weight). This difference in Lysine content as measured by this bacterial assay is in close agreement with the results obtained by complete amino acid analysis (Table 3.3). Similarly however, the difference in Lysine content between the positive and negative sibling kernels in this event as measured in this bacterial assay was not significant in a t-test (P=0.06). In P64-18, however, the α-La positive kernels exhibited a significant 27.27% difference in Lysine (P=0.04) when compared to their null siblings (from 0.11 g/100 g to 0.14 g/100 g dry weight). In P45-11 and P45-22, the amounts of Lysine in α-La positive kernels are higher though not significantly different from their negative sibling counterparts (Table 3.4) while in P45-3 and P45-16, α-La positive kernels have slightly lower Lysine values compared to their negative siblings. Among the events analyzed, those that accumulated relatively higher amounts of α-La in the endosperm also contained relatively higher values in terms of percent difference in Lysine content (Table 3.4). Correlation showed a positive linear relationship between levels of porcine α-La and Lysine content (r = 0.16). No significant differences were detected
between α-La positive and negative sibling kernels in percent total Nitrogen and Carbon or in total crude protein content in most events analyzed as detected in a combustion analysis except for P45-11 F2 where the percent total Nitrogen, Carbon and crude protein content were significantly lower in α-La positive kernels (Table 3.5).

To further confirm the difference in Lysine levels that were detected in the microbial assay, 4 events that displayed the highest difference in Lysine content were selected. These were P64-14, P64-18, P45-11 and P45-22 (Table 3.4). Endosperm tissue samples from 3 α-La positive and 3 α-La negative F2 and F3 sibling kernels from these events were submitted to the Experiment Station Chemical Laboratories in the University of Missouri for a complete amino acid analysis. As summarized in Table 3.6, a significant 49% difference in Lysine content was detected in α-La positive kernels of P64-18 F2 when compared to their negative siblings (0.25 g/100g vs 0.17 g/100 endosperm tissue; P=0.05). Also, a significant 35% difference in Lysine content was detected in α-La positive kernels of P64-14 F2 when compared to their negative siblings (0.27 g/100g vs 0.20 g/100g endosperm tissue; P=0.006). When F3 kernels were analyzed, a similar difference in Lysine content was detected in α-La positive F3 kernels from P64-18 and P64-14 when compared to their negative siblings. The higher Lysine content in α-La positive kernels in P64-14 F2 and P64-18 F2 as detected in this complete amino acid assay is in agreement with the results obtained in the initial complete amino acid analysis of P64-14 F2 (Table 3.3) as well as with the results obtained by microbial analysis (Table 3.4). Intriguingly, the α-La positive kernels of P45-11 F2 exhibited a significantly lower (-16.7%) Lysine content when compared to their negative siblings (0.15 g/100g vs 0.18 g/100g endosperm tissue; P=0.016) while no
significant difference in Lysine levels were detected between α-La positive and negative kernels in P45-22 F2. The significantly lower Lysine content in the α-La positive kernels in P45-11 F2 as detected by this complete amino acid assay is in contrast to the results obtained by microbial assay where the α-La positive kernels of this event contained higher Lysine values when compared their negative siblings (Table 3.4; this difference though is not significant at P=0.13).

The difference in Lysine levels as detected in this complete amino acid assay was also accompanied by changes in levels of the other amino acids (Table 3.7). A significant 6.1 % difference in Threonine and 7.1 % difference in Cysteine was detected in P64-14 F2. Also, a significant 8.1 % difference in Isoleucine was detected in P64-14 F2 and 16.9 % difference in P64-14 F3 while an 11.7 % difference in Aspartic Acid was detected in P64-14 F3. Additionally, differences in the levels of several amino acids were also detected in P64-18 F2 which include among others a 24.7 % difference in Aspartic acid, 21.5 % in Serine, 22.2 % difference in Tyrosine and 21.4 % difference in Methionine. Moreover, a 13.9 % difference in Aspartic Acid and a 10.8 % difference in Tyrosine was also detected in P64-18 F3. For some amino acids, a significant difference was detected only in F2 kernels but not in F3 kernels such as Threonine and Cysteine in P64-14 F2 and P64-14 F3 while for others, the change involved a significantly lower amino acid levels in α-La positive kernels, such as Serine and Proline in P64-14 F3 (Table 3.7). Interestingly, in P45-11 F2, the levels of all the amino acids were significantly lower in α-La positive kernels. It is also noteworthy that Lysine exhibited the lowest change (-16.3 %) when compared to the rest of the amino acids.
In terms of total crude protein, α-La positive kernels in P64-18 F2 showed a significant 10.3% difference in total crude protein content compared to its negative siblings (Table 3.7) while in P45-11 F2, total crude protein was significantly lower (-22.3%) in α-La positive kernels. No significant change in total crude protein was detected between α-La positive and negative sibling kernels in the rest of the events analyzed.

Discussion

A synthetic gene encoding the modified milk protein porcine α-La has been transformed into maize in an effort to improve the nutritional value of the maize grain. In P64 events driven by the 27 kDa gamma-zein promoter, analysis of the tissue specificity of α-La expression showed that the α-La protein accumulated only in the endosperm (Figure 3.2a and 3.2b). The α-La protein was not detected in other tissues (i.e embryo, root, leaf and pollen) demonstrating that the zein promoter is sufficient and appropriate for an endosperm-specific expression of porcine α-La protein in maize. This result is comparable to other studies that used endosperm-specific promoters (Lamacchia et al., 2001; Stoger et al., 1999). For example in wheat, GUS activity was detected specifically in the endosperm and not in other tissues such as the aleurone. The UidA gene encoding β-glucuronidase (GUS) was under the control of the endosperm-specific Glu-ID-1 promoter (Lamacchia et al., 2001). The results of this study do not support the findings of Russell and Fromm (1997a). These authors detected a strong GUS activity in the endosperm for a UidA transgene under the control of the 27 kDa zein promoter but also reported significant low level GUS expression in other parts of the kernel including the embryo.
Differential tissue specificity of α-La was detected among events in P45. The α-La protein accumulated primarily in the endosperm but not in the embryo. In some events, the α-La protein was also detected in the roots and pollen but not in the leaves (i.e. P45-11, Table 3.1). These results were unexpected as the ubiquitin promoter is a strong constitutive promoter that has been widely used for the expression of transgenes in several crops (Christensen and Quail, 1996; Cornejo et al., 1993; Stoger et al., 1999). In this study, the ubiquitin promoter did not show constitutive activity. It is highly plausible that the non-detection of α-La protein in some tissues may be due to a quantitative difference in the levels of the α-La protein. It is likely that α-lactalbumin was highly expressed in the endosperm but expressed at very low amounts in the embryo and other tissues so that the α-La protein in these low expressing tissues are below the detection limits. Quantitative differences in the expression levels of transgenes driven by the ubiquitin promoter have been reported. In wheat, highest expression levels for GUS were found in the endosperm and low transgene expression was detected in some tissues, while no GUS activity was detected in other tissues examined (Stoger et al., 1999). In some events, GUS expression was confined in the pollen (Stoger et al., 1999) while in a related study, a ubiquitin promoter-driven β-glucorinidase (GUS) delivered by particle bombardment showed GUS expression only in the embryo (Muhitch and Shatters, 1998). Also, the ubiquitin promoter showing strong tissue preference in the kernel embryo has been reported in transgenic maize (Hood et al., 1997). The differential tissue specificity of α-La expression and the higher frequency of events that showed transgene inactivation in P45, suggest that the ubiquitin promoter may not be the appropriate promoter for a consistent and stable expression of α-La.
The levels of α-La protein that were detected in the endosperm vary among events in a construct. For P64, the mean levels of α-La among events range from 0.0014 – 0.062 g/100g endosperm tissue in F2 kernels and from 0.003 – 0.095 g/100g endosperm tissue in F3 kernels as summarized in Table 3.4. For P45, the mean levels of porcine α-La among events range from 0.023 – 0.062 g/100 endosperm tissue. This event to event variation is somewhat desirable as this allowed selection of events with relatively higher α-La values for amino acid composition analysis. Variation in the levels of foreign protein among events in a construct has been observed in several transgenic crops and has been postulated to be partly due to position effects - the chromosomal site where the transgene integrated into the maize genome - and to differences in copy number (Gendloff et al., 1990). It was shown in chapter 2 that independent events of P64 and P45 constructs had different transgene integration sites as indicated by their unique Southern blot hybridization patterns. It is not unlikely that for some events, the α-La transgene may have been integrated nearby hypermethylated chromosomal regions resulting in reduced α-La expression (Ye and Signer, 1996). Also, each event differed in α-La copy number as each event displayed multiple hybridizing bands. Variation in the levels of the α-La was also detected among sibling kernels from an event. While part of this can be due to experimental error, measurement was performed in duplicate to reduce experimental error. This variation in α-La levels among sibling kernels in an event could also be attributed to the microenvironment in which the kernels developed. Additionally, this variation could be due to the unique genotype of each sibling kernel as these kernels were from segregating F2 and F3 ears. The F2 ear was obtained by self-pollination of an F1 plant derived from a cross between T0 and B73 plants. For this reason, the α-La levels from
positive kernels in an event were averaged to represent the mean $\alpha$-La level for that event. In P64 events, the levels of $\alpha$-La in F3 kernels were 2-4-fold higher when compared to the $\alpha$-La levels in F2 kernels (Table 3.2). From these data, it may be possible to conclude that the levels of $\alpha$-La can be improved by self-pollination. Self-pollination in succeeding generations may increase the dose of the transgene in the endosperm, which might result in increased $\alpha$-La levels. In other studies, subsequent generations have been shown to have increasing levels of the novel protein when compared to the F1 generation (Zhong et al., 1999). In this study however, the levels of $\alpha$-La in F2 and F3 kernels are confounded with the environmental effects. The F2 and F3 kernels were grown in different environments and so comparison of $\alpha$-La levels in the F2 and F3 kernels is not appropriate.

One important reason to quantify the levels of porcine $\alpha$-La in the endosperm is to determine if expression of porcine $\alpha$-La would have an observable effect on amino acid composition. A question of particular interest was do high levels of Lysine accumulate in the endosperm of high $\alpha$-La kernels? Expression of the modified porcine $\alpha$-La protein in the endosperm resulted in an increase in levels of Lysine. Two events, P64-14 and P64-18, showed relatively higher levels of porcine $\alpha$-La protein in the endosperm as measured in an ELISA and these events displayed consistently higher Lysine levels in a microbial assay and in a complete amino acid composition analysis using F2 and F3 kernels. The percent difference in Lysine levels ranges from 20 % to 35 % for P64-14 and from 21 % to 49 % for P64-18. While the two methods of analysis are different and so the Lysine values obtained from each method can not be directly compared, the difference in Lysine levels as detected consistently in P64-14 and P64-18 demonstrate that in these events the levels of Lysine in the
plant (or in the maize kernel) were improved as a result of the expression of the modified porcine α-La, indicating that the use of a synthetic gene encoding a modified milk protein can indeed be a useful strategy to improve maize grain quality.

In addition to Lysine, there were also significant changes in the content of other amino acids of the endosperm as a result of the expression of porcine α-La (Tables 3.7). There was a 24.7 % difference in Aspartic Acid in P64-18 F2, 13.9 % difference in P64-18 F3 and an 11.7 % difference in P14-4 F3 kernels. A significant 20 % difference in Aspartic Acid levels was also observed in an initial complete amino acid composition analysis of P64-14 F2 kernels. In the modified porcine α-La coding sequence, Aspartic Acid is the most frequent amino acid (15 of 141) followed by Lysine (12 of 141) (Yang et al., 2002). It is expected that as the maize plant expressed α-La, changes in the levels of Aspartic Acid and Lysine would be observed. Also, significant changes in Threonine and Cysteine were detected in P64-14 F2 and Isoleucine in P64-14 F2 and P64-14 F3 while no change was detected in the other amino acids. In P64-18 F2, expression of porcine α-La resulted in significant changes in the levels of several amino acids including a 21.5 % difference in Serine, 22.2 % difference in Tyrosine and a 21.4 % difference in Methionine. These significant changes in the levels of several amino acids might have contributed to the significant increase in total protein content detected in P64-18 F2. Significant changes in the amino acid composition as a result of the expression of a foreign protein has been reported in several crops. In lupin, which is deficient in sulfur containing-amino acids Methionine and Cysteine, expression of a sulfur-rich sunflower seed albumin resulted in a 94% increase in Methionine but a 12% reduction in the amount of Cysteine was also observed. In addition, there was no significant change in other amino acids or in total nitrogen or sulfur content of
the seeds (Molvig et al., 1997). In potato, expression of a seed albumin gene from *Amaranthus hypochondriacus* did not only increase Lysine, but a significant 2 to 4-fold increase in Methionine, Cysteine and Tyrosine contents, which are otherwise limited in potato, were also detected. Moreover, this increase in amino acids was accompanied by an increase in total protein content as well as a striking increase in the size of tubers (Chakraborty et al., 2000).

While the expression of the modified porcine α-La resulted in significant differences in Lysine content, a question of particular interest is whether the detected changes in Lysine levels were due to the amount of Lysine contained in the modified porcine α-La? From the current data, it seemed that the changes in Lysine levels can not be solely attributed to the expression of the porcine α-La. For example, in P64-18 F2, the amount of α-La protein in this event as measured by ELISA was 0.025 g/100g endosperm tissue (Table 3.4). The level of Lysine in this event as measured in a complete amino acid analysis assay was 0.25 g/100g endosperm tissue (Table 3.6). Lysine constitutes about 14% of the total porcine α-La protein. Therefore, from its current level of 0.025 g/100g in P64-18, the porcine α-La is expected to supply only about 0.0035 g/100g of Lysine (or 1.4 %) in the total Lysine detected in α-La positive kernels (0.25 g/100g) in P64-18. So about 0.246 g/100g of Lysine (or 98.6%) detected in α-La positive kernels were not contributed by the porcine α-La protein. This amount of Lysine may have been contributed by the other proteins in the endosperm. How the expression of the modified porcine α-La caused Lysine levels in other proteins to change is not clear. It is possible that an interaction occurred between the α-La transgenes and other
maize genes that are involved in the synthesis of storage proteins resulting in altered amino acid profile and higher Lysine levels in the endosperm.

In this study, abnormal phenotypes were detected under field conditions from plants derived from α-La positive and negative kernels including stunted growth, abnormal anther dehiscence, reduced male fertility and seed set and early senescence. In P64-14 and P64-15, frequent occurrence of albino seedlings and chlorotic plants were detected in F2 and F3 plants from these events. With the current data, it is not possible to ascertain if these phenotypes are associated with the expression of porcine α-La or due to somaclonal variation arising from the stress in tissue culture.

It is interesting that P45-11 contained relatively higher α-La values than P64-14 and P64-18 F2 plants (Table 3.2) and the α-La positive kernels in this event showed a significantly lower amino acid levels including a 16.6 % reduction in Lysine levels as measured in a complete amino acid composition analysis. Also, a significantly lower (22.28 %) total protein content was detected in this event. It is possible that an interaction occurred between α-La transgenes and maize genes that are involved in the synthesis of storage proteins, resulting in an altered amino acid profile and reduced total protein content in the endosperm. In P45-22, expression of α-La did not result in substantial changes in amino acid composition. It is possible that the level α-La in this event (0.062 g/100g) was not sufficient to change the amino acid composition or to increase the amount of Lysine to desired levels.

Efforts need to focus on expressing α-La to higher levels in order to achieve the desired improvement in maize grain quality. One strategy is the inclusion of enhancer
sequences in the construct to improve transgene expression. This approach has been found to improve the expression of a Methionine-rich storage 2S albumin from Brazil nut (Aragao et al., 1999; Datla et al., 1993). Additionally, the inclusion of MAR (matrix attachment region) sequences in transgene constructs was also found to reduce variation in transgene activity and prevented lower levels of transgene expression (Mylnarova et al., 1994; Petersen et al., 2002; Vain et al., 1999).

The α-La expression levels can also be optimized by choosing the right promoters. The low levels of human α-La detected in tobacco have been attributed to the intrinsic properties of 35S promoter used to direct human α-La transcription and not to the codon usage or signal peptide used. (Takase and Hagiwara, 1998). In this study, it was demonstrated that the 27 kDa zein promoter can direct stable and high levels of α-La expression in the endosperm. The 27 kDa zein promoter appears to be the strongest compared with all other zein gene promoters (Ueda and Messing, 1991). This is partly due to the existence of cis–elements that regulate transcription. One of these cis-regulatory elements is the prolamin box (P-box), a highly conserved 7-bp sequence motif which is found in the promoters of many cereal seed storage protein genes and generally located about 300 bp upstream of the ATG translation initiation point in most zein genes (Boronat et al., 1986). The P-box has been found to enhance transcription when fused to truncated zein promoters (Ueda et al., 1994) and has been suggested to play a central role in the common regulatory mechanism and mediate the endosperm-specific expression of zein genes (Ueda et al., 1994). An endosperm nuclear factor called prolamin box-binding factor 1 (PBF-1) encodes a 38 kDa protein that is synthesized only in the endosperm tissue (Wang et al., 1998) and specifically recognizes the P-box of the 27 kDa zein promoter or promoters of genes encoding zeins and
other endosperm-specific proteins. The PBF-1 is thought to have domains that can interact with other proteins to increase gene expression (Wang et al., 1998). It was suggested that PBF-1 might function as a recruiting protein and enhance the interaction with transcription factors. In addition to using a strong promoter such as the 27 kDa gamma-zein promoter, an alternative strategy to further increase transcription is to transfer or introduce the transgene into suitable genetic background such as the opaque 2 background. The transcription factor Opaque-2 (O2) is present in the opaque 2 background. The 27 kDa zein gene promoter does not have an O2 binding site unlike the 22 kDa and 15 kDa gene promoters (Neto et al., 1995; Schmidt et al., 1992) but it is conceivable that PBF-1 enhances the interaction with O2 or other class-specific transcription factors and position transcription factors like the O2 closer to the TATA-box in the 27 kDa zein gene promoter thereby enhancing transcription.

In summary, it was demonstrated that the modified porcine α-La protein accumulated in the endosperm in constructs driven by the 27 kDa gamma-zein promoter (P64) and ubiquitin promoter (P45). The data indicate that the 27 kDa zein promoter directed higher levels and a more stable expression of α-La protein in the endosperm than the ubiquitin promoter. The expression of α-La resulted in a significant difference in Lysine content and altered the kernel amino acid composition, indicating that the plant was able to increase its supply of Lysine for protein synthesis as a result of the expression of the modified porcine α-Lactalbumin. However, efforts need to focus on developing strategies to express α-La to higher levels, as the current amounts obtained may not be sufficient to increase Lysine content to the levels desired for improved maize grain quality.
References


Figure 3.1  A. Dot blot analysis to detect the presence of α-La protein in the root, leaf, pollen and endosperm tissues among events in P64 construct. H= human α-La as positive control. B. Tissue print analysis to detect the presence of the synthetic porcine α-La in the embryo and endosperm of a transformed kernel. En = endosperm portion of the kernel; Em = embryo portion of the kernel. In both assays, accumulation of the porcine α-lactalbumin protein was detected only in the endosperm and not in other tissues indicating an endosperm-specific expression of the porcine α-lactalbumin protein in P64.
Figure 3.2 A. Western blot analysis to detect the presence of the synthetic porcine \( \alpha \)-La protein in the endosperm and embryo in P45 construct. Ha = human \( \alpha \)-La as positive control. B73 = untransformed kernel as negative control. En = protein extracts from the endosperm; Em = protein extracts from the embryo of the same kernel. 125.1-44 and 125.1-45 are positive siblings from event P45-16. B. Tissue print analysis to detect the accumulation of the synthetic porcine \( \alpha \)-La in the embryo and endosperm of a transformed kernel. En= endosperm portion of the kernel; Em = embryo portion of the kernel. In both assays, accumulation of the porcine \( \alpha \)-lactalbumin protein was detected only in the endosperm and not in the embryo suggesting a differential expression of the porcine \( \alpha \)-lactalbumin in P45.
Table 3.1. Summary of α-La protein accumulation in different tissues among events in P45 as detected by Western and dot blot analysis.

<table>
<thead>
<tr>
<th>Event</th>
<th>F3 siblings</th>
<th>Tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Endosperm</td>
</tr>
<tr>
<td>P45-11-3</td>
<td>31-12</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>31-7</td>
<td>-</td>
</tr>
<tr>
<td>P45-22-1</td>
<td>32-8</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>32-7</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>32-12</td>
<td>-</td>
</tr>
<tr>
<td>P45-19-7</td>
<td>33-14</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>33-8</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>33-9</td>
<td>-</td>
</tr>
<tr>
<td>P45-3-3</td>
<td>34-15</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>34-5</td>
<td>-</td>
</tr>
<tr>
<td>P45-16-2</td>
<td>35-7</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>35-8</td>
<td>+</td>
</tr>
</tbody>
</table>

(-) no detectable α-La protein; (+) presence of α-La protein.
Table 3.2 Levels of porcine α-lactalbumin protein in the endosperm of transformed maize kernels among different events of P64 and P45 constructs as measured by ELISA.

<table>
<thead>
<tr>
<th>Events</th>
<th>α-La ± SE (g/100g)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F2</td>
<td>F3</td>
</tr>
<tr>
<td>P64-18-5</td>
<td>0.0250 ± 0.0014</td>
<td>0.089 ± 0.020</td>
</tr>
<tr>
<td>P64-14-4</td>
<td>0.0390 ± 0.0065</td>
<td>0.095 ± 0.020</td>
</tr>
<tr>
<td>P64-24-4</td>
<td>0.0072 ± 0.0012</td>
<td>0.012 ± 0.0018</td>
</tr>
<tr>
<td>P64-1-4</td>
<td>0.0014 ± 0.0001</td>
<td>0.003 ± 0.0001</td>
</tr>
<tr>
<td>P64-15-6</td>
<td>-</td>
<td>0.074 ± 0.025</td>
</tr>
<tr>
<td>P45-11-3</td>
<td>0.062 ± 0.0001</td>
<td>*</td>
</tr>
<tr>
<td>P45-3-3</td>
<td>0.039 ± 0.0096</td>
<td>*</td>
</tr>
<tr>
<td>P45-16-2</td>
<td>0.023 ± 0.0090</td>
<td>*</td>
</tr>
<tr>
<td>P45-22-1</td>
<td>0.062 ± 0.0462</td>
<td>*</td>
</tr>
</tbody>
</table>

* Values are means of five to fifteen positive sibling kernels. F2 and F3 were grown in different environments. * = kernels showed transgene inactivation as described in Chapter 2. Consequently, levels of α-La protein were not measured by ELISA.
Table 3.3. Amino acid composition of maize kernels positive and negative for α-La protein in the endosperm as measured by complete amino acid analysis in P64-14.

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Mean ± SE(^a) (g/100g)</th>
<th>% difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>0.74 ± 0.04</td>
<td>0.61 ± 0.04</td>
</tr>
<tr>
<td>Threonine</td>
<td>0.35 ± 0.02</td>
<td>0.33 ± 0.02</td>
</tr>
<tr>
<td>Serine</td>
<td>0.45 ± 0.03</td>
<td>0.44 ± 0.02</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>2.27 ± 0.22</td>
<td>2.22 ± 0.12</td>
</tr>
<tr>
<td>Proline</td>
<td>0.90 ± 0.08</td>
<td>0.91 ± 0.05</td>
</tr>
<tr>
<td>Glycine</td>
<td>0.33 ± 0.02</td>
<td>0.31 ± 0.02</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.86 ± 0.08</td>
<td>0.84 ± 0.05</td>
</tr>
<tr>
<td>Cysteine</td>
<td>0.23 ± 0.02</td>
<td>0.22 ± 0.02</td>
</tr>
<tr>
<td>Valine</td>
<td>0.51 ± 0.03</td>
<td>0.50 ± 0.03</td>
</tr>
<tr>
<td>Methionine</td>
<td>0.21 ± 0.01</td>
<td>0.22 ± 0.01</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>0.46 ± 0.03</td>
<td>0.42 ± 0.02</td>
</tr>
<tr>
<td>Leucine</td>
<td>1.64 ± 0.17</td>
<td>1.60 ± 0.08</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>0.34 ± 0.03</td>
<td>0.33 ± 0.02</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>0.61 ± 0.05</td>
<td>0.59 ± 0.03</td>
</tr>
<tr>
<td>Histidine</td>
<td>0.27 ± 0.01</td>
<td>0.26 ± 0.01</td>
</tr>
<tr>
<td><strong>Lysine</strong></td>
<td><strong>0.24 ± 0.02</strong></td>
<td><strong>0.20 ± 0.02</strong></td>
</tr>
<tr>
<td>Arginine</td>
<td>0.34 ± 0.01</td>
<td>0.32 ± 0.03</td>
</tr>
<tr>
<td><strong>Crude Protein</strong></td>
<td><strong>10.79 ± 0.87</strong></td>
<td><strong>10.36 ± 0.62</strong></td>
</tr>
</tbody>
</table>

\(^a\) = mean values are from three F2 sibling kernels from each positive and negative class ± standard error.\(^*\) = difference in values between positive and negative kernels is significant in a t-test at \(\alpha=0.05\); \(\text{ns}\) = not significant.
Table 3.4. Correlated levels of Lysine among events in P64 and P45 expressing $\alpha$-La in the endosperm as measured by a microbial assay using *E. coli* auxotrophs.

<table>
<thead>
<tr>
<th>Events</th>
<th>$\alpha$-La $\pm$ SE $^a$ (g/100g)</th>
<th>Lysine $\pm$ SE $^b$ (g/100g)</th>
<th>% difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
<td></td>
</tr>
<tr>
<td>P64-18-5</td>
<td>0.025 $\pm$ 0.0014</td>
<td>0.14 $\pm$ 0.005</td>
<td>27.27 $^*$</td>
</tr>
<tr>
<td>P64-14-4</td>
<td>0.038 $\pm$ 0.0065</td>
<td>0.15 $\pm$ 0.012</td>
<td>25.00 ns</td>
</tr>
<tr>
<td>P64-24-4</td>
<td>0.007 $\pm$ 0.0012</td>
<td>0.13 $\pm$ 0.008</td>
<td>18.00 ns</td>
</tr>
<tr>
<td>P45-11-3</td>
<td>0.062 $\pm$ 0.0001</td>
<td>0.11 $\pm$ 0.007</td>
<td>11.6 ns</td>
</tr>
<tr>
<td>P45-3-3</td>
<td>0.038 $\pm$ 0.0096</td>
<td>0.11 $\pm$ 0.005</td>
<td>-0.90 ns</td>
</tr>
<tr>
<td>P45-16-2</td>
<td>0.023 $\pm$ 0.0090</td>
<td>0.11 $\pm$ 0.011</td>
<td>-9.35 ns</td>
</tr>
<tr>
<td>P45-22-1</td>
<td>0.063 $\pm$ 0.0462</td>
<td>0.13 $\pm$ 0.011</td>
<td>30.00 ns</td>
</tr>
</tbody>
</table>

$^a$ $\alpha$-La values are means of 5-15 positive sibling kernels $\pm$ SE. $^b$ For Lysine, endosperm tissues from 5 siblings kernels from each positive and negative class were bulked. Lysine levels were measured in a microbial assay using this bulked sample. Lysine values presented are means of 4 replications in a microbial assay $\pm$ SE. $^* =$ difference in Lysine content between $\alpha$-La positive and negative sibling kernels is significant in a t-test at $\alpha=0.05$; ns = not significant. Correlation between $\alpha$-La and Lysine levels is $r=0.16$. 

106
Table 3.5. Mean percent total Carbon, Nitrogen and Crude Protein in maize endosperms positive and negative for the porcine α-La protein as measured in a combustion analysis.

<table>
<thead>
<tr>
<th>Event</th>
<th>Mean % Total C ± SE</th>
<th>t-test</th>
<th>Mean % Total N ± SE</th>
<th>t-test</th>
<th>Mean % Crude Protein ± SE</th>
<th>t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>positive</td>
<td>negative</td>
<td></td>
<td>positive</td>
<td>negative</td>
<td></td>
</tr>
<tr>
<td>P64-14 F2</td>
<td>40.73 ± 0.005</td>
<td>40.76 ± 0.050</td>
<td>0.26 ns</td>
<td>1.65 ± 0.017</td>
<td>1.62 ± 0.040</td>
<td>0.44 ns</td>
</tr>
<tr>
<td>P64-14 F3</td>
<td>40.89 ± 0.104</td>
<td>40.78 ± 0.017</td>
<td>0.17 ns</td>
<td>1.93 ± 0.070</td>
<td>1.83 ± 0.060</td>
<td>0.23 ns</td>
</tr>
<tr>
<td>P64-18 F3</td>
<td>40.86 ± 0.058</td>
<td>40.36 ± 0.649</td>
<td>0.24 ns</td>
<td>1.90 ± 0.020</td>
<td>1.95 ± 0.107</td>
<td>0.34 ns</td>
</tr>
<tr>
<td>P45-11 F2</td>
<td>40.51 ± 0.040</td>
<td>40.83 ± 0.047</td>
<td>0.004 *</td>
<td>1.35 ± 0.014</td>
<td>1.60 ± 0.031</td>
<td>0.001 *</td>
</tr>
<tr>
<td>P45-22 F2</td>
<td>40.69 ± 0.058</td>
<td>40.92 ± 0.015</td>
<td>0.06 ns</td>
<td>1.62 ± 0.049</td>
<td>1.66 ± 0.039</td>
<td>0.31 ns</td>
</tr>
</tbody>
</table>

*Values are means of three positive and negative sibling kernels ± standard error. Combustion analysis was performed by the ISU Soil and Plant Analysis Laboratory. * = difference between positive and negative values is significant in a t-test at α=0.05.; ns= not significant
b Values for % Crude Protein were obtained by % N x 6.25.
Table 3.6. Mean Lysine levels and Total amino acids in maize endosperms positive and negative for the porcine α-La protein from selected events of P64 and P45 constructs as measured by a complete amino acid analysis.

<table>
<thead>
<tr>
<th>Event</th>
<th>Mean Lysine ± SE *</th>
<th>% difference</th>
<th>Mean Total Amino Acid ± SE *</th>
<th>% difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive (g/100g)</td>
<td>Negative (g/100g)</td>
<td>Positive (g/100g)</td>
<td>Negative (g/100g)</td>
</tr>
<tr>
<td>P64-14 F2</td>
<td>0.27 ± 0.014</td>
<td>0.20 ± 0.005</td>
<td>11.15 ± 0.156</td>
<td>11.27 ± 0.181</td>
</tr>
<tr>
<td>P64-14 F3</td>
<td>0.22 ± 0.008</td>
<td>0.17 ± 0.003</td>
<td>11.90 ± 0.054</td>
<td>11.79 ± 0.139</td>
</tr>
<tr>
<td>P64-18 F2</td>
<td>0.25 ± 0.040</td>
<td>0.17 ± 0.005</td>
<td>11.55 ± 0.470</td>
<td>10.47 ± 0.046</td>
</tr>
<tr>
<td>P64-18 F3</td>
<td>0.24 ± 0.016</td>
<td>0.18 ± 0.008</td>
<td>12.63 ± 0.375</td>
<td>12.15 ± 0.276</td>
</tr>
<tr>
<td>P45-11 F2</td>
<td>0.15 ± 0.003</td>
<td>0.18 ± 0.008</td>
<td>8.80 ± 0.256</td>
<td>11.32 ± 0.397</td>
</tr>
<tr>
<td>P45-22 F2</td>
<td>0.17 ± 0.006</td>
<td>0.17 ± 0.000</td>
<td>11.19 ± 0.089</td>
<td>10.89 ± 0.234</td>
</tr>
</tbody>
</table>

* Values are means of three sibling kernels from each positive and negative class ± standard error. *= difference between positive and negative values is significant in a t-test at α = 0.05; ns = not significant.
Table 3.7. Percent difference in amino acid and protein content between maize kernels positive and negative for porcine α-La protein in the endosperm in events of P64 and P45.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>P64-14 F2 % difference</th>
<th>P64-14 F3 % difference</th>
<th>P64-18 F2 % difference</th>
<th>P64-18 F3 % difference</th>
<th>P45-11 F2 % difference</th>
<th>P45-22 F2 % difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspartic Acid</td>
<td>7.11 ns</td>
<td>11.73 *</td>
<td>24.73 *</td>
<td>13.95 *</td>
<td>-20.73 *</td>
<td>2.76 ns</td>
</tr>
<tr>
<td>Threonine</td>
<td>6.06 *</td>
<td>1.89 ns</td>
<td>17.89 *</td>
<td>5.41 ns</td>
<td>-22.64 *</td>
<td>2.00 ns</td>
</tr>
<tr>
<td>Serine</td>
<td>5.93 ns</td>
<td>-6.43 *</td>
<td>21.49 *</td>
<td>6.67 ns</td>
<td>-21.58 *</td>
<td>5.38 ns</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>-0.99 ns</td>
<td>0.00 ns</td>
<td>11.18 *</td>
<td>-2.86 ns</td>
<td>-25.10 *</td>
<td>1.99 ns</td>
</tr>
<tr>
<td>Proline</td>
<td>-4.31 ns</td>
<td>-6.32 *</td>
<td>-1.86 ns</td>
<td>-7.98 ns</td>
<td>-19.43 *</td>
<td>1.15 ns</td>
</tr>
<tr>
<td>Glycine</td>
<td>0.00 ns</td>
<td>1.33 ns</td>
<td>19.12 *</td>
<td>2.67 ns</td>
<td>-20.00 *</td>
<td>-2.86 ns</td>
</tr>
<tr>
<td>Alanine</td>
<td>-0.76 ns</td>
<td>-1.06 ns</td>
<td>12.99 *</td>
<td>-0.97 ns</td>
<td>-24.55 *</td>
<td>3.49 ns</td>
</tr>
<tr>
<td>Cysteine</td>
<td>7.14 *</td>
<td>2.63 ns</td>
<td>5.88 ns</td>
<td>-1.47 ns</td>
<td>-16.67 *</td>
<td>1.41 ns</td>
</tr>
<tr>
<td>Valine</td>
<td>3.21 ns</td>
<td>1.94 ns</td>
<td>6.80 ns</td>
<td>-6.71 ns</td>
<td>-21.02 *</td>
<td>1.33 ns</td>
</tr>
<tr>
<td>Methionine</td>
<td>9.30 ns</td>
<td>-1.09 ns</td>
<td>21.43 *</td>
<td>6.67 ns</td>
<td>-16.87 *</td>
<td>6.90 ns</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>8.06 *</td>
<td>16.95 *</td>
<td>15.93 *</td>
<td>-3.55 ns</td>
<td>-21.60 *</td>
<td>4.35 ns</td>
</tr>
<tr>
<td>Leucine</td>
<td>-0.40 ns</td>
<td>-0.36 ns</td>
<td>12.35 *</td>
<td>-0.49 ns</td>
<td>-23.14 *</td>
<td>3.84 ns</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>-1.71 ns</td>
<td>-2.54 ns</td>
<td>22.22 *</td>
<td>10.83 *</td>
<td>-28.23 *</td>
<td>1.74 ns</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>1.08 ns</td>
<td>2.03 ns</td>
<td>15.25 *</td>
<td>-0.89 ns</td>
<td>-23.59 *</td>
<td>4.42 ns</td>
</tr>
<tr>
<td>Histidine</td>
<td>0.00 ns</td>
<td>1.16 ns</td>
<td>-1.30 ns</td>
<td>-9.64 ns</td>
<td>-18.82 *</td>
<td>2.47 ns</td>
</tr>
<tr>
<td>Lysine</td>
<td>33.33 *</td>
<td>28.85 *</td>
<td>49.02 *</td>
<td>33.96 *</td>
<td>-16.36 *</td>
<td>-1.96 ns</td>
</tr>
<tr>
<td>Arginine</td>
<td>2.68 ns</td>
<td>1.85 ns</td>
<td>15.38 ns</td>
<td>0.93 ns</td>
<td>-21.24 *</td>
<td>0.95 ns</td>
</tr>
<tr>
<td>Crude Protein</td>
<td>-1.06 ns</td>
<td>0.93 ns</td>
<td>10.31 *</td>
<td>3.98 ns</td>
<td>-22.28 *</td>
<td>2.81 ns</td>
</tr>
</tbody>
</table>

* = difference between positive and negative values is significant in a t-test at α = 0.05; ns = not significant.
Chapter 4. The inheritance and expression patterns of the selectable marker gene BAR in maize (Zea mays L.).

Earl H. Bicar¹, Paul M. Scott²* and Michael Lee¹*

Key words: transgene, porcine α-lactalbumin, BAR, expression, inheritance

Abstract

In many studies on transgenes, much emphasis is given on the inheritance and expression of the transgene of interest while the inheritance and expression of the selectable marker genes are largely ignored despite the potential of selectable marker genes to contribute to the expression patterns of the transgene of interest. Three constructs that contain the synthetic porcine α-La and the selectable marker gene BAR were introduced into maize callus cells in an attempt to improve maize grain quality. The objectives of this study were to: 1) confirm the integration of the BAR transgene into the maize genome; 2) determine whether the BAR gene is functional; 3) assess the stability of BAR integration; and 4) evaluate the inheritance and expression patterns of BAR for three generations.

¹Department of Agronomy, Iowa State University, Ames, IA 50011
²USDA-ARS Corn Insects and Crop Genetics Research Unit
*Authors for correspondence
BAR DNA sequences were detected by PCR analysis of progenitor F1 plants from 7 events of P64 construct, 6 events of P45 construct and 4 events of P57 construct, indicating that the BAR transgene was integrated into the maize genome and transmitted from the T0 to the F1 generation. In a herbicide assay, the progenitor F1 plants that contained BAR sequences by PCR analysis were resistant to the herbicide Liberty, indicating that the BAR gene was not only transmitted to the F1 generation but it was also functional. In a Southern blot analysis, each progenitor F1 plant derived from different events of P64, P45 and P57 constructs had unique integration patterns and the integration patterns in the F2 and F3 progenies were identical to their corresponding F1’s, indicating that no new rearrangements were detected, demonstrating a stable integration of the BAR gene. In a phenotypic segregation analysis of the α-La and BAR transgenes, 5 P64 events and 4 P45 events contained F2 phenotypes that mostly belong to parental classes (i.e. α-La positive kernel/herbicide resistant plant or α-La negative kernel/herbicide susceptible plant). In 3 of the 5 P64 events, no recombinant phenotypes were observed suggesting that in these events the α-La and BAR transgenes co-segregated and were linked.

The expression and inheritance of BAR was assessed for 3 generations. In 6 P64 events, both BAR and the α-La transgenes were expressed and inherited as single locus in the F2 and F3 generations. In 6 P45 events, 3 events showed single locus inheritance and consistent expression of BAR in the F2 and F3 generations while in the remaining 3 events the α-La protein and herbicide resistant plants were not detected despite the presence of the α-La and BAR transgene DNA sequences in PCR and Southern blot analyses, indicating that both the α-La and BAR transgenes were present but not functional. For the 2 events of P57,
herbicide resistant plants were consistently detected and BAR was inherited as a single locus in F2 and F3 generations while the $\alpha$-La protein was not detected in kernels that were positive for the presence of $\alpha$-La transgene DNA sequences in PCR and Southern blot analyses suggesting that the $\alpha$-La transgene was inactive.

Introduction

The success of the genetic modification of plants is determined by both the ability to deliver foreign DNA into the host plant and by the efficiency with which shoots or somatic embryos can be regenerated from transformed cells. The poor transformation efficiency of most crops necessitates the use of selectable marker genes to identify transgenic plants. Selectable marker genes are therefore required to ensure efficient genetic modification.

In many studies on transgenes, emphasis is given on the inheritance and expression of the transgene of interest while the selectable marker genes are largely ignored. One reason is that the expression of selectable marker genes is no longer necessary when successful transformants are obtained. The presence however and continued expression of selectable marker genes can potentially affect the expression of the transgene of interest. Their existence could conceivably activate gene silencing mechanisms with adverse effects on the expression of the transgene of interest. While some selectable marker genes would have no impact on plant growth and development in the absence of the selection agent, continued expression of some selectable markers can cause substantial metabolic burden and undermine expression levels of the transgene of interest or endogenous genes. Reduction in agronomic performance attributed to the presence of selectable marker genes has been reported (Hucl et al., 1998). For these reasons, the continued presence and expression of selectable marker
genes has been considered as undesirable and several strategies have been developed to eliminate them from the genome. Most transformation systems however are not yet at an efficiency level where the use of selectable markers are not required. Additionally, the proposed techniques to remove selectable marker genes from the genome still need to be refined. Until these techniques can be perfected, the inheritance and expression of selectable marker genes needs to be assessed in conjunction with the transgene of interest.

The synthetic porcine α-La transgene has been transformed into maize together with the gene BAR as a selectable marker as described in chapter 2. BAR encodes the enzyme phosphinotricin acetyltransferase (PAT). PAT converts phosphinotricin (PPT), the active ingredient of the herbicide bialaphos, into a non-herbicidal acetylated form by transferring the acetyl group from acetyl CoA to the free amino group of PPT, thus conferring PPT resistance on transgenic tissues and regenerated plants (D'Halluin et al., 1992). In chapter 2, the inheritance and expression patterns of the synthetic porcine α-La transgene was characterized. In this study, the objectives were to: 1) confirm the integration of the BAR transgene into the maize genome; 2) determine whether the BAR transgene is functional; 3) assess the stability of transgene integration; and 4) evaluate the inheritance and expression patterns of the BAR gene for three generations.

Materials and Methods

Gene constructs

A modified porcine alpha-lactalbumin (α-La) coding sequence (Yang et al., 2002) was used to produce 3 different constructs (Figure 4.1). In construct pUbi/z-sig/lactal/KDEL
(YPS 118), the synthetic α-La coding sequence is transcriptionally regulated by the maize Ubi-1 promoter and the nos 3' untranslated region. In this construct, the coding region encodes mature porcine α-lactalbumin and is translationally fused to the maize 27-kDa gamma-zein signal sequence at the N-terminus and an endoplasmic reticulum (ER) retention sequence (KDEL) at the C-terminus of the protein (Figure 4.1A). Transgenic maize calli and plants transformed by this construct were designated as P45. Construct pZein/z-sig/lactal/KDEL (YPS 112) is identical to the first construct except that the maize Ubi-1 promoter was replaced by the maize 27-kDa gamma-zein promoter (Figure 4.1B). Transgenic maize calli and plants transformed by this construct were designated as P64. Construct pZein/z-sig/lactal (YPS 117) is identical to the second construct except that the coding sequence does not encode the ER retention sequence (KDEL) at the C-terminus of the protein (Figure 4.1C). Transgenic maize calli and plants transformed by this construct were designated as P57. The selectable marker gene BAR, driven by the maize Ubi-1 promoter, is contained on the same plasmid as the α-La gene in pUbi/z-sig/lactal/KDEL. In pZein/z-sig/lactal/KDEL and pZein/z-sig/lactal, the BAR gene was on a separate plasmid (pBAR 184) and the two plasmids were co-bombarded (Figure 4.1 B and C).

Plant transformation and tissue culture

Plant transformation and tissue culture was done at the Plant Transformation Facility at Iowa State University as described (Frame et al., 2000). Highly embryogenic, Type II callus was initiated from embryos of the Hi II genotype (Armstrong et al., 1991). Established Type II embryogenic callus lines were bombarded with pUbi/zsig/lactal/KDEL (YPS 118) or co-bombarded with pZein/zsig/Lactal/KDEL (YPS112) and pBAR184 or pZein/zsig/Lactal
(YPS 117) and pBAR184 using particle bombardment-mediated transformation with a PDS-1000/HE Biolistic Particle Delivery System (Bio-Rad) following the manufacturer's recommendations. One week after bombardment, calli were moved to selective media containing Bialaphos. After 8 weeks, bialaphos resistant calli were moved to regeneration media supplemented with bialaphos. The plants obtained from regeneration were designated as T0 plants and were then transplanted into pots.

Development of segregating populations

The F1 generation was produced by matings between a T0 plant and inbred line B73 using the T0 plants as females. Five F1 kernels from each event were planted and the F1 plants obtained were backcrossed as males to B73 to obtain BC1F1 kernels. The F1 plants were also self-pollinated to obtain the F2 kernels. From each F2 ear, 40 to 65 F2 kernels from each event were planted and the resulting plants were self-pollinated to obtain F3 kernels. In each event, 2 to 20 F3 ears were obtained and from these, 2 to 5 F3 ears were selected. Seventeen F3 kernels from each selected F3 ear were planted and self-pollinated to obtain the F4 generation.

Analysis of herbicide resistance

Plants grown directly from α-La positive and negative sibling kernels in the F1, F2 and F3 generations were screened for herbicide resistance. The tip of the leaf was cut with scissors and about 6 inches of the leaf was dipped in herbicide solution (Liberty, 200 mg/l). The plants were scored for resistance to the herbicide after one week. A plant was scored resistant when the portion of the leaf dipped in the herbicide remained green and unburned. A plant was scored susceptible when the portion of the leaf dipped in the herbicide turned
brown and burnt (Figure 4.2). The frequency of herbicide resistant and susceptible plants was obtained to determine the phenotypic segregation ratios of the BAR gene. Deviation of these ratios from expected segregation ratios was determined by a chi-square test.

**PCR analysis**

F1 kernels that were positive and negative for the presence of the α-La protein in a Western blot analysis were planted in the field. Three weeks after germination, leaf tissue was collected from the F1 plants. DNA was extracted from the fresh leaf tissues using the Puregene DNA Isolation Kit (Gentra Systems) following the manufacturer’s recommendations. The DNA pellet was resuspended in 50 µl of 50 mM Tris, 10 mM EDTA, pH 8.0. PCR was carried out using 1 µl of isolated genomic DNA (~ 50 ng) in a 20 µl total reaction volume containing 100 µM of each dNTP, 2 µl of 10X PCR buffer (Gibco BRL), 1 unit Platinum Taq DNA Polymerase (Gibco BRL), 2.85 mM of MgCl2 and 0.2 µM of each primer. The primers used were BAR-Forward=GGTCAACTTCCGTACCGAGCCG and BAR-Reverse = CATCGCAAGACCGGCAACAGG. These primer sequences corresponded to sequences in the BAR coding region and amplified a fragment of about 536 bp (Figure 4.1 B and C). PCR conditions were 35 cycles of 30 seconds at 94 °C, 30 seconds at 58 °C and 2 minutes at 72 °C, with 5 minutes at 94 °C prior to the reaction and 5 minutes of final extension at 72 °C in a Rapidcycler (Idaho Technologies, Inc). PCR products were analyzed on 1.5 % (w/v) agarose gels stained with ethidium bromide and photographed.

**Southern blot analysis**

DNA was prepared from leaf tissue by the CTAB method (Saghai-Maroof et al., 1984). Genomic DNA was digested with enzymes that cut the plasmid once outside the α-La
transgene. P64 and P57 genomic DNA was digested with Sca I and P45 DNA was digested with Ssp I, and then fractionated by electrophoresis on an 0.8 % agarose gel. The α-La-containing plasmids used to produce P64 and P57 are about 4.5 kb while the plasmid used in P45 is about 8.2 kb in size (Figure 4.1). Transfer to nylon membranes (Hybond, Amersham) was performed as described (Southern, 1975). Approximately 5 µg of genomic DNA was used per digest. Hybridization was carried out using a 32P-labelled probe at 65 °C for 12-16 hrs. The α-La probe was obtained by amplifying the α-La coding region from the plasmid used to transform these plants, YPS 112, in a PCR (Figure 4.1 B). The amplified band was resolved in a low melting point agarose gel, cut out and then labeled with (α-32P) dCTP according to Feinberg and Vogelstein (1983). The blots obtained for α-La were washed to remove the labeled α-La probe. The blots were then reprobed with a BAR-containing restriction fragment. The BAR probe was obtained by a restriction digest of the plasmid YPS 118 using the enzyme Pst I. A 600 bp Pst I restriction fragment was isolated and labeled with (α-32P) dCTP according to Feinberg and Vogelstein (1983). This fragment corresponded to the BAR coding region (Figure 4.1 A).

Results and Discussion

Integration of the BAR gene into the genome and transmission and expression in the F1 generation

Three constructs that contain the synthetic porcine α-La and the selectable marker gene BAR were introduced into maize callus cells by the Iowa State University Plant Transformation Facility using their standard protocol for particle bombardment (Frame et al.,
2000). T0 plants were obtained from regeneration and crossed as females to the non-transformed inbred line B73 and F1 kernels were obtained from 8 different P64 T0 plants (i.e. 8 P64 events, 1 T0 plant per event), 7 different P45 T0 plants and 5 different P57 T0 plants as described in chapter 2 (Table 2.1). Even though the parent T0 plants were derived from callus grown in a media supplemented with bialaphos, PCR or Southern blot analysis was not performed in callus or in T0 plants to confirm the integration of the BAR transgene in the maize genome. It is important to confirm the integration of the BAR gene in the genome as well as its expression in the F1 generation because chimeric expression of transgenes including preferential regeneration of non-transformed cells from the chimeras has been reported to occur in transgenic maize (Fromm et al., 1990a).

To confirm the integration of the BAR gene into the maize genome, 5 F1 kernels from each event of P64, P45 and P57 constructs that were initially screened for the expression of the α-La protein in the endosperm were grown in the field and DNA was obtained from leaf tissues of these F1 plants for PCR analysis. Using primers corresponding to the BAR coding sequences, a PCR product (~500 bp) was detected in 6 of 7 P64 progenitor F1 plants, in 6 of 6 P45 progenitor F1 plants and in 4 of 4 P57 progenitor F1 plants, indicating the presence of BAR DNA sequences in these progenitor F1 plants. Additionally, BAR-hybridizing fragments were detected in these PCR-positive F1 plants in a Southern blot analysis (Table 4.1 and Figure 4.3B). Since the F1 plants were obtained from a cross between T0 plants and non-transformed B73 plants, the presence of BAR DNA sequences in these progenitor F1 plants as revealed in PCR and Southern blot analyses, suggests that the BAR transgene was indeed integrated into the maize genome and was transmitted from T0 to the F1 generation. In a leaf-dipping assay, the F1 plants that were
positive for BAR DNA sequences were resistant to the herbicide, indicating that the BAR gene was not only transmitted to the F1 generation but it was also functional. It is noteworthy that the herbicide resistant F1 plants came from kernels that were positive for α-La protein in the endosperm as summarized in Table 4.1. This may suggest that the α-La and BAR transgenes were not only co-transmitted from the T0 to the F1 generation but both transgenes were also expressed in the F1 generation.

It is the general observation that co-bombarded transgenes are integrated at linked sites in the genome (Pawlowski and Somers, 1996). It is noteworthy that in P64-27, where the α-La and BAR transgenes were co-bombarded, BAR DNA sequences were not detected in the P64-27 progenitor F1 plant while α-La DNA sequences were present as revealed in a PCR analysis. In a herbicide assay, the progenitor F1 plant was susceptible to the herbicide. In addition, Southern blot analysis showed α-La-hybridizing fragments but BAR-hybridizing fragments were not detected in this progenitor F1 plant (Table 4.1). Since the T0 plant was regenerated in media supplemented with bialaphos, it is highly likely that the BAR transgene was expressed in T0 but the BAR transgene may have been integrated at a different chromosome or at an unlinked site from the α-La transgene. It is possible that the BAR transgene was segregated away during transmission from T0 to the F1 generation resulting in the absence of BAR transgene DNA sequences in the selected progenitor F1 plant. This result shows that co-bombarded transgenes may not always integrate at linked sites in the genome. This result is not in agreement with the general finding that co-bombarded transgenes are integrated at linked sites in the genome (Pawlowski and Somers, 1996).
The same phenomenon can be used to explain the data for P64-23. In this event, α-La DNA sequences were not detected by PCR analysis in the progenitor F1 plant but the F1 plant was resistant to the herbicide. Furthermore, Southern blot analysis showed BAR-hybridizing fragments but no α-La-hybridizing fragments were detected in the progenitor F1 plant (Table 4.1). It is possible that in P64-23, the α-La transgene was integrated at a different chromosome or at an unlinked site from the BAR transgene and the α-La transgene may have been segregated away from BAR during transmission from T0 to the F1 generation resulting in the absence of α-La transgene DNA sequences in the progenitor F1 plant as detected by PCR and Southern blot analyses. This result further supports the observation that co-bombarded transgenes may not integrate at linked sites in the genome. This results do not support the general finding that co-bombarded transgenes are integrated at linked sites in the genome as reported in oats (Pawlowski and Somers, 1996).

*Stable integration and transmission of the BAR gene*

Southern blot analyses were performed on one F1 plant and its F2 and F3 progenies from each event in order to characterize BAR transgene integration and determine transgene stability through meiosis. The α-La probe was removed from the filters shown in Figure 4.3A and the filters were re-probed with a BAR-containing restriction fragment. As illustrated in Figure 4.3B, each progenitor F1 plant had a unique pattern of bands, indicating that each event had a unique integration pattern for the BAR transgene. The integration patterns in the F2 and F3 progenies were identical to their corresponding progenitor F1s, indicating that the BAR transgenes had not undergone new rearrangements in the production of F2 and F3 progenies, demonstrating a stable integration of the BAR gene.
Co-segregation and linkage of BAR and α-La transgenes

To determine if BAR and the porcine α-La transgenes were linked, F2 kernels and their corresponding plants were categorized into phenotypic classes relative to the presence of α-La protein in the endosperm and the reaction of the plants to herbicide treatment. The F2 kernels were obtained by selfing a herbicide resistant F1 plant derived from an α-La positive kernel while the F1 plant was obtained from a cross between a T0 and a non-transformed B73 inbred plant. As summarized in Table 4.2, 3 events of P64 (P64-14, P64-15, P64-1) contained F2 phenotypes that belong to parental classes (i.e. α-La positive kernel/herbicide resistant plant or α-La negative kernel/herbicide susceptible plant). No recombinant phenotypes were observed in these events, suggesting that the α-La and BAR transgenes co-segregated and were linked. Also, as shown in Figure 4.3 A and B, some α-La and BAR fragments in a Southern blot analysis of 2 P45 and 2 P64 events were of the same size (i.e 15 kb α-La and BAR fragments in P45-11 and P64-24). This may indicate that the α-La and BAR transgenes in these events were contained on the same restriction fragment in the genomic DNA.

In P64-18 and P64-24, most of the F2 kernels and plants displayed phenotypes that belong to the parental classes. However, a small number of recombinants was also detected in these events (Table 4.2). The occurrence of rare recombinants in these events may suggest that though α-La and BAR were integrated at linked sites, the linkage was not complete resulting to recombination between α-La and BAR loci. A similar pattern was observed among the four events of P45 (P45-3, P45-22, P45-16 and P45-11; Table 4.2). Most of the F2 phenotypes belong to the parental classes, but recombinants were also detected, suggesting
that the linkage between α-La and BAR loci in these events was incomplete. However, it can also be possible that the linkage between α-La and BAR is indeed complete in those events where recombinants were detected. As the analysis was based solely on the segregation of the transgene phenotypes (i.e. presence of α-La protein and herbicide resistance), it is possible that transgene DNA sequences may be present in these recombinants that were negative for the α-La protein and/or susceptible to the herbicide, indicating that the transgenes were present but were inactive.

About three-fourths (¾) of the progenies are expected to express the transgene phenotype when a single dominant locus is in independent assortment in the F2. In 5 P64 and 4 P45 events evaluated, the percentage of expression of both the α-La and BAR transgenes was close to 75% (Table 4.2), indicating that both linked transgenes were inherited as a single genetic locus. This result agrees well with the general finding that transgenes introduced by particle bombardment are integrated at linked sites in the genome and are inherited as single genetic locus (Pawlowski and Somers, 1996).

Inheritance and expression patterns of BAR in the F2 and F3 generations

The inheritance and expression patterns of the selectable gene BAR was assessed for three generations. As summarized in Table 4.3, in 5 events of P64 (P64-18, P64-14, P64-15, P64-1, P64-24), the proportion of plants that were herbicide resistant and susceptible conformed to a 3:1 phenotypic ratio in the F2 and F3 generations, indicating that the BAR transgene was expressed and inherited as a single locus in these events. It is noteworthy that in 4 of these 5 events, the α-La transgene was also expressed and inherited as a single locus in the F2 and F3 generations, indicating that both the α-La and BAR transgenes were
consistently expressed in these events as single locus for three successive generations. No BAR DNA sequences were observed in P64-27 and no α-La DNA sequences were observed in P64-23 as detected by PCR and Southern blot analysis of progenitor F1 plants (Table 4.1). Consequently, no α-La expressing kernels in P64-23 nor herbicide resistant plants in P64-27 were detected in the F2 and F3 generations.

For P45, three events (P45-3, P45-22, P45-16) showed a phenotypic segregation ratio in the F2 that conformed to a single locus inheritance pattern for both for α-La and BAR genes indicating both the α-La and BAR transgenes were expressed and inherited as single locus in these events (Table 4.3). In the F3 generation however, P45-3 and P45-22 displayed a phenotypic ratio for the α-La transgene that deviated from the single locus ratio due to the prevalence of kernels with no detectable α-La protein as a result of transgene inactivation as described in chapter 2. Some of these α-La null kernels when planted grew into plants that were herbicide resistant and the phenotypic ratio of resistant to susceptible plants was not different from a single locus inheritance of the BAR gene. This results show that unlike α-La, the expression of BAR in these events was consistent, suggesting that the mechanisms that contributed to the inactivation of the α-La transgene in these events did not inactivate BAR. Examples of transgenes that were inactivated or silenced when inserted into a particular location of the genome, and did not affect the expression of any other element or transgenes have been reported (Assaad et al., 1993; Kumpatla et al., 1997a).

In three other events of P45 (P45-11, P45-4, P45-19), the segregation ratio for herbicide resistance did not conform to the 3:1 phenotypic ratio in the F2 and F3 generations. The inheritance also of α-La in these events in the F2 and F3 generation did not follow the
3:1 phenotypic ratio. In P45-4, the \( \alpha \)-La transgene was inactivated in the F2 and F3 generation as described in chapter 2. The F2 and F3 plants derived from these \( \alpha \)-La null kernels were herbicide susceptible (Table 4.3). PCR analysis showed that BAR DNA sequences were present in these F2 and F3 susceptible plants, indicating that the BAR transgene was also inactivated in this event. So in P45-4, both the \( \alpha \)-La and BAR transgenes were inactivated in the F2 and F3 generations. It is possible that the mechanisms that contributed to the inactivation \( \alpha \)-La in these events might have also inactivated BAR or vice versa. One possible scenario is that in these events, the \( \alpha \)-La transgene was inactivated and targeted BAR, imposing its inactivated state into BAR thereby inactivating also the BAR transgene. The presence of homologous promoter sequences in BAR might have provided the target for the inactivation of the BAR transgene. The reverse scenario is equally plausible, with the inactivated BAR transgene targeting and imposing its inactivated state on the \( \alpha \)-La transgene. It is for this reason why in breeding systems that utilize transgenes, single transgene copies are preferred and the continued presence of selectable markers is considered undesirable. The results in this study reinforces the notion that selectable marker genes can contribute to the expression patterns of the transgene of interest or to endogenous genes and should therefore be given importance when assessing the expression and inheritance of transgenes.

Of all the events analyzed, 3 of 6 events in P45 displayed inactivation of the BAR transgene both in the F2 and F3 generations. In contrast, all the 7 events of P64 displayed consistent expression of BAR in F2 and F3 generations. In addition, P57-35 and P57-24 displayed segregation ratios for herbicide resistance in the F2 and F3 generations that
conformed to the single locus inheritance pattern demonstrating a consistent expression of BAR in these events while the α-La transgene was inactivated. The higher number of P45 events that displayed BAR inactivation (and also α-La inactivation as described in chapter 2) relative to P64 and P57 events may have been related to the presence of homologous promoter sequences in P45. In P45, both α-La and BAR transgenes are driven by the ubiquitin promoter and it is likely that this homology in the promoter region might have provided a target for the inactivation of either the α-La or BAR genes in P45. Presence of homologous sequences in promoter regions has been reported as one of the mechanisms that contribute to inactivation of transgenes (Finnegan and McElroy, 1994). Homologous sequences can lead to methylation of promoter regions blocking transcription (Davies et al., 1997; Kumpatla et al., 1997).

It was shown in this study that it is possible that co-bombarded transgenes may not integrate at linked sites in the genome. Although it is the general finding that co-bombarded transgenes are integrated at linked sites in the genome as shown in oats (Pawlowski and Somers, 1996), partial linkage can also result to a recombination between a selectable marker and the transgene of interest. A possible consequence of these observations is that in some breeding systems, the expression of linked selectable markers such as herbicide resistance is used to indicate the presence of the transgenes of interest during the transfer of transgene locus to elite lines. This procedure however may not be always reliable as recombination may result in some individuals that express the selectable marker but do not carry the transgenes of interest. For such a procedure to be effective requires selection of events where the selectable marker and the transgenes of interest are completely linked. In addition, the
occurrence of transgene inactivation in later generations may require constant monitoring to determine consistent expression of both genes.
References


Table 4.1. Summary of the genotypes and phenotypes of different progenitor F1s for the α-La and BAR transgenes in P64, P45 and P57 constructs.

<table>
<thead>
<tr>
<th>Event</th>
<th>PCR α-La</th>
<th>Western blot α-La</th>
<th>PCR Bar</th>
<th>Herb. assay</th>
<th>Southern blot α-La</th>
<th>Southern blot Bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P64-14-4)-3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>resistant</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P64-18-5)-5</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>resistant</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P64-15-6)-4</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>resistant</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P64-1-4)-1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>resistant</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P64-27-5)-2</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>susceptible</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>(P64-23-1)-5</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>resistant</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>(P64-24-2)-1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>resistant</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P45-3-3)-4</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>resistant</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P45-22-1)-1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>resistant</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P45-16-2)-4</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>resistant</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P45-11-3)-1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>resistant</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P45-4-5)-3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>resistant</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P45-19-7)-1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>susceptible</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P57-35-8)-1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>resistant</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P57-24-10)-1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>resistant</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P57-25-2)-1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>susceptible</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(P57-27-3)-1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>resistant</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

PCR α-La = plus and minus signs indicate presence (+) and absence (-) of α-La DNA sequences as detected by PCR analysis. Western α-La = plus and minus signs indicate presence (+) and absence (-) of α-La protein as detected by Western blot analysis. PCR BAR = plus and minus signs indicate presence (+) and absence (-) of BAR DNA sequences as detected by PCR analysis. Plus and minus signs in a Southern blot indicate the presence (+) or absence (-) of α-La or BAR hybridizing sequences.
Table 4.2. Segregation of α-La and herbicide resistance among events in P64, P45 and P57 in the F2 generation.

<table>
<thead>
<tr>
<th>Event</th>
<th>α-La/Res (+/+</th>
<th>α-La/Res (+/-)</th>
<th>α-La/Res (-/+</th>
<th>α-La/Res (-/-)</th>
<th>% expression of α-La and BAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>64-18-5</td>
<td>17</td>
<td>2</td>
<td>2</td>
<td>11</td>
<td>53.1</td>
</tr>
<tr>
<td>64-14-4</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>84.6</td>
</tr>
<tr>
<td>64-15-6</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>63.2</td>
</tr>
<tr>
<td>64-1-4</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>64.3</td>
</tr>
<tr>
<td>64-24-4</td>
<td>17</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>65.4</td>
</tr>
<tr>
<td>45-3-3</td>
<td>19</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>76.0</td>
</tr>
<tr>
<td>45-22-1</td>
<td>28</td>
<td>1</td>
<td>0</td>
<td>14</td>
<td>65.1</td>
</tr>
<tr>
<td>45-16-2</td>
<td>25</td>
<td>0</td>
<td>1</td>
<td>9</td>
<td>71.4</td>
</tr>
<tr>
<td>45-11-3</td>
<td>19</td>
<td>1</td>
<td>1</td>
<td>15</td>
<td>52.8</td>
</tr>
<tr>
<td>57-35-8</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>12</td>
<td>0.0</td>
</tr>
<tr>
<td>57-24-10</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>11</td>
<td>0.0</td>
</tr>
</tbody>
</table>

(+/+) = positive for presence of the porcine α-La protein in the endosperm and the plant is herbicide resistant; (+/-) = positive for presence of the porcine α-La protein in the endosperm but the plant is herbicide susceptible; (-/+ ) = no detectable porcine α-La protein in the endosperm but the plant is herbicide resistant; (-/-) = no detectable porcine α-La protein in the endosperm and the plant is herbicide susceptible.
Table 4.3 Inheritance and expression of α-La and herbicide resistance among events in P64, P45 and P57 constructs in F2 and F3 generations.

<table>
<thead>
<tr>
<th>Events</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>α-La (+/-)</td>
<td>Herb. Res. (+/-)</td>
</tr>
<tr>
<td>64-18-5</td>
<td>46/14 ns</td>
<td>19/13 ns</td>
</tr>
<tr>
<td>64-14-4</td>
<td>49/18 ns</td>
<td>22/4 ns</td>
</tr>
<tr>
<td>64-15-6</td>
<td>34/12 ns</td>
<td>12/8 ns</td>
</tr>
<tr>
<td>64-1-4</td>
<td>50/16 ns</td>
<td>9/5 ns</td>
</tr>
<tr>
<td>64-24-4</td>
<td>39/27 **</td>
<td>18/8 ns</td>
</tr>
<tr>
<td>64-27-5</td>
<td>23/43 **</td>
<td>0/23 **</td>
</tr>
<tr>
<td>64-23-1</td>
<td>0/60 **</td>
<td>34/6 ns</td>
</tr>
<tr>
<td>45-3-3</td>
<td>34/10 ns</td>
<td>21/4 ns</td>
</tr>
<tr>
<td>45-22-1</td>
<td>43/20 ns</td>
<td>28/15 ns</td>
</tr>
<tr>
<td>45-16-2</td>
<td>53/11 ns</td>
<td>26/8 ns</td>
</tr>
<tr>
<td>45-11-3</td>
<td>33/28 **</td>
<td>20/16 **</td>
</tr>
<tr>
<td>45-4-5</td>
<td>0/44 **</td>
<td>0/29 **</td>
</tr>
<tr>
<td>45-19-7</td>
<td>0/44 **</td>
<td>2/37 **</td>
</tr>
<tr>
<td>57-35-8</td>
<td>0/40 **</td>
<td>30/12 ns</td>
</tr>
<tr>
<td>57-24-10</td>
<td>0/61 **</td>
<td>14/11 ns</td>
</tr>
</tbody>
</table>

ns = not significantly different from 3:1 phenotypic ratio for a single dominant locus model in the F2 and F3; ** significantly different; chi square (0.01, 1df) = 6.64; H = homozygous for α-La protein expression. For the α-La transgene, (+) indicates presence of α-La protein and (-) absence of α-La protein as detected in a Western blot analysis. For the BAR transgene, (+) indicates the plant is herbicide resistant and (-) the plant is susceptible in a herbicide leaf-dipping assay.
Figure 4.1. Schematic representation of plasmids A. YPS 118; B. YPS 112 and C. YPS 117 used in transformation. Calli transformed by these plasmids and the T0 plants obtained were designated as P45, P64 and P57, respectively. For Southern blot analysis, the α-La probe was obtained by PCR amplification of the α-La coding region in YPS 112. Arrows indicate the position of primer binding sites. The BAR probe was obtained by a *Pst* I digest of YPS 118. The position and size of the α-La and BAR probes used in Southern blot analysis are indicated by the dotted black box below the coding region of the genes of interest. Genomic DNA was digested with enzymes that cuts once in the plasmid outside the transgene cassette. P45 genomic DNA was digested with *Ssp* I, P64 and P57 with *Sea* I. For pBAR184, arrows indicate the position of the primers used in a PCR analysis to detect presence of BAR transgene DNA sequences in transformed plants. Bold boxes – zein signal sequence; striped boxes – ER retention signal.
Figure 4.2 Analysis of herbicide resistance in maize plants transformed with a synthetic porcine α-La transgene and the selectable marker gene BAR. The tip of the leaf was cut with scissors and about 6 inches of the leaf was dipped in herbicide solution (Liberty, 200mg/l). After one week, a plant is scored susceptible if the portion of the leaf dipped in the herbicide turned brown and burnt (left) and resistant if the leaf remained green (right).
Figure 4.3. Southern blot analysis of porcine α-La and BAR transgenes in progenitor F1s and the corresponding F2 and F3 progenies in P45 and P64 constructs. P1 = plasmid used to transform P45; P2 = plasmid used to transform P64 that contain the α-La transgene as positive controls. It contained 5 pg of plasmid DNA representing 1 copy per diploid genome. B = untransformed DNA from B73 as negative control. P45-11, P45-16, P64-24 and P64-15 are independent events. In each event, lanes 1 contain DNA from the progenitor F1 plant; lanes 2 contain DNA from an F2 plant; lanes 3 and 4 contain DNA from 2 F3 sibling progenies. M = molecular weight standards in kilobases (kb). A. Genomic DNA was digested with Sea I for P64, Ssp I for P45 and probed with α-La coding sequence. B. The blot shown in A was washed to remove the α-La probe and reprobed with BAR. No BAR hybridizing fragment was observed in P2 since P64 was co-bombarded with pBAR184.
General Conclusions

In an attempt to improve the nutritional quality of the maize grain, a codon adjusted version of a gene encoding the milk protein porcine α-lactalbumin (α-La) was synthesized and different versions of this gene were constructed. Three constructs that differed in their promoters and in their subcellular targeting domains, namely; pUbi/zsig/lac/kdel, pZein/zsig/lac/kdel and pZein/zsig/lac, were introduced into maize cells by particle bombardment and the plants obtained from regeneration were called P45, P64 and P57, respectively. After transformation, the integration of the transgenes into the genome and their expression and inheritance patterns were not known. It was important to characterize the expression and inheritance of α-La transgenes as transgenes often display unpredictable patterns of inheritance and expression. Transgenes can only be of practical use when they are inherited and expressed in a predictable manner.

In Chapter 2, it was established by PCR and Southern blot analyses that in six events of P64, six events of P45 and four events of P57, the synthetic porcine α-La transgene was integrated successfully into the maize genome and was transmitted to the next generation. In each event, the transgene integration pattern in the F1 generation was unique and exhibited rearrangements indicated by the appearance of additional hybridizing restriction fragments of altered molecular weights. The transgene integration patterns were stable as the hybridization patterns remained identical in succeeding F2 and F3 generations. It was also established by Western blot that not only were the transgenes transmitted to the next generation but they were also functional since protein bands of similar molecular weight to human α-
Lactalbumin were detected in the endosperm of transformed kernels indicating that the porcine α-La accumulated in the endosperm.

The inheritance and expression patterns of the synthetic porcine α-La transgene was assessed for three generations. The general pattern observed was that in events of P64 and P45, the α-La transgene was inherited as single locus. In P64, the expression of the α-La protein was consistent for 3 generations. However, transgene inactivation was also observed as shown by the absence of α-La protein in kernels that contained the α-La transgene. In P64, 5 in 6 events (83%) showed consistent expression of α-La. In contrast, 5 in 6 events (83%) of P45 displayed transgene inactivation. Inactivation was detected as early as in the BC1F1 and this persisted in the F3 generation. The explanation put forward for the higher frequency of gene inactivation in P45 when compared to P64 was the presence of homologous promoter sequences in the same vector in P45. In P45, the α-La transgene and the selectable marker gene BAR were on the same plasmid and both were under the transcriptional control of the ubiquitin promoter. This homology in the promoter sequences might have provided a target for the inactivation of the α-La transgenes. Homologous sequences can lead to methylation of the promoter region blocking transcription. It has been shown in other studies that the presence of homologous or identical promoter sequences on the same vector can lead to transgene inactivation. This observation have important practical implications in the future designs of α-La constructs as repeated elements in a construct need to be minimized if not avoided to maximize transgene expression. The transgene inactivation was not stable as α-La expression was observed in some F3 lines derived from self-pollination of inactivated F2 plants suggesting putative reactivation of α-La expression. In
P57, the α-La protein was detected in the F1 but not in the F2 and F3 generations. It was difficult to ascertain if a gene silencing mechanism was involved in the non-detection of the α-La protein among events in P57 as P57 lacks the ER retention signal KDEL contained in P64. P64 contains KDEL and consistent expression of α-La protein was detected among events in P64. It was postulated that the non-detection of α-La protein in P57 is related to the absence of KDEL in this construct suggesting that post-translational factors may be involved in the accumulation of porcine α-La in maize cells.

Having established that the synthetic porcine α-La transgenes were integrated into the genome, inherited in the next generation and functional, studies were conducted to assess the activity and tissue specificity of the α-lactalbumin expression under the control of the zein and ubiquitin promoters. This was important in order to determine the suitable promoters to drive α-La transcription and obtain the required level and pattern of expression. Studies were also conducted to assess the levels of the porcine α-La protein and determine whether the expression of porcine α-La in the endosperm can improve the amino acid content. It was established in chapter 3 that the accumulation of the modified porcine α-La protein was restricted in the endosperm while differential tissue specificity of α-La was detected for events driven by the ubiquitin promoter. This result demonstrated that the zein promoter is sufficient to direct an endosperm-specific expression of α-La while the ubiquitin promoter may not be suitable to direct transcription of the α-La in the whole plant.

The levels of α-La protein that accumulated in the endosperm vary among events in a construct. It was postulated that this event to event variation in the levels of porcine α-La protein was due to the differences in the site of transgene integration within the genome or to
transgene copy number. The mean levels of porcine α-La that accumulated in the endosperm ranged from 0.0014 g/100g – 0.095 g/100g of endosperm tissue. Importantly, the expression of porcine α-La in the kernel endosperm resulted in an increase in Lysine content. Kernels positive for α-La protein had a 29% - 49% difference in Lysine when compared to their negative sibling counterparts. A positive correlation \((r = 0.16)\) was also detected between levels of α-La in the endosperm and Lysine content. The difference in Lysine levels was construct and event specific, as most events that showed differences in Lysine levels were from P64 and the Lysine levels varied among the events of this construct.

The difference in Lysine levels was also accompanied by significant changes in the content of other amino acids including a significant difference in the levels of Aspartic Acid and Methionine, among others. Expression of α-La usually did not lead to significant changes in Nitrogen, Carbon and total protein content. The results of this study showed that plants expressing the modified porcine α-La protein were able to increase their supply of Lysine for protein synthesis, demonstrating that the use of a gene encoding a milk protein can indeed be useful strategy to improve maize grain quality. To our knowledge, this is the first report on an improvement in Lysine levels in maize by a transgenic approach using a synthetic gene encoding a modified milk protein.

It was important to characterize the integration as well as the inheritance and expression patterns of the selectable marker gene BAR. In many studies on transgene inheritance and expression, the selectable marker genes are mostly ignored despite their potential to contribute to the inheritance and expression patterns of the transgene of interest. It was established in chapter 4 that the integration patterns of BAR was unique for each
event. The BAR integration patterns were stable as the hybridization patterns remained identical in the F1, F2 and F3 generations. In most cases, the BAR transgene was co-integrated with the α-La transgene into the maize genome as the different events of P64, P45 and P57 constructs that contained BAR-hybridizing sequences also contained α-La hybridizing sequences. Phenotypic segregation analysis in the F2 generation showed that BAR was linked with the α-La transgene. Analysis of the inheritance and expression patterns of BAR showed that in five of six P64 events, both BAR and α-La transgenes were consistently expressed and inherited as single locus in the F2 and F3 generations. In P45, variable expression patterns for BAR were observed. In three of six events, both BAR and α-La transgenes were inactivated. It was postulated that a similar mechanism that contributed to the inactivation of α-La among events in P45 contributed to the variable expression patterns for BAR or vice versa. In P57, BAR was consistently expressed and inherited as a single locus in the F2 and F3 generations while the α-La transgene was inactivated. The results of this study supported the notion that expression of selectable marker genes can contribute to the inheritance and expression patterns of the transgenes of interest.
Appendix

Table 1. Amino acid composition and total protein content of maize kernels positive and negative for porcine α-La protein in the endosperm in event P64-14 as measured in a complete amino acid analysis assay.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Positive (Mean ± SE)</th>
<th>Negative (Mean ± SE)</th>
<th>% Difference</th>
<th>Positive (Mean ± SE)</th>
<th>Negative (Mean ± SE)</th>
<th>% Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspartic Acid</td>
<td>0.75 ± 0.023</td>
<td>0.70 ± 0.046</td>
<td>7.11 ns</td>
<td>0.73 ± 0.015</td>
<td>0.65 ± 0.003</td>
<td>11.73*</td>
</tr>
<tr>
<td>Threonine</td>
<td>0.35 ± 0.005</td>
<td>0.33 ± 0.005</td>
<td>6.06*</td>
<td>0.36 ± 0.005</td>
<td>0.35 ± 0.003</td>
<td>1.89 ns</td>
</tr>
<tr>
<td>Serine</td>
<td>0.42 ± 0.012</td>
<td>0.39 ± 0.014</td>
<td>5.93 ns</td>
<td>0.44 ± 0.003</td>
<td>0.47 ± 0.008</td>
<td>-6.43*</td>
</tr>
<tr>
<td>Glutamic Acid</td>
<td>2.34 ± 0.000</td>
<td>2.36 ± 0.059</td>
<td>-0.99 ns</td>
<td>2.58 ± 0.020</td>
<td>2.58 ± 0.020</td>
<td>0.00 ns</td>
</tr>
<tr>
<td>Proline</td>
<td>1.04 ± 0.006</td>
<td>1.08 ± 0.029</td>
<td>-4.31 ns</td>
<td>1.14 ± 0.018</td>
<td>1.21 ± 0.008</td>
<td>-6.32*</td>
</tr>
<tr>
<td>Glycine</td>
<td>0.26 ± 0.006</td>
<td>0.26 ± 0.012</td>
<td>0.00 ns</td>
<td>0.25 ± 0.006</td>
<td>0.25 ± 0.000</td>
<td>1.33 ns</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.87 ± 0.003</td>
<td>0.88 ± 0.023</td>
<td>-0.76 ns</td>
<td>0.94 ± 0.008</td>
<td>0.95 ± 0.006</td>
<td>-1.06 ns</td>
</tr>
<tr>
<td>Cysteine</td>
<td>0.25 ± 0.000</td>
<td>0.23 ± 0.006</td>
<td>7.14*</td>
<td>0.26 ± 0.000</td>
<td>0.25 ± 0.003</td>
<td>2.63 ns</td>
</tr>
<tr>
<td>Valine</td>
<td>0.54 ± 0.012</td>
<td>0.52 ± 0.010</td>
<td>2.21 ns</td>
<td>0.53 ± 0.013</td>
<td>0.52 ± 0.006</td>
<td>1.94 ns</td>
</tr>
<tr>
<td>Methionine</td>
<td>0.31 ± 0.008</td>
<td>0.29 ± 0.013</td>
<td>9.30 ns</td>
<td>0.30 ± 0.003</td>
<td>0.31 ± 0.006</td>
<td>-1.09 ns</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>0.45 ± 0.006</td>
<td>0.41 ± 0.012</td>
<td>8.06*</td>
<td>0.46 ± 0.005</td>
<td>0.39 ± 0.012</td>
<td>16.95*</td>
</tr>
<tr>
<td>Leucine</td>
<td>1.65 ± 0.008</td>
<td>1.66 ± 0.052</td>
<td>-0.40 ns</td>
<td>1.82 ± 0.025</td>
<td>1.83 ± 0.020</td>
<td>-0.36 ns</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>0.38 ± 0.008</td>
<td>0.39 ± 0.005</td>
<td>-1.71 ns</td>
<td>0.38 ± 0.012</td>
<td>0.39 ± 0.008</td>
<td>-2.54 ns</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>0.63 ± 0.003</td>
<td>0.62 ± 0.023</td>
<td>1.08 ns</td>
<td>0.67 ± 0.010</td>
<td>0.66 ± 0.008</td>
<td>2.03 ns</td>
</tr>
<tr>
<td>Histidine</td>
<td>0.28 ± 0.006</td>
<td>0.28 ± 0.008</td>
<td>0.00 ns</td>
<td>0.29 ± 0.005</td>
<td>0.29 ± 0.003</td>
<td>1.16 ns</td>
</tr>
<tr>
<td>Lysine</td>
<td>0.27 ± 0.014</td>
<td>0.20 ± 0.005</td>
<td>33.33*</td>
<td>0.22 ± 0.008</td>
<td>0.17 ± 0.003</td>
<td>28.85*</td>
</tr>
<tr>
<td>Arginine</td>
<td>0.38 ± 0.008</td>
<td>0.37 ± 0.012</td>
<td>2.68 ns</td>
<td>0.37 ± 0.008</td>
<td>0.36 ± 0.000</td>
<td>1.85 ns</td>
</tr>
</tbody>
</table>

Crude Protein | 11.53 ± 0.156 | 11.27 ± 0.181 | -1.06 ns | 11.90 ± 0.054 | 11.79 ± 0.139 | 0.93 ns |

*Mean values are from three sibling kernels from each positive and negative class ± standard error.
* = difference between positive and negative values is significant in a t-test at α = 0.05; ns = not significant.
Table 2. Amino acid composition and total protein content of maize kernels positive and negative for porcine α-La protein in the endosperm in event P64-18 as measured in a complete amino acid analysis assay.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>P64-18 F2 positive Mean ± SE</th>
<th>P64-18 F2 negative Mean ± SE</th>
<th>% positive difference</th>
<th>P64-18-F3 positive Mean ± SE</th>
<th>P64-18-F3 negative Mean ± SE</th>
<th>% negative difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspartic Acid</td>
<td>0.76 ± 0.054</td>
<td>0.61 ± 0.020</td>
<td>24.73 *</td>
<td>0.82 ± 0.023</td>
<td>0.72 ± 0.020</td>
<td>13.95 *</td>
</tr>
<tr>
<td>Threonine</td>
<td>0.37 ± 0.020</td>
<td>0.32 ± 0.008</td>
<td>17.89 *</td>
<td>0.39 ± 0.011</td>
<td>0.37 ± 0.000</td>
<td>5.41 ns</td>
</tr>
<tr>
<td>Serine</td>
<td>0.49 ± 0.020</td>
<td>0.40 ± 0.014</td>
<td>21.49 *</td>
<td>0.53 ± 0.017</td>
<td>0.50 ± 0.017</td>
<td>6.67 ns</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>2.52 ± 0.040</td>
<td>2.27 ± 0.029</td>
<td>11.18 *</td>
<td>2.72 ± 0.047</td>
<td>2.80 ± 0.075</td>
<td>-2.86 ns</td>
</tr>
<tr>
<td>Proline</td>
<td>1.05 ± 0.031</td>
<td>1.07 ± 0.014</td>
<td>-1.86 ns</td>
<td>1.15 ± 0.027</td>
<td>1.25 ± 0.039</td>
<td>-7.98 ns</td>
</tr>
<tr>
<td>Glycine</td>
<td>0.27 ± 0.020</td>
<td>0.23 ± 0.003</td>
<td>19.12 *</td>
<td>0.26 ± 0.014</td>
<td>0.25 ± 0.005</td>
<td>2.67 ns</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.96 ± 0.026</td>
<td>0.85 ± 0.013</td>
<td>12.99 *</td>
<td>1.02 ± 0.016</td>
<td>1.03 ± 0.029</td>
<td>-0.97 ns</td>
</tr>
<tr>
<td>Cysteine</td>
<td>0.24 ± 0.010</td>
<td>0.23 ± 0.003</td>
<td>5.88 ns</td>
<td>0.22 ± 0.008</td>
<td>0.23 ± 0.003</td>
<td>-1.47 ns</td>
</tr>
<tr>
<td>Valine</td>
<td>0.52 ± 0.018</td>
<td>0.49 ± 0.011</td>
<td>6.80 ns</td>
<td>0.51 ± 0.011</td>
<td>0.55 ± 0.023</td>
<td>-6.71 ns</td>
</tr>
<tr>
<td>Methionine</td>
<td>0.34 ± 0.015</td>
<td>0.28 ± 0.005</td>
<td>21.43 *</td>
<td>0.27 ± 0.012</td>
<td>0.25 ± 0.005</td>
<td>6.67 ns</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>0.44 ± 0.014</td>
<td>0.38 ± 0.008</td>
<td>15.93 *</td>
<td>0.45 ± 0.003</td>
<td>0.47 ± 0.020</td>
<td>-3.55 ns</td>
</tr>
<tr>
<td>Leucine</td>
<td>1.82 ± 0.030</td>
<td>1.62 ± 0.020</td>
<td>13.35 *</td>
<td>2.05 ± 0.043</td>
<td>2.06 ± 0.079</td>
<td>-0.49 ns</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>0.44 ± 0.020</td>
<td>0.36 ± 0.010</td>
<td>22.22 *</td>
<td>0.44 ± 0.013</td>
<td>0.40 ± 0.005</td>
<td>10.83 *</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>0.68 ± 0.023</td>
<td>0.59 ± 0.005</td>
<td>15.25 *</td>
<td>0.74 ± 0.010</td>
<td>0.75 ± 0.023</td>
<td>-0.89 ns</td>
</tr>
<tr>
<td>Histidine</td>
<td>0.25 ± 0.008</td>
<td>0.26 ± 0.006</td>
<td>-1.30 ns</td>
<td>0.25 ± 0.005</td>
<td>0.28 ± 0.012</td>
<td>-9.64 ns</td>
</tr>
<tr>
<td>Lysine</td>
<td>0.25 ± 0.040</td>
<td>0.17 ± 0.005</td>
<td>49.02 *</td>
<td>0.24 ± 0.016</td>
<td>0.18 ± 0.008</td>
<td>33.96 *</td>
</tr>
<tr>
<td>Arginine</td>
<td>0.40 ± 0.028</td>
<td>0.35 ± 0.008</td>
<td>15.38 ns</td>
<td>0.36 ± 0.015</td>
<td>0.36 ± 0.008</td>
<td>0.93 ns</td>
</tr>
</tbody>
</table>

Crude Protein  11.55 ± 0.470   10.47 ± 0.046   10.31 *   12.63 ± 0.375   12.15 ± 0.276   3.98 ns

* Mean values are from three sibling kernels from each positive and negative class ± standard error.

* = difference between positive and negative values is significant in a t-test at α = 0.05; ns = not significant.
Table 3. Amino acid composition and total protein content of maize kernels positive and negative for porcine α-La protein in the endosperm in events P45-11 and P45-22 as measured in a complete amino acid analysis assay.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>P45-11 F2 Mean ± SE a</th>
<th>% difference</th>
<th>P45-22 F2 Mean ± SE a</th>
<th>% difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspartic Acid</td>
<td>0.51 ± 0.015</td>
<td>0.64 ± 0.006</td>
<td>-20.73 *</td>
<td>0.62 ± 0.011</td>
</tr>
<tr>
<td>Threonine</td>
<td>0.27 ± 0.006</td>
<td>0.35 ± 0.003</td>
<td>-22.64 *</td>
<td>0.34 ± 0.005</td>
</tr>
<tr>
<td>Serine</td>
<td>0.36 ± 0.006</td>
<td>0.46 ± 0.014</td>
<td>-21.58 *</td>
<td>0.46 ± 0.008</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>1.88 ± 0.041</td>
<td>2.51 ± 0.023</td>
<td>-25.10 *</td>
<td>2.40 ± 0.038</td>
</tr>
<tr>
<td>Proline</td>
<td>0.94 ± 0.020</td>
<td>1.17 ± 0.014</td>
<td>-19.43 *</td>
<td>1.17 ± 0.025</td>
</tr>
<tr>
<td>Glycine</td>
<td>0.20 ± 0.005</td>
<td>0.25 ± 0.010</td>
<td>-20.00 *</td>
<td>0.23 ± 0.003</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.70 ± 0.018</td>
<td>0.92 ± 0.012</td>
<td>-24.55 *</td>
<td>0.89 ± 0.025</td>
</tr>
<tr>
<td>Cysteine</td>
<td>0.20 ± 0.005</td>
<td>0.24 ± 0.005</td>
<td>-16.67 *</td>
<td>0.24 ± 0.005</td>
</tr>
<tr>
<td>Valine</td>
<td>0.41 ± 0.008</td>
<td>0.52 ± 0.012</td>
<td>-21.02 *</td>
<td>0.51 ± 0.008</td>
</tr>
<tr>
<td>Methionine</td>
<td>0.23 ± 0.005</td>
<td>0.28 ± 0.006</td>
<td>-16.87 *</td>
<td>0.31 ± 0.005</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>0.33 ± 0.008</td>
<td>0.42 ± 0.003</td>
<td>-21.60 *</td>
<td>0.40 ± 0.01</td>
</tr>
<tr>
<td>Leucine</td>
<td>1.34 ± 0.040</td>
<td>1.74 ± 0.017</td>
<td>-23.14 *</td>
<td>1.71 ± 0.048</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>0.30 ± 0.003</td>
<td>0.41 ± 0.008</td>
<td>-28.23 *</td>
<td>0.39 ± 0.005</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>0.50 ± 0.013</td>
<td>0.65 ± 0.011</td>
<td>-23.59 *</td>
<td>0.63 ± 0.015</td>
</tr>
<tr>
<td>Histidine</td>
<td>0.23 ± 0.005</td>
<td>0.28 ± 0.006</td>
<td>-18.82 *</td>
<td>0.28 ± 0.003</td>
</tr>
<tr>
<td>Lysine</td>
<td>0.15 ± 0.003</td>
<td>0.18 ± 0.008</td>
<td>-16.36 *</td>
<td>0.17 ± 0.006</td>
</tr>
<tr>
<td>Arginine</td>
<td>0.30 ± 0.008</td>
<td>0.38 ± 0.012</td>
<td>-21.24 *</td>
<td>0.35 ± 0.006</td>
</tr>
<tr>
<td>Crude Protein</td>
<td>8.8 ± 0.256</td>
<td>11.32 ± 0.397</td>
<td>-22.28 *</td>
<td>11.19 ± 0.089</td>
</tr>
</tbody>
</table>

a Mean values are from three sibling kernels from each positive and negative class ± standard error.

* = difference between positive and negative values is significant in a t-test at α = 0.05; ns = not significant.