2007-07-12

The Prolamins of Pearl Millet

Christian B. Ricks
Brigham Young University - Provo

THE PROLAMINS OF PEARL MILLET

by

Christian B. Ricks

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the degree requirements for

Master of Science

Department of Plant and Wildlife Sciences

Brigham Young University

August 2007

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Christian B. Ricks

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

_____	_____
Date	Craig E. Coleman, Chair
_____	_____
Date	Joshua A. Udall
_____	_____
Date	Steven R. Herron

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Christian B. Ricks its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and departmental style requirements; (2) its illustrative material including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

_____ _____

Date

Craig E. Coleman
Chair, Graduate Committee

Accepted for the Department

_____

Loreen Woolstenhulme
Graduate Coordinator

Accepted for the College

_____

Rodney J. Brown
Dean, College of Biology and Agriculture

ABSTRACT

THE PROLAMINS OF PEARL MILLET

Christian B. Ricks

Department of Plant and Wildlife Sciences

Master of Science

Although work on the prolamins of pearl millet (*Pennisetum glaucum*) has revealed partial amino acid sequences for several alcohol-soluble storage proteins (Marcellino *et al.* 2002) the genes encoding them have not yet been isolated. We constructed a cDNA library from developing *P. glaucum* seed tissue and screened it using maize zein gene probes to isolate several α-prolamin-like gene sequences. The proteins encoded by these genes generally fall into two size classes: 20.6kD and 27.1kD, which we call the 21kD and 27kD pennisetins. Both proteins are similar in composition and sequence to α-prolamins from maize, sorghum and *Coix*. Protein bodies that appear as occlusions within the rough ER of *P. glaucum* endosperm cells are also very similar in size and shape to maize and sorghum protein bodies. The SDS-PAGE gel of the alcohol soluble protein fraction shows two distinct bands in the region corresponding to the 19kD and 22kD of maize α-zein. Both classes of pennisetins appear to be more similar to the 19kD α-zein of maize than to the 22-kD α-zein judging from sequence homology and

maize antibody binding. Phylogenetic reconstruction suggests that *P. glaucum* may have branched from maize prior to the gene duplication which created the 19kD and 22kD α-zein families.

ACKNOWLEDGMENTS

I thankfully acknowledge the support and guidance of many whose assistance proved invaluable in completion of this thesis. My graduate advisor Dr. Craig Coleman provided important mentoring and expertise throughout my undergraduate and graduate work. I also thank graduate committee member Dr. Joshua Udall for his laboratory insight and for the many cycling pulls he selflessly gave. A special thanks to family members and Lauren Douglas for their love and understanding. I thank the many students who throughout the years contributed in this research, in particular Jared Allen and Derek O'Neil who did much of the protein work. In addition I gratefully acknowledge the contributions of Mike Standing for his technical assistance on the transmission electron microscope, Dr. David McClellan for his support with the phylogenetic analysis, and the kindness of Brian Larkins for sharing the γ-zein, 22kD-*α*-zein, 19kD-B1-*α*-zein, and 19kD-D1-*α*-zein antiserum.

TABLE OF CONTENTS

LIST OF FIGURES

THE PROLAMINS OF PEARL MILLET

Introduction

Millets are classified with maize, sorghum, and *Coix* (Job's tears) in the grass sub-family Panicoideae. Genetic diversity of regional millets suggests that pearl millet (*Pennisetum glaucum)* originated in the Sahelian area of Africa and was later taken to East and Central Africa and India to be planted in drier environments there (Obilana and Manyasa 2002). Pearl millet (2n=2x=14) is the most widely grown of all millets and represents the fifth most important of the world's cereal crops when measured by total production and area planted. Well suited for semi-arid tropic regions with short growing seasons and sandy soil, 40 million ha of pearl millet is annually harvested. In Africa and Asia pearl millet is consumed directly by the local populations in the form of breads, beverages, porridges, steamed foods, and boiled foods. In more developed nations pearl millet is used predominantly as a feed grain (Rai *et al*. 1999).

Millets typically contain higher quantities of essential amino acids methionine and cysteine and are higher in fat content than maize, rice, wheat, and sorghum (Obilana and Manyasa 2002). Because pearl millet is a drought-tolerant, ruddy crop with higher protein content and quality than other cereal crops, many people hope to increase its use as a feed grain for both animals and humans alike. Seed protein is responsible for up to 50% of the dry weight of most seeds, and makes a significant contribution to their processing properties and nutritional quality (Shewry and Halford 2002). Yet although millions of hectares of pearl millet are harvested worldwide, providing the main source of protein for over 500 million people, little is known about its important seed storage proteins (Kumar 2002; Virk and Mangat 1997). Biology of pearl millet has been studied little and genetic resources untapped, therefore attempts to improve protein content in

harvested varieties could yield significant improvements (Rai *et al*. 1999). Further understanding the structures of these storage proteins, their biophysical and functional properties, the mechanisms which determine their synthesis, and the genes which encode them will be essential to improve the quantity and quality of these proteins in harvested varieties (Shewry and Halford 2002).

Plants rely upon seed storage proteins as a reserve for sulfur, carbon, and nitrogen during seedling growth and germination. Even though these proteins are chemically and structurally diverse, they are classified into four distinct groups based on their solubility. Albumins are soluble in water, globulins in saline solutions, prolamins in alcohol, and glutelins in alkali solutions (Osborne 1924). Although every plant stores proteins in its seeds by producing some albumins, globulins, prolamins and glutelins, different species have a larger proportion of certain proteins than others. Prolamins are the main contributors of protein body deposition in the endosperm of all cereal grains except oats and rice. The low content of essential amino acids lysine and tryptophan in cereals such as maize, *Coix*, sorghum, and millets is attributable to the negligible amounts of these residues in the prolamin fraction. Thus prolamins have been the focus of research for those trying to improve protein quality in agronomically important harvested crops (Shewry and Casey 1999).

A study of the characterized seed storage proteins of related panicoid grasses might yield valuable information concerning these proteins in pearl millet. The prolamins of maize are called zeins, and are formed within the endosperm of seeds beginning 10-14 days after pollination (Lee *et al*. 1976). While cells in the middle of the endosperm store starch, the outer tissues store a large quantity of zein. Four zein classes:

alpha(*α)*, beta(*β)*, gamma(*γ)*, and delta(*δ)* are responsible for the majority of protein storage in maize. Approximately 70% of maize prolamin comes from *α*-zein of two classes, M_r 19kD and 22kD, which are commonly found as disulfide-bridged dimers. Deduced amino acid sequences from cDNA clones shows the absolute molecular weight for M_r 19kD zeins range from 23-24kD and 26.5-27kD for the M_r 22kD zeins (Coleman and Larkins 1999). Similar to prolamins in other sub-families like Triticeae, *α*-zein is encoded at multiple gene loci, and as many as 40 *α*- zein isoforms have been identified (Wall *et al*. 1984). Gamma is the second most abundant of the maize zeins. Like the multiple sizes of *α*-zein, *γ*-zein has two separate proteins with M_r of 16kD and 27kD. Beta and delta are the least abundant of the zeins, but still contribute significantly to the content and quality of the maize protein. Both rich in methionine, *β*-zein has a single protein of M_r 15kD, while *δ*-zein has both 10kD and 18kD M_r proteins. Structurally, *δ*-zein is different from the other three zeins, and is the only one of the zeins to have tryptophan and lysine amino acids (Swarup *et al*. 1995).

The prolamins of sorghum (kafirins) and *Coix* (coixins) are similar to those of maize, and are also divided into *α*, *β*, *γ*, and *δ* classes (Leite *et al*. 1999). Kafirins of *Sorghum bicolor*, when separated on SDS-PAGE, form major bands at M_r 25kD and 23kD, with minor bands at M_r 28kD, 20kD, 18kD, and 16kD (Shull *et al*. 1991). Protein fraction from sorghum seeds typically consists of 80% *α*-kafirin, 7-8% *β*-kafirin, and 9-12% *γ*-kafirin (Watterson *et al*. 1993). SDS-PAGE of *Coix* seed protein separates the coixins into five polypeptides with M_r 27kD, 25kD, 22kD, 17kD and 15kD (Ottoboni *et al*. 1990). Analysis of several *α*-prolamins from sorghum and *Coix* show amino acid sequence, signal peptide, and a predicted secondary protein structure similar to 22kD

maize α-zein (Ottoboni *et al*. 1993). Like α-zein in maize, the genes encoding α-kafirins and α-coixins are found at multiple loci and are without introns. Coixins and kafirins of sizes M_r 27-28kD with high homology to the 27kD γ –zein have been isolated and shown to be encoded by one or two genes. Sequence analysis of these genes reveals a pattern of conserved cysteine motifs which suggests the protein may play an essential role in protein packaging and folding during synthesis and deposition (Leite *et al*. 1991). The β- and δ –zein-like kafirins and coixins seem to be less conserved than α- and γ -prolamins across species. A coixin with high similarity to β-zein has been isolated (Leite *et al*. 1992), and a 20kD kafirin was shown to cross-react with β-zein antiserum (Shull *et al*. 1991). A kafirin cDNA with homology to the 10kD δ -zein has been sequenced, but no δ –zein-like coixins have been identified (Shewry and Halford 2003).

Storage proteins are cotranslationally localized to the endoplasmic reticulum (ER) by a signal peptide which is cleaved as the peptide enters the lumenal space (Von Heijn 1984). Within the ER these deposited proteins are folded and processed by chaperone proteins and lumenal enzymes which allow for their proper aggregation and transportation (reviewed in Vitale and Denecke 1999). In maize, the 11S and 7S globulin proteins aggregate into smooth ER derived protein storage vacuoles (PSVs) which are then transported to vacuoles through the endomembrane system. The prolamins, however, are retained in the rough ER where they form large protein accretions (Larkins and Hurkman 1978). Proper formation of zein protein bodies within the lumen of the endoplasmic reticulum depends on various protein-protein interactions. The youngest cells contain mainly hydrophilic β- and γ-zeins in the developing protein bodies. As the protein bodies pass from the subaleurone layer into the central endosperm they increase

in size and weight as hydrophobic *α*- and *δ*-zein fill the core and *γ*- and *β*-zein are pushed to the periphery of the aggregate (Lending and Larkins 1989, Esen and Stetler 1992). Thus *α*-zein is not uniformly distributed across the endosperm. Although protein bodies in sorghum appear similar to those in maize, research suggests that they may aggregate in a slightly different manner, with protein bodies consisting of mostly α-kafirins in the peripheral endosperm and accumulating more *β*- and *γ*-kafirins in the central endosperm (Shull *et al*. 1992).

Transgenic experiments with tobacco have shown an interdependence of the different maize prolamins for proper protein body formation and deposition. Without the presence of *γ*-zein, *α*-zein will not accumulate into protein bodies (Coleman *et al*. 1996). Similarly, *α*-zein is stabilized by *β*-zein (Coleman *et al*. 2004). Accumulation of *δ*-zein formed abnormal bodies without the presence of *β*-zein (Bagga *et al*. 1997). These results are consistent with the observations that *β*- and *γ*-zein are synthesized in the peripheral endosperm before *α*- and *δ*-zein, thus allowing for proper packaging as the cells enter the central endosperm. Although similar transgenic experiments with *Coix* and sorghum have not been conducted, homologs of the *β*-, *γ*-, and *δ* -zeins have been identified, and kafirin protein bodies with differential staining observed, suggesting protein bodies in sorghum and *Coix* accumulate in a relatively similar interdependent manner. The interdependence of these prolamins in forming proper protein bodies is especially significant when considering transgenic incorporation of these proteins into other organisms.

The structure of *α*-zein and its relatives has been the subject of numerous studies and prediction algorithms. Sequence analysis of *α*-zein like prolamins has revealed that

the center of the protein consists of a series of 20 amino acids, repeated 9-10 times (Pedersen *et al*. 1982). In maize, the 19kD and 22kD α-zein differ by one such repeat. Most estimates agree that these repeated regions compose α-helices which are arranged in an anti-parallel fashion, although whether these aggregate in a hexagonal net (Garratt *et al*. 1993), coiled-coils to form a superhelical conformation (Bugs *et al*. 2004), or individual α-helices lying antiparallel in a plane (Matushima *et al*. 1997) remains unclear. Most recently Momany *et al*. (2006) predicted a structure of α-zein with coiled-coil tendencies with α-helices creating a triple superhelix.

In *P. glaucum*, as in sorghum, maize, and *Coix*, the major protein storage fraction consists of the alcohol soluble prolamins (Okoh *et al*. 1985). These prolamins, called pennisetins, comprise 33.1-49.5% of the total protein fraction. Alkali-soluble glutelins varied from 30.6-45.3%, saline-soluble globulins 11.6-16.8%, and water-soluble albumins 6.4-9.6% (Chandna and Matta 1990). Amino acid composition is similar to that of zeins and kafirins, with high concentrations of glutamate, glutamine, alanine, leucine, and relatively low contents of cyteine, methionine, and proline (Okoh *et al*. 1985). Extraction of the alcohol soluble protein fraction showed 3 distinct bands. These pennisetins of M_r 27kD, 22kD, and 12kD were named A, B, and C pennisetins, respectively (Marcellino *et al*. 2002). The major protein band of A-pennisetin, when analyzed on a two-dimensional electrophoretic gel, showed at least 4 major polypeptides while B-pennisetin had three (Marcellino *et al*. 2002). N-terminal sequencing of A, B, and C-pennisetins showed no homology to those of other panicoid cereals. Although peptide sequences with related cereals are seemingly divergent, the antiparallel α-helices and protein models of pennisetin and α-zein are nevertheless quite similar (Bugs *et al*.

2004). Evidence supports the structure of pennisetin to be a rod-shaped structure packed with α-helices (Sainani *et al.* 1989).

Despite preliminary protein work on these pennisetins, the genes have not yet been characterized, proteins have not been studied for structural and sequence similarity to relatives, and the manner of protein body deposition within developing seed remains unclear. We present here the characterization of several α-zein-like pennisetins and analyze their expression, deposition, evolution, and homology to related α-prolamins.

Results

Total Prolamin Fraction

Prolamins can be effectively isolated in a single extraction using alcohol in the presence of a reducing agent as outlined in Wallace *et al*., (1990). Using these solubility properties, the prolamin fraction was extracted from seeds of maize, sorghum, *Coix*, and pearl millet and separated using SDS-PAGE (Fig. 1). Pearl millet exhibits two major bands of M_r 22kD and 19kD in the same region as the 19kD and 22kD α-zeins of maize. Pearl millet does not appear to have any major banding larger than 22kD. Several smaller bands of lower relative molecular weights are also present.

Cloning of α-zein-like Pennisetins

To identify full length pennisetin genes a cDNA library was created using mRNA taken from seeds during development. To ensure that our library contained zein-like mRNA several mRNA extractions were used from different stages of development to create a northern blot. The mRNA with greatest hybridization to the 22kD α-zein probe

(results not shown) was used to construct a cDNA library for screening with a maize 22kD α-zein probe. From this dot blot screening a clone encoding a predicted 20.86kD peptide with 57% homology to a *Panicum sumatrense* zein-like seed storage protein (AAW82166) and 48% homology to a maize 22 kD α-zein (AAK32953) was isolated.

Based on the cDNA sequence primers were designed to amplify a portion of the 21kD pennisetin to use as a probe to re-screen the genomic library. Hybridization yielded several more α-zein-like clones encoding putative mature proteins of 20.7kD and 27.1kD. Re-screening with the 27.1kD pennisetin genomic fragment as a probe yielded several more 27.1kD α-like clones. A total of nine unique clones were isolated judging from differences in the translated peptides. These clones encode proteins of predicted mature molecular weights of 20.86kD (21kD pennisetin), 22.06kD (22kD pennisetin), and 27.1kD (27kD pennisetin). Screening the cDNA library with γ-, β-, and δ-zein probes did not yield any positive hits.

Northern Blot Analysis

After identification of several α-pennisetin clones the expression of α-pennisetin mRNA was analyzed. A northern blot was performed with 1µg of mRNA taken from seeds collected every two days after anthesis (DAA) (Fig. 2A). Probing the blot with the 27kD α-pennisetin gene showed that expression of α-pennisetin mRNA begins approximately at day 10 and continues throughout seed maturation until day 24. Accumulation of α-pennisetin mRNA appears to peak at 12 DAA. These data are similar to the expression of zein mRNA in maize which begins 10-14 days after pollination (Lee *et al*. 1976).

Western Blotting

The accumulation of pennisetins was observed by extracting the alcohol-soluble protein fraction from endosperm of seeds taken every two days following anthesis. The SDS-PAGE gel shows accumulation of pennisetins starting at day 12 and increasing throughout seed maturation, consistent with the expected delayed protein synthesis according to the mRNA expression profile (Fig. 2B).

The SDS-PAGE separation of the alcohol-soluble fraction of seed protein in the presence of a reducing agent shows two distinct bands at M_r 22kD, 19kD, and various minor bands of lower molecular weight (Fig.1). Maize polyclonal *α*-zein antibody reacted with the 19kD and 22kD zeins and the *α*-kafirins as expected, but only cross-linked to the larger M_r 22kD pennisetin band (Fig. 3A). Anti-γ-zein serum reacted with the 27kD γ-zein and *γ*-kafirin, but failed to react with any of the major pennisetin fractions (Fig. 3B). Anti-22kD-*α*-zein serum reacted with both zein and kafirin fractions, but failed to cross react with pearl millet (Fig. 3C). Anti-19kD-B1-*α*-zein serum bound the 19kD-zein and cross-reacted with the 23kD kafirin as well as the M_r 22kD pennisetin (Fig. 3D) while anti-19kD-D1-*α*-zein serum failed to cross react with prolamins from either sorghum or pearl millet (Fig. 3E).

Southern Blot Analysis

With several unique *α*-pennisetin clones identified we were interested in estimating the number of genes encoding these *α*-pennisetins in the pearl millet genome. A southern blot with four wells was created using 10μg genomic DNA digested with *BamHI, HindIII, EcoRV* or *KpnI* and probed it with 27kD and 21kD *α*-pennisetin gene

probes (Fig. 4). The 21kD southern blot shows 5,2,1, and 2 bands in lanes 1 through 4, respectively. This suggests a multigenic family of approximately five loci encoding 21kD pennisetins. We identified 5 unique clones encoding the 21kD pennisetins, consistent with our southern blot. The 27kD probe shows 5, 6, 5, and 4 bands in lanes 1 through 4, respectively. This indicates a family of approximately six loci encoding the 27kD pennisetins. We identified 3 unique 27kD pennisetin clones, also consistent with the multiple loci shown in the southern blot.

Electron Microscopy

Given the unique prolamin profile of pearl millet, we were interested in the molecular structure of these protein bodies in the endosperm. Protein bodies found in other species have been found to be largely interdependent in the aggregation, yet we failed to identify *γ-, β-,* and *δ*-zein-like homologs in our cDNA library or with *γ*-zein antisera. *P. glaucum* seeds taken at 12 days after anthesis were sectioned and stained with OsO_4 prior to microscopy. Protein bodies were seen to accumulate in a manner similar to maize and sorghum, with accumulation of spheroid protein bodies from 0.5-1.5μm in diameter (Fig. 5). In maize and sorghum, prolamins form spherical protein body accretions with diameters of .3-2μm surrounded by rough endoplasmic reticulum. Pearl millet protein bodies were surrounded by a rough endoplasmic reticulum membrane and were often found adjacent to the ER network. Protein bodies were also observed with light and dark staining, indicating regions of different densities (Fig. 5C). Dark regions often appeared in the center core of the protein body or as darker inclusions.

Multiple Sequence Alignment

To better understand the relationships of α-prolamins across species the peptide sequences of the 21, 22, and 27kD pennisetins were aligned against the α-prolamins from maize, sorghum, and *Coix* (Fig. 6). Alignment of these α-prolamins showed a clear conservation of sequence and structure. The N-terminus, C-terminus, and signal peptide regions exhibit high similarity across species. Some differences do arise when looking at the repeated motifs which comprise the α-helical repeats necessary for efficient packing into protein bodies. Like previous work done by Petersen *et al.* (1982), our results show the 19kD and 22kD α-zeins differ by the presence of one repeated motif, or one α-helix. Our alignment shows that α-kafirins and α-coixins are seemingly more like the 22kD α-zein than the 19kD α-zein judging by sequence homology and number of repeats. The major 21kD and 27kD pennisetins have 8 and 10 repeats, similar to the 9 and 10 repeats of the 19 and 22kD α-zeins. The 21kD pennisetin is more similar to the 19kD α-zein, judging by the absence of repeat 3 in both alignments and other similar deleted regions. The three α -pennisetin classes (21, 22, and 27kD) are also closely related to each other, as evidenced by their unique absence of repeat 6. The 27kD pennisetin appears quite unique from the other α-prolamins. Although it has 10 repeats like the other 22kD α-zein-like prolamins, repeats 6 and 8 are not the same, and even shared repeats show sequence divergence. The 22kD pennisetin most closely resembles the 27kD pennisetin except for the absence of two α-helical repeats, giving it a total of eight.

Phylogenetic Tree Analysis

To compare our characterized pennisetins with those of related species we searched the non-redundant database for α-prolamin sequences from maize, sorghum, and *Coix*. We created a Bayesian consensus cladogram to analyze the relationship of these various α-prolamins (Fig. 7). Two distinct clades appeared in all alignments, corresponding to the 19kD and 22kD α-zeins and their homologs. Kafirins and coixins all grouped with the 22kD clade, while all pennisetins grouped with the 19kD clade. The pennisetins are not monophyletic despite having some sequence similarity in the amino acid alignment. The separation of these α-prolamins is supported by high posterior probabilities.

Discussion

Identification and Characterization of α-Pennisetins

Mature peptides from the deduced amino acid sequences of complete pennisetin cDNAs generally fall into two groups, with five unique peptides of sizes ~20.7kD and two unique peptides of sizes ~27.1kD identified in our library. We also identified a 22.07kD and a 12.0kD pennisetin. These predicted proteins are referred to as 27kD, 22kD, 21kD, and 12kD pennisetins. Southern blotting of *P. glaucum* genomic DNA using homologous 27kD and 21kD pennisetin probes shows a multigenic family for both pennisetin classes (Fig. 4). Banding suggests at least six loci encoding the 27kD pennisetins and five loci encoding the 21kD pennisetins. Given the highly inbred USDA strains we assume that allelic differences are minimal, and thus banding represents the number of loci present. The related α-prolamins of panicoid grasses maize, sorghum, and

Coix are similarly organized into multigene families with abundant pseudogenes. Maize is the best characterized, and contains over 23 members of the 22kD gene family alone and numerous pseudogenes (Song *et al*. 2001). Because of multiple bands on a Southern blot, previous quantification of *α*-zein in the maize genome predicted 75-100 loci (Hagen and Rubenstein 1981). It is therefore apparent from our southern blots that pearl millet has a genome with fewer loci encoding *α*-prolamins than maize, and may have a more similar organization to the sorghum and *Coix* genomes.

Our finding that the pennisetins exist as multigene families is consistent with 2-D analysis by Marcellino *et al*. showing at least four-27kD and three-21kD major peptide groups (2002). This analysis was unable to separate peptides with similar charged substitutions however, so the actual number of pennisetins encoding loci can be greater, as our estimates suggest. The 27.1kD protein matches the A-pennisetin short N-terminal peptide sequence Marcellino *et al*. (2002) reported. We were, however, unable to correlate our sequences with his partial B- and C-pennisetin N-terminal codes, despite identifying deduced amino acid sequences of corresponding sizes (12kD and 22kD).

Two clones encoding putative proteins of mature sizes 22kD and 12kD with homology to the 27kD pennisetins were also identified. Sequence analysis shows a 47bp deletion causing a frame-shift accounts for the truncated 12kD peptide, while an in-frame deletion of 135bp from the 27kD pennisetin created the 22kD peptide. It is not clear whether these peptides form functional protein bodies, if they are even translated. Finding nonsense mutations in pennisetin sequences is not surprising given the numerous *α*-prolamin-like pseudogenes found in related grasses. However, alignment of amino acids (Fig. 6) shows that the deletion in the 22kD peptide accounts for exactly two *α*-

helical repeats, strongly suggesting this peptide has retained its function to assemble and aggregate properly, even though a strong M_r 22kD band in the SDS-PAGE gel is not present (Fig. 1).

The SDS-PAGE gel of the alcohol soluble protein fraction of pearl millet shows two distinct major bands in the region corresponding to the 19kD and 22kD maize *α*-zeins. We have already correlated our 27kD pennisetin with the larger of these bands based on alignment with the N-terminal sequence (Marcellino *et al.* 2002). Based on size and expression of transcripts in our cDNA library we predict the smaller band to be the 21kD pennisetin. This would seemingly suggest that these pennisetins have homology to the 19kD and 22kD *α*-zeins, respectively. Western blotting with zein antibodies, however, showed that the only pennisetin to cross-react with maize antibody was the larger 27kD pennisetin to the 19kD-B1 antibody. This may simply be a function of the quantity of protein present, as the 21kD pennisetin class clearly accumulates in lower concentrations (Fig 1). Despite this ambiguity, antibody binding seems to suggest that the 27kD pennisetin must have a structure closer to the 19kD *α*-zein despite having a similar size to the 22kD *α*-zein. We do not have evidence of a pennisetin class with structural homology to the maize 22kD *α*-zein. This result is particularly curious because the 22kD *α*-zein-like prolamins are well conserved across maize, sorghum, and *Coix,* whereas 19kD *α*-zein-like prolamins are seemingly less conserved.

Alignment of the amino acid sequences of *α*-prolamins from maize, sorghum, *Coix*, and pearl millet emphasizes the many similarities between the encoded peptides. Conservation of the signal peptide in pearl millet suggests these proteins are also deposited to the ER. This conservation is not surprising given the way we observed

rough endoplasmic reticulum bound pearl millet protein bodies forming in the endosperm cells in a manner resembling maize and sorghum prolamin deposition. The 21kD pennisetin has some structural and sequence homology with the 19kD α-zein, as both lack repeat 3 and contain similar deleted regions (Fig. 6). Some divergence has clearly occurred, evidenced by the 19kD α-zein containing 9 repeated motifs while the 21kD pennisetin has 8, and by the failure of the 21kD pennisetin to cross react with 19kD α-zein antisera. The 27kD and 22kD pennisetins more closely resemble the 21kD pennisetins than the α-prolamins of other species. This, combined with cross reaction with the 19kD-B1 antibody might suggest that this larger class of pennisetins, despite having 10 repeats and a similar size to the 22kD α-zeins, arose from a 19kD α-zein-like precursor.

Evolution of α -Prolamins

Most prolamins, including the β- and γ-zeins of maize, share common ancestral domains and peptide repeats which link them to a widely distributed "Prolamin Superfamily". The α-prolamins of the panicoid cereals, however, show no sequence similarity to this superfamily and are thought to have evolved independently (Shewry and Tatham 1999). Our phylogenetic analysis of the characterized α-prolamins from maize, sorghum, *Coix*, and pearl millet may shed light into the evolution of this agronomically important protein family (Fig. 7). The kafirins and coixins all group with the 22kD α-zein, while the pennisetins all belong to the 19kD α-zein clade. The 21kD pennisetin is more closely related to the 19kD α-zein than the other pennisetins, as our sequence alignment also suggested. This is somewhat surprising given that the 21kD pennisetin

did not cross react to the 19kD-B1 antibody while the 27kD pennisetin did. Perhaps the 27kD pennisetins have retained a structural similarity to the 19kD α-zein despite sequence divergence. No pennisetin with homology to the 22kD α-zein was identified.

Previous molecular systematics studies of the panicoid grasses have shown pearl millet to be more distantly related to maize than sorghum or *Coix* (Martel *et al*. 2004). The ancestral species may have contained a 19kD α-zein like sequence, which duplicated prior to pearl millet speciation creating the 21kD and 27kD pennisetin clades. Subsequent modification of the 27kD pennisetin produced the 22kD pennisetin class. After speciation of pearl millet the 19kD α-zein-like prolamin in the sorghum, *Coix*, and maize ancestor again duplicated to create the 22kD α-zein-like family. The duplication of prolamins to create new proteins differing in the number of α-helical repeats is consistent with the evolution of the α-prolamins in other species.

While our construction of the phylogenetic tree helps illucidate some events in the evolution of this unique α-prolamin family, it still presents a few ambiguities. It is possible that a gene duplication event created the 19kD and 22kD α-zein-like families before the 19kD duplicated to create the 21kD and 27kD pennisetins prior to pearl millet speciation. If this were the case we would expect to find a 22kD α-zein-like sequence in the pearl millet genome. Although such a sequence has not been identified, it could have been anciently silenced and exist only as a pseudogene.

In either case, because the speciation of *Coix* and sorghum from maize occurred post-duplication of these genes, we would predict that 27kD pennisetin-like, 19kD zein-like, and 22kD zein-like α-prolamins are represented in the sorghum, *Coix*, and maize genomes. Searching available sequences, however, did not identify 19kD zein-like α-

prolamins of sorghum and *Coix* in the non-redundant database. Cross reaction of the smaller kafirin band (Fig. 3D) with 19kD-D1 α-zein antiserum and 22kD α-zein antiserum suggests some structural similarities exist to the 19kD α-zein despite greater sequence homology to the 22kD α-zein. By searching the EST database we were able to identify one EST (BG048546) found on supercontig1 with high (2e-58) homology to 19kD α-zein. Interestingly, this was from an EST library taken from immature sorghum ovaries. It is unclear if this EST is even expressed within developing sorghum endosperm. We were unable to find any 19kD α-zein homologs in *Coix.* The similar alteration of the 19kD α-zein-like kafirins and coixins might suggest that these two species split from each other after branching off maize, allowing for 19kD α-zein alteration before their subsequence speciation. No 27kD pennisetin homologs were identified in either *Coix*, sorghum, or maize, suggesting this gene family may have been completely silenced and exist in the genome only as a pseudogene, or lost altogether. The selective silencing of these multigene families might also suggest that the expansion of these families occurred post speciation.

It is interesting that the 27kD pennisetins should arise from 19kD sequence, but through gene duplication and rearrangement come to contain 10 repeats just like the 22kD α-zein. This suggests an interesting example of concerted evolution, indicating some importance of this larger prolamin class. This is also supported by the greater abundance of the 27kD pennisetin than the 21kD pennisetin (Fig. 1), and by the greater conservation of the 22kD α-zein-like prolamins across maize, sorghum, and *Coix*.

Protein Body Formation and Deposition

We were unable to detect any *γ-, β-,* and *δ*-zein like clones in our cDNA library screening. Although SDS-PAGE of pearl millet pennisetins shows minor bands of low molecular weights, it is unclear of their relation to the *β-,* and *δ*-zeins or the other minor bands of sorghum and *Coix* (Fig. 1). There does not appear to be any corresponding *P. glaucum* protein band to the 27kD γ-zein. 27kD γ-prolamins have previously been identified in *Coix* and sorghum, suggesting that this class of prolamins is well conserved across species. This is not surprising given the importance of γ-zein in proper protein body formation and deposition within the developing endosperm. The identification of γ-coixins and kafirins and the conservation of cysteine residues suggests a similar function in sorghum and *Coix* may be the reason for this conservation. Despite being unable to identify a γ-zein-like pennisetin, the protein bodies that appear as occlusions within the rough ER of *P. glaucum* endosperm cells are very similar to maize and kafirin protein bodies in size, shape, and pattern of deposition. Several of these protein bodies demonstrated light and dark staining (Fig. 5C), possibly indicative that these bodies are composed of aggregates of several pennisetins. This is similar to maize and kafirin protein bodies, which demonstrate lighter staining of *α*-prolamins and darker staining of *β*- and *γ*-prolamins (Lending and Larkins 1989, Esen and Stetler 1992, Shull *et al*. 1992). The ability of *P. glaucum* to properly package and deposit protein bodies in the absence of a γ-zein-like protein raises interesting questions concerning the formation of protein aggregates within developing endosperm. It may be that the 22kD pennisetin plays a key role in protein body formation. Alternatively, the 27kD or 19kD pennisetins may have evolved to aggregate properly in the absence of *γ-, β-,* and *δ*-prolamins. More work is

needed to determine if a γ-zein-like pennisetin, or any other protein is needed for proper protein body formation.

The prolamins of *P. glaucum* are in fact quite distinctive from those of the other panicoid grasses. With most of the prolamin fraction consisting of a unique 27kD *α*-pennisetin class and lacking any 22kD *α*-zein homolog it adds an interesting piece in the analysis of the evolution of this important prolamin family. Yet despite such divergence these seed storage proteins have conserved their form and function, perhaps highlighting the significance of these proteins to both the plants and the people who rely upon them.

Materials and Methods

Plant Material

P. glaucum (PI 532742 01 SD) was obtained from a USDA germplasm resource center. Plants were grown under standard greenhouse conditions at Brigham Young University, Provo, UT, U.S.A. Plants were allowed to self-fertilize prior to seed collection.

cDNA Library Construction and Screening

Developing seed taken 14 days after pollination was collected and immediately frozen in liquid nitrogen. Total RNA was obtained using a Plant RNeasy kit (Qiagen, Valencia, CA, U.S.A.) according to manufacturer's instructions. A 1%-agarose formaldehyde gel was prepared and loaded with 1μg total RNA and electrophoresed at 40V overnight to ensure quality of mRNA as measured by prevalence of 18s and 28s rRNA. Northern blot was prepared as described below and probed using a maize 22kd *α*-

zein as template to ensure prevalence of zein-like mRNA. Using the SMART cDNA kit (Clontech Laboratories, Mountain View, CA, U.S.A.) cDNA was synthesized according to manufacturers instructions. cDNAs were inserted into plasmids using the pGEM-T Vector System 1 (Promega, Madison, WI, U.S.A.) was followed by heat shock transformation into XL1-Blue competent cells (Stratagene, La Jolla, CA, U.S.A.) and plating out onto X-Gal plates. Fourteen 384-well plates were created and double-stamped onto Hybond-N+ membrane followed by liberation and UV-crosslinking according to manufacturers instructions (Amerisham Biosciences, Piscataway, NJ, U.S.A.). Probes for hybridization were initially generated using a maize 22kd α-zein, then from previously sequenced 20.7kD pennisetin amplified from genomic DNA using the Prime-a-Gene labeling reaction (Promega, Madison, WI, U.S.A.) with [^{32}P] dCTP (PerkinElmer, Wellesley, MA, U.S.A.). Denatured probes were hybridized to the cDNA library membranes in Ekono hybridization solution (Research Products International Corp., Mt. Prospect, IL, U.S.A.) at 60-65C. Blots were washed at low and high stringency washes according to the Hybond-N+ protocol then exposed to a phosphor imaging screen (Kodak, Rochester, NY, U.S.A.) for 3 hours. The imaging screen was scanned using a Molecular Imager FX system and images of the blots obtained using Quantity One PC software (Bio-Rad Laboratories, Hercules, CA, U.S.A.).

DNA Extraction and Southern Blot Analysis

Plant tissue was immediately frozen in liquid nitrogen and DNA extracted using a DNeasy kit according to manufacturer's instructions (Qiagen, Valencia, CA, U.S.A.). DNA was quantified using a NanoDrop-1000 spectrophotometer (Nanodrop

Technologies, Wilmington, DE, U.S.A.). 10 µg DNA was digested using *BamHI, HindIII, EcoRV* or *KpnI* and added to each well before electrophoresis in .8% agarose at 40V for 12 hours. Transfer to Hybond-N+ membrane according to manufacturer's instructions (Amerisham Biosciences, Piscataway, NJ, U.S.A.). Probes for hybridization were generated from 27.1kD or 20.7kD pennisetin genes amplified from genomic *P. glaucum* DNA. Using the Prime-a-Gene labeling reaction (Promega, Madison, WI, U.S.A.) imaging were performed as described above for screening the cDNA library.

RNA Extraction and Northern Blot Analysis

Total RNA was extracted from seed harvested every two days after anthesis and placed immediately in liquid nitrogen and stored at -80C. The RNA was extracted using a Plant RNeasy kit (Qiagen, Valencia, CA, U.S.A.) according to manufacturer's instructions. RNA was quantified using a NanoDrop-1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE, U.S.A.). For Northern blotting a 1%-agarose formaldehyde gel was prepared and loaded with 1µg total RNA in each well and electrophoresed at 40V overnight. Following electrophoresis the RNA was transferred to Hybond-N+ membranes according to manufacturer's protocol (Amerisham Biosciences, Piscataway, NJ, U.S.A.). Probes for hybridization were generated from 27.1kD pennesetin gene DNA amplified from genomic *P. glaucum* DNA. Using the Prime-a-Gene labeling reaction (Promega, Madison, WI, U.S.A.), hybridization, and imaging were performed as described above for screening the cDNA library.

Protein Extraction and Immunoblot Analysis

Pennisetins from *P. glaucum* were extracted according to the procedure outlined in Wallace *et al*., (1990). For an analysis of protein accumulation in developing endosperm tissue pennisetins were extracted from nine seeds frozen in liquid nitrogen collected every two days after anthesis until maturity. Extracted pennisetins were resuspended in 1X Laemmli buffer. Proteins were electrophoresed in a 12% SDS-polyacrylamide gel and run at 155V for one hour. The gel was fixed overnight in 40% ethanol, 10% glacial acetic acid solution and stained with 1X flamingo fluorescent stain (Biorad, Hercules, CA, U.S.A.). Alternatively, SDS-polyacrylamide gels were stained in 0.15% Coomassie Blue.

Comparative analysis of the total prolamin fraction from maize (W64A), sorghum (PI 597969 01 SD), *Coix* (PI 320865) and pearl millet was also done according to the procedure outlined in Wallace *et al*., (1990) and as indicated above.

For western blot analysis total prolamins from maize, sorghum, and pearl millet were extracted according to the protocol above. Alcohol-soluble protein was diluted in loading dye, subjected to SDS-PAGE, and blotted onto Immun-Blot PVDF membrane (Biorad, Hercules, CA, U.S.A.) using a Trans-Blot SD Semi-Dry Electrophoretic Transfer Cell (Biorad, Hercules, CA, U.S.A.) according to manufacturers protocol. Membranes were soaked in a 1X TBS solution for five minutes after which the membranes were soaked in 20ml of a 1X TBS/Casein Blocker solution for 15 minutes. .5µl of rabbit primary antibody was added to the solution and agitated gently overnight at room temperature. Polyclonal α-zein antiserum was obtained from Coleman *et al*. (2004), γ-zein antiserum from Lending *et al*. (1988), and 22kD-α-zein antiserum, 19kD-

B1-α-zein antiserum, and19kD-D1-α-zein antiserum from Woo *et al*. (2001). Membranes were subsequently washed three times with 1X TBS solution for five minutes each. The membrane was then agitated in a 1:30,000 solution of anti-rabbit secondary antibody conjugated to alkaline phosphatase (Sigma-Aldrich, St, Louis, MO, U.S.A) for two hours. Membranes were again washed three times with 1X TBS solution and then two times with ddH20 for 30 seconds. Membranes were covered and soaked with resuspended Sigma Fast 5-Bromo-4-Chloro-3-Indolyl Phosphate/Nitro Blue Tetrazolium Tablets (Sigma-Aldrich, St. Louis, MO, U.S.A) until proteins were visualized.

Transmission Electron Microscopy

For transmission electron microscopy *P. glaucum* seeds taken at 12 days after anthesis were sliced into eighths and immersed in ice-cold fixative (2% glutaraldehyde, 60mM sodium cacodylate, pH 7.3) for two hours. Samples were brought to room temperature and washed six times (10 min each) in a 30 mM sodium cacodylate (pH 7.2-7.4) buffer, followed by staining in a 1% OsO_4 solution buffered in 60 mM sodium cacodylate for two hours. Samples were washed six times, 10 min each, with 30 mM sodium cacodylate and placed in a 0.5% uranyl acetate solution overnight. Samples were dehydrated in a graded ethanol series: 10 min each in 10%, 30%, 50%, 70%, and 95% ethanol followed by 3 washes, 10 min each, in 100% ethanol and 3 washes, 10 min each, of 100% acetone. Samples were suspended for 1 hour each in 30%, 70%, and 100% Spurr's resin in acetone (Spurr 1969). The samples were transferred to fresh 100% Spurr's resin before polymerizing at 70C overnight. 80nm sections were cut with a diamond knife on a RMC MT-X ultramicrotome (Boeckeler Instruments, Tuscan, AZ,

U.S.A) and collected in copper grids prior to microscopy using the FEI Tecnai T-12 transmission electron microscope (FEI Company, Hillsboro, OR, U.S.A.).

Sequence Alignment

Sequences of the α-prolamins from maize, sorghum, *Coix*, and pearl millet were aligned to look for reasons for patterns of divergence. We chose 22kD α-zein-like sequences from maize (AF371277), sorghum (Y17556), and *Coix* (X63113). 19kD α-zein-like sequences have not been identified in *Coix* (DeRose *et al*. 1989) and sorghum (Ottoboni *et al*. 1993), so we only included the maize 19kD-B1 (AF371269) sequence. We included three pennisetin sequences: a 21kD (EF608506), 22kD (EF608504), and 27kD (EF608507) for comparative analysis. Sequences were divided into their signal peptide, N-terminus, repeat motifs, and C-terminus referencing the alignments of Argos *et al*. (1982) for maize, DeRose *et al.* (1989) for sorghum, and Ottoboni *et al*. (1993) for *Coix*. This alignment of multiple α-prolamins agreed with that of Garratt *et al.* (1993). Using the N-terminal pennisetin sequences identified by Marcellino *et al*. (2002) the signal peptide of the 27kD pennisetin was deduced to consist of the first 21 amino acids. The pennisetins were then aligned to maximize sequence homology between repeat motifs and place poly-glutamine sequences at the repeat ends. Gaps were introduced for optimal alignment.

Phylogenetic Tree Analysis

The GenBank non-redundant database was searched for α-prolamin sequences from maize, sorghum, and *Coix*. These raw sequence data were used to make a neighbor

joining tree to arrive at a strategy for taxonomic sampling. Representative sequences from closely related clades were selected to maximize divergence and increase the phylogenetic signal. ModelTest was used to determine which model used the fewest number of parameters yet failed to significantly deviate from the most general model (Posada and Crandall 1998). The result was GTR+I+G, which tracks 10 free parameters. This model was then used in MrBayes, sampling every 1000 generations (Ronquist and Huelsenbeck 2003). From this sampling PAUP was used to generate a consensus with posterior probabilities to estimate the branch support (Swoffard 2002).

References

Argos P, Pedersen K, Marks MD, Larkins BA (1982) A structural model for maize zein proteins. J Biol Chem 257: 9984-9990

Bagga S, Adams HP, Rodriguez FD, Kemp JD, Sengupta-Gopalan C (1997) Coexpression of the maize δ-zein and β-zein genes results in stable accumulation of δ-zein in endoplasmic reticulum-derived protein bodies formed by β-zein. Plant Cell 9: 1683-1696

Bietz, JA (1982) Cereal prolamin evolution and homology revealed by sequence analysis. Biochem Gen 20:1039-1051

Bugs MR, Forato LA, Borteleto-Bugs RK, Fischer H, Mascarenhas YP, Ward RJ, and Colnago LA (2004) Spectroscopic characterization and structural modeling of prolamin from maize and pearl millet. Eur Biophys J 33:335-343

Chandna M, Matta NK (1990) Characterization of pearl millet protein fractions. Phytochem 29: 3395-3399

Coleman CE, Herman EM, Takasaki K, Larkins BA (1996) The maize γ-zein sequesters α-zein and stabilizes its accumulation in protein bodies in transgenic tobacco endosperm. Plant Cell 8: 2335-2345

Coleman CE, Larkins BA (1999) The prolamins of maize. In: Shewry PR, Casey R (eds) Seed Proteins. Kluwar Academic Publishers, Dordrecht, 109-139

Coleman CE, Lopes MA, Gillikin JW, Boston RS, Larkins BA (1995) A defective signal peptide in the maize high-lysine mutant floury 2. Plant Biol 92: 6828-6831

Coleman CE, Yoho PR, Escobar S, Ogawa M (2004) The accumulation of α-zein in transgenic tobacco endosperm is stabilized by co-expression of β-zein. Plant Cell Physiol 45(7): 864-871

DeRose RT, Ma DP, Kwon IJ, Hasmain SE, Klassy RC, Hall TC (1989) Characterization of the kafirin gene family from sorghum reveals extensive homology with zein from maize. Plant Mol Biol 12:245-256

Esen A, Stetler DA (1992) Immunocytochemical localization of δ-zein in protein bodies of maize endosperm cells. Amer J Bot 79: 243-248

Garrat R, Oliva G, Caracelli I, Leite A, Arruda P (1993) Studies of the zein-like α-prolamins based on an analysis of amino acid sequences: Implications to the evolution and three-dimensional structure. Proteins: Structure, Function and Genetics 15: 88-99

Hagan G, Rubenstein I (1981) Complex organization of zein genes in maize. Gene 13: 239-249

Kumar KA (2002) Pearl Millet. In: Chopra VL, Prakash S (eds) Evolution and adaptation of cereal crops. Science Publishers Inc., Enfield, 239-290

Larkins BA, Hurkman WJ (1978) Synthesis and deposition of zein in protein bodies of maize endosperm. Plant Physiol 62: 256-263

Lee KH, Jones RA, Dalby A, Tsai Cy (1976) Genetic regulation of storage protein content in maize endosperm. Biochem Genet 14: 641-650

Leite A, Freitas FA, Yunes JA, and Arruda P (1991) Nucleotide sequence of a cDNA clone encoding *γ*-coixin from *Coix lacryma-jobi* seeds. Plant Physiol 97: 1604-1605

Leite A, Neto GC, Vettore AL, Yunes JA, and Arruda P (1999) The prolamins of sorghum, *Coix* and millets. In: Shewry PR, Casey R (eds) Seed Proteins. Kluwar Academic Publishers, Dordrecht, 141-155

Leite A, Yunes JA, Turcinelli SR, and Arruda P (1992) Cloning and characterization of a cDNA encoding a sulfur-rich coixin. Plant Mol Biol 18: 171-174

Lending CR, Kriz AL, Larkins BA, Bracker CE (1988) Structure of maize protein bodies and immunocytochemical localization of zeins. Protoplasma 143: 51-62

Lending CR, Larkins BA (1989) Changes in zein composition of protein bodies during maize endosperm development. Plant Cell 1: 1011-1023

Marcellino LH, Bloch C, Gander ES (2002) Characterization of pearl millet prolamins. Protein and Peptide Letters 9: 237-244

Martel E, Poncet V, Lamy F, Siljak-Yakovlev S, Lejeune B, Sarr A (2004) Chromosome evolution of *Pennisetum* species (Poaceae): implications of ITS phylogeny. Plant Syst Evol 249: 139-149

Matsushima N, Danno G, Takezawa H, Izumi Y (1997) Three-dimensional structure of maize *α*-zein proteins studied by small-angle X-ray scattering. Biochem Biophys Acta 1339: 14-22

Momany FA, Sessa DJ, Lawton JW, Selling GW, Hamaker SAH, Willett JL (2006) Structural characterization of α-zein. J Agric Food Chem 54: 543-547

Obilana AB, Manyasa E. (2002) Millets. In 'Pseudocereals and less common cereals: grain properties and utilization potential' (P.S. Belton and J.R.N. Taylor eds), Springer-Verlag, New York, 177-217

Okoh PN, Nwasike CC, Ikediobi CO (1985) Studies on seed proteins of pearl millets. J Agric Food Chem 33: 55-57

Osborne TB (1924) The Vegetable Proteins. Longmans Green, London

Ottoboni LMM, Leite A, Targon MLPN, da Silva MJ, Arruda P (1990) Heterogeneity of *Coix*, maize and teosinte prolamins detected by isoelectric focusing. Rev Bras Genet 13: 313-322

Ottoboni LMM, Leite A, Yunes JA, Targon MLPN, Souza Filho GA, Arruda P (1993) Sequence analysis of 22 kDa–like α-coixin genes and their comparison with homologous zein and kafirin genes reveals highly conserved protein structure and regulatory elements. Plant Mol Biol 21: 765-778

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14: 817-818

Pedersen K, Devereux J, Wilson DR, Sheldon E, Larkins BA (1982) Cloning and sequence analysis reveal structural variation among related zein genes in maize. Cell 29: 1016-1026

Rai KN, Murty DS, Andrews DJ, Bramel-Cox PJ (1999) Genetic enhancement of pearl millet and sorghum for the semi-arid tropics of Asia and Africa. Genome 42: 617-628

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574

Sainani MN, Lachke AH, Sahasrabudhe NA, Ranjekar PK (1989) Viscometric characterization of pennisetin from pearl millet. Biochem Biophys Res Commun 30: 334-341

Shewry PR, Casey R (1999) Seed proteins. In: Shewry PR, Casey R (eds) Seed Proteins. Kluwar Academic Publishers, Dordrecht, 1-9

Shewry PR, Halford NG (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. Exper Botany 53: 947-958

Shewry PR, Halford NG (2003) The prolamin storage proteins of sorghum and millets. http://www.afripro.org.uk/papers/Paper03Shewry.pdf cited 06/06/07

Shewry PR, Tatham AS (1999) The characteristics, structures, and evolutionary relationships of prolamins. In: Shewry PR, Casey R (eds) Seed Proteins. Kluwar Academic Publishers, Dordrecht, 11-33

Shull JM, Watterson JJ and Kerleis AW (1991) Proposed nomenclature of the alcohol soluble proteins (kafirins) of *Sorghum bicolor* (L Moench) based of molecular weight, solubility and structure. J Agric Food Chem 39: 83-87

Shull JM, Watterson JJ, Kirleis AW (1992) Purification and immunocytochemical localization of kafirins in *Sorghum bicolor* (L. Moench) endosperm. Protoplasma 171: 64-74

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

Spurr AR (1969) A low-viscosity epoxy resin embedding medium for electron microscopy. J Ultrasound Res 26: 31-43

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

Swofford DL (2002) PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Vitale A, Denecke J (1999) The endoplasmic reticulum-Gateway of the secretory pathway. Plant Cell 11: 615-628.

Von Heijn G (1984) How signal sequences maintain cleavage specificity. J Mol Biol 173: 243-251

Virk DS, Mangat BK (1997) Pearl Millet. In: Bahl PN, Salimath PM, Mandal AK (eds) Genetics, Cytogenetics and Breeding of Crop Plants. Science Publishers Inc., Enfield, 241-265

Wall JS, Fey DA, Paulis JW (1984) Improved two dimensional electrophoretic separation of zein proteins: application to study of zein inheritance in com genotypes. Cereal Chem 61: 141-146

Wallace JC, Lopes MA, Paiva E, Larkins BA (1990) New methods for extraction and quantitation of zeins reveal a high content of γ-zein in modified opaque-2 maize. Plant Physiol 92: 191-196

Watterson JJ, Shull JM, Kirleis AW (1993) Quantitation of α-, β-, and γ-kafirins in vitreous and opaque endosperm of Sorghum bicolor. Cereal Chem 70: 452-457

Woo MW, Hu DWH, Larkins BA, Jung R (2001) Genomics analysis of genes expressed in maize endosperm identifies novel seed proteins and clarifies patterns of zein gene expression. Plant Cell 13: 2297-2317

Tables and Figures

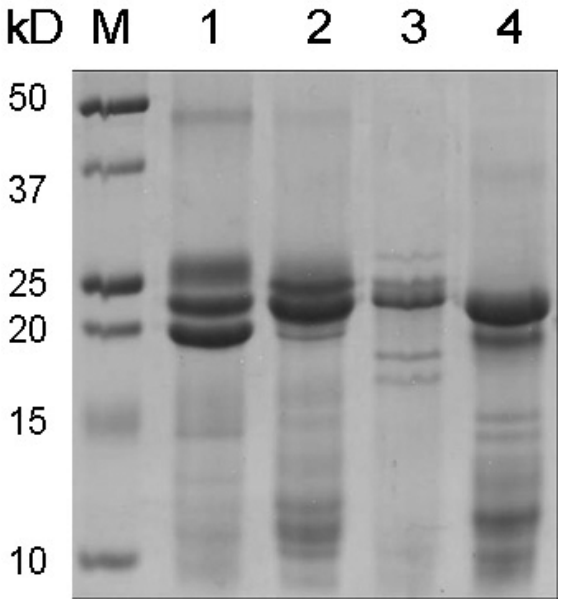

Fig.1.
Prolamin proteins from maize, sorghum, *Coix*, and pearl millet. Alcohol-soluble prolamins were separated by SDS-PAGE and stained with Coomassie Blue. Lane 1 = maize, lane 2 = sorghum, lane 3 = *Coix*, lane 4 = pearl millet.

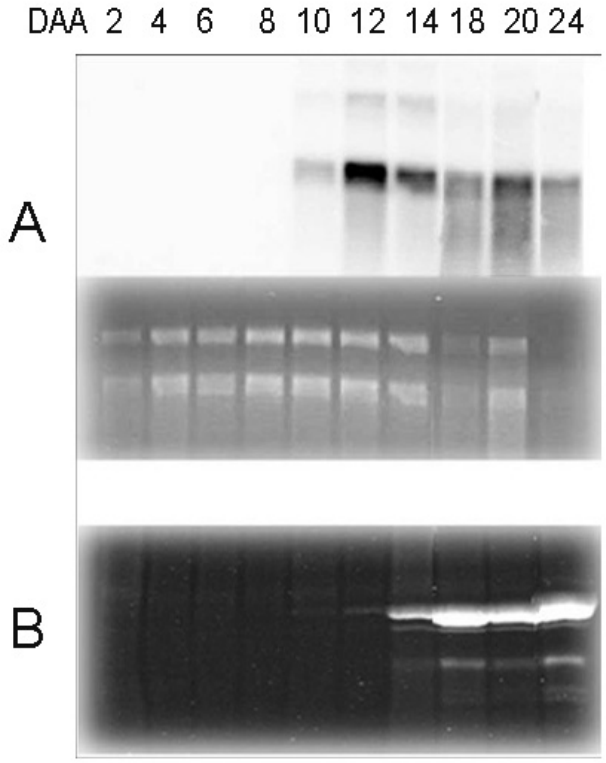

Fig. 2.
Expression of pearl millet mRNA and protein accumulation during seed development. (A) Northern blot of pearl millet seed mRNA taken at two day intervals following anthesis was transferred to a membrane and probed using homologous 27kD-pennisetin template. Formaldehyde gel shows loading of 1 μg total RNA. Intensity of rRNA bands is largely consistent across all time points. (B) Alcohol-soluble prolamins were separated by SDS-PAGE and stained with flamingo orange.

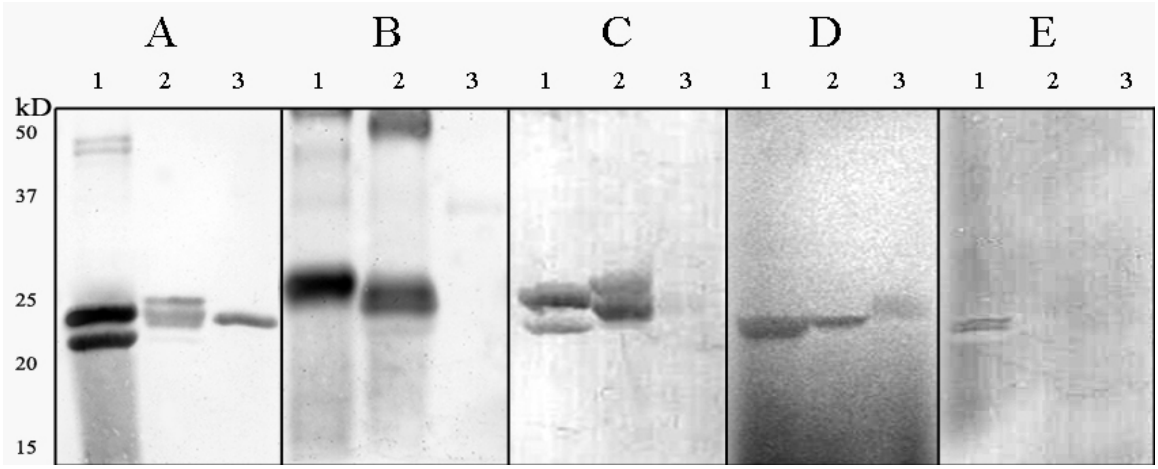

Fig. 3.
Prolamin proteins from maize, sorghum and pearl millet. Alcohol-soluble prolamins were separated by SDS-PAGE and reacted with α-zein antiserum (A), γ-zein antiserum (B), 22kD-α-zein antiserum (C), 19kD-B1-α-zein antiserum (D), and 19kD-D1-α-zein antiserum (E). Lane 1 = maize, lane 2 = sorghum, lane 3 = pearl millet.

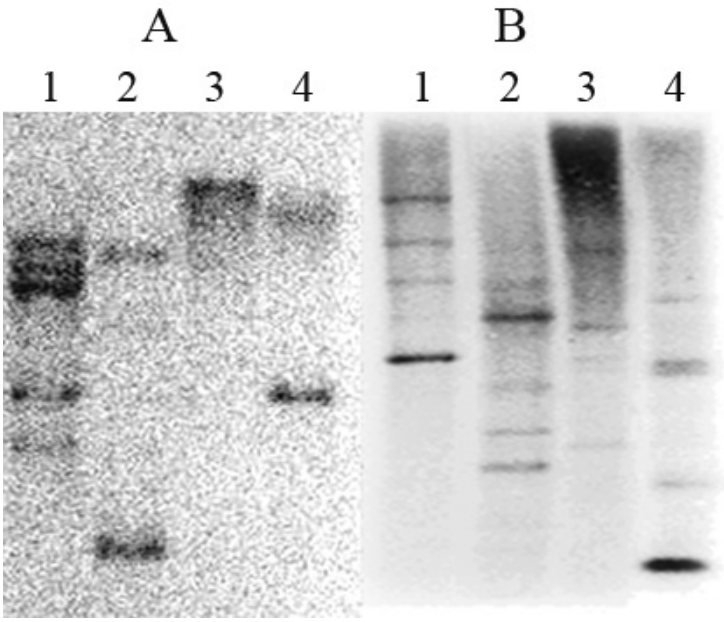

Fig. 4.
Southern blot using *P. glaucum* genomic DNA digested using *BamHI*(lane 1), *HindIII*(lane 2), *EcoRV*(lane 3), and *KpnI*(lane 4) and probed with 21kD (A) and 27kD (B) pennisetin homologous probe.

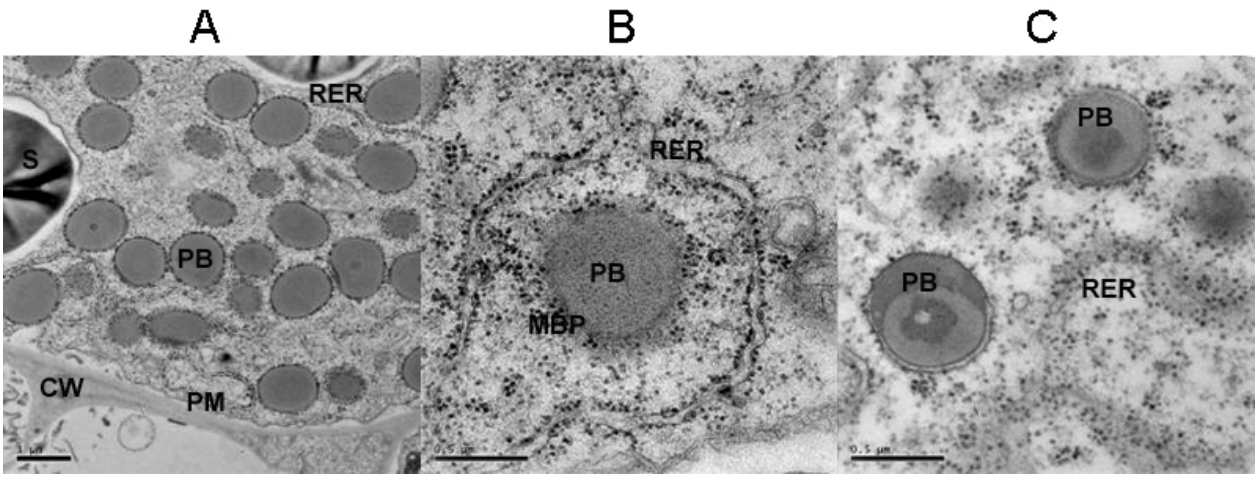

Fig. 5.
Electron micrographs of 12-day *P. glaucum* endosperm cells. Protein bodies (PB), rough endoplasmic reticulum (RER), cell wall (CW), plasma membrane (PM), starch (S), and membrane –bound polyribosomes (MBP) labeled as indicated.

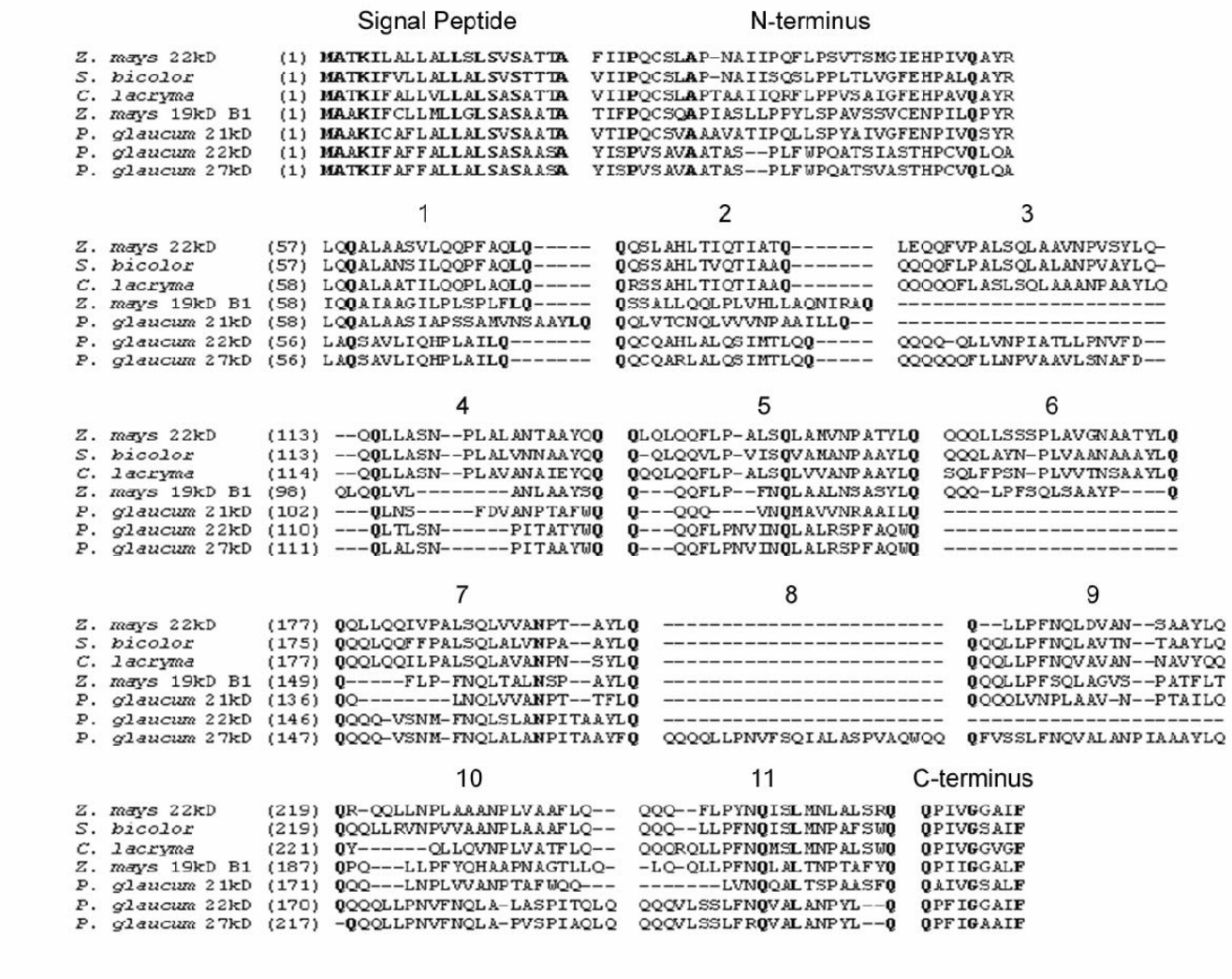

Fig. 6.
Comparison of the predicted amino acid sequence for α-prolamins from maize (*Z. mays*), sorghum (*S. bicolor*), *Coix* (*C. lacryma*), and pearl millet (*P. glaucum*). Signal peptide, N- and C-terminus, and repeat motifs (numbers 1-11 above) forming α -helices are indicated above. Dashes were introduced to maximize alignment, and bold font indicates conserved residues in all sequences.

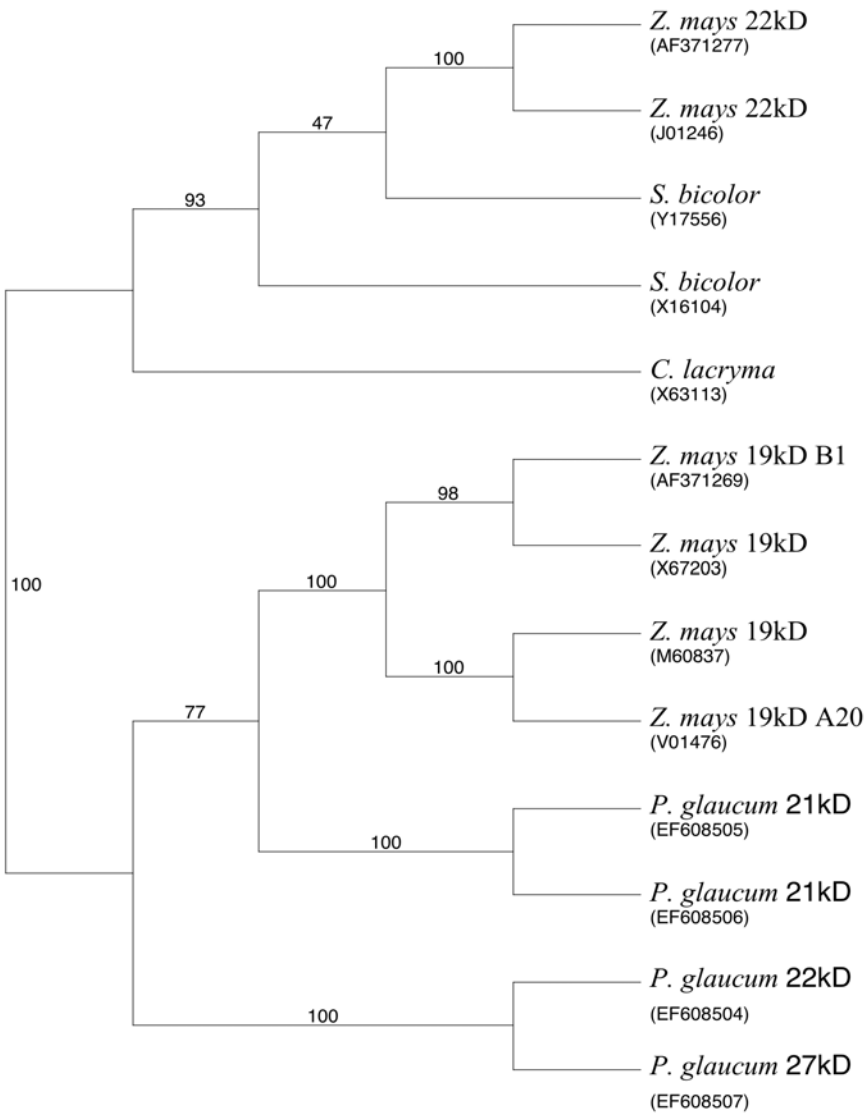

Fig. 7
Bayesian consensus cladogram based on ML reconstruction is shown. Bayesian posterior probabilities of lineal sister relationships are indicated on the branches of the cladogram.