Effect of suppressing the synthesis of different kafirin sub-classes on grain endosperm texture, protein body structure and protein nutritional quality in improved sorghum lines

Laura S. da Silva [a,b], Rudolf Jung [c], Zuo-yo Zhao [c], Kimberly Glassman [c], Andile W. Grootboom [d], Luke Mehlo [d], Maretha M. O'Kennedy [d], Janet Taylor [a], John R.N. Taylor [a,*]

[a] University of Pretoria, Department of Food Science, Private Bag X20, Hatfield 0028, South Africa

[b] Tshwane University of Technology, Department of Biotechnology and Food Technology, Pretoria 0001, South Africa

[c] Pioneer Hi-Bred International Inc, 7300 NW 62nd Avenue, Johnston, IA 50131, USA

[d] CSIR/BioSciences, Meiring Naude Road, Brummeria, Pretoria, 0001, South Africa

*e-mail address: john.taylor@up.ac.za (Prof John Taylor)

Keywords: sorghum, transgenic; protein bodies, digestibility.

Abbreviations used: AAS - amino acid score; ABS - Africa Biofortified Sorghum; EAA – essential amino acid; IVPD - *in vitro* protein digestibility; LKR - lysine ketoglutarate reductase; NC – null control; PDCAAS - protein digestibility corrected amino acid score; TEM - transmission electron microscopy; TG - transgenic.

ABSTRACT

To improve sorghum grain protein nutritional quality, improved sorghum lines were transformed to suppress the synthesis of different kafirin sub-classes, or back-crossed into transgenic lines with improved protein quality. Co-suppression of the alpha-, gamma- and delta-kafirin sub-classes and removal of the tannin trait resulted in transgenic sorghum lines with high cooked protein digestibility (±80%), improved Amino Acid Score (0.8) and Protein Digestibility Corrected Amino Acid Score (0.7) compared to the non-transgenic null controls (±50%, 0.4 and 0.2, respectively). These high protein quality lines had a floury endosperm. They also had modified protein body structure, where the protein bodies were irregular shaped with few to numerous invaginations and were less densely packed, with a dense protein matrix visible around the protein bodies. When fewer sub-classes were suppressed, i.e. gamma 1 and delta 2, the endosperm was corneous with normal protein body structure but the improvement in cooked protein digestibility appeared to be less. Apparently, co-suppression of several kafirin sub-classes is required to obtain high protein nutritional quality sorghum lines, but this seems to result in floury-type grain endosperm texture.

HIGHLIGHTS

Suppression of major kafirins results in transgenic sorghum with high protein quality.

High protein quality transgenic sorghum has a floury phenotype.

High protein quality transgenic sorghum has somewhat irregular shaped protein bodies.

1. Introduction

Sorghum is a major source of protein for people in tropical and subtropical developing countries (FAO, 1995). However, the nutritional quality of sorghum protein is of concern. Sorghum proteins are very deficient in the indispensable (essential) amino acid lysine, due to the kafirin storage proteins being essentially free of lysine (reviewed by Shewry, 2007). Additionally, sorghum proteins have lower cooked protein digestibility compared to other cereals, reducing the bioavailability of the protein (reviewed by Duodu et al., 2003). The reasons for the lower protein digestibility of cooked sorghum are multifactorial, including extensive polymerisation of the kafirins upon cooking, the location and organisation of the different kafirin sub-classes in the protein bodies, and the presence of tannins in certain sorghum lines (reviewed by Duodu et al., 2003).

Efforts to address sorghum protein nutritional quality started with identification of native high-lysine sorghum genotypes from Ethiopia (Singh and Axtell, 1973) and was followed by chemical mutagenesis to develop a high-lysine genotype (P721 opaque) (reviewed by Mertz et al., 1993). The different high-lysine native and mutant genotypes where found to have 50 and 60% increased lysine content, respectively (Singh and Axtell, 1973; reviewed by Mertz et al., 1993). Improved lysine contents were attributed to decreased levels of kafirin proteins and increased levels of lysine–rich, non-kafirin proteins in the grain endosperm (reviewed by Shewry, 2007). However, poor grain quality, especially soft and floury endosperm texture, is common in high-lysine cereals (reviewed by Shewry, 2007), including sorghum (Tesso et al., 2006).

Breeding using P721 opaque line has been undertaken to produce high-lysine genotypes with improved grain hardness and protein digestibility after cooking (Tesso et al., 2006; Weaver et al., 1998). Electron microscopy and immunological studies of these high-

lysine high-protein digestibility mutants showed their grain to have modified protein bodies (irregular shapes with deep invaginations), compared to the spherical protein bodies of normal sorghums (Oria et al., 2000). The location and organisation of different kafirin sub-classes within the protein bodies also differed with the gamma-kafirins being located at the bottom of folds, exposing the more digestible alpha-kafirins to digestive enzymes.

Advances in sorghum tissue culture and transformation research have led to the development of the first nutritionally improved transgenic sorghum (Zhao et al., 2003). Nutritionally improved sorghum lines are being developed using genetic engineer techniques by the "Africa Biofortified Sorghum (ABS) project", under the Bill and Melinda Gates Foundation Grand Challenges in Global Health initiative (Africa Biofortified Sorghum Project, 2009). Early transformation work was using a tannin sorghum line (P890812), with poor end-use quality. This study describes the effect of suppressing different kafirin sub-classes on the grain endosperm texture, protein body structure and protein nutritional quality in different transgenic lines produced by *Agrobacterium*-mediated transformation and through backcrossing into improved normal sorghum lines, with the aim of developing sorghum types of improved protein quality and good functional properties.

2. Materials and Methods

2.1 Grain samples and whole grain flour preparation

Six different transgenic (TG) sorghum lines (plus six non-transgenic null controls) (NCs) developed for the ABS project by Pioneer Hi-Bred using *Agrobacterium*-mediated transformation, as described by Zhao et al (2000). Of these, three were obtained through back-crossing into an normal sorghum variety, Macia (see below). Three different gene constructs (ABS032, ABS166, and ABS149), each designed to suppress the synthesis of different kafirin sub-classes within the grain endosperm, were used, namely: Alpha-kafirin

A1, B1 and B2; gamma-kafirin 1 and 2, and delta-kafirin 2, for ABS032 gene construct (Table 1). Alpha-kafirin A1 and gamma-kafirin 1 for ABS166 gene construct; and delta-kafirin 2 and gamma-kafirin 1 and 2 for ABS149 gene construct. In addition, reduced expression of Lysine Ketogluterate Reductase (LKR) was included in the ABS032 and ABS149 gene constructs. NC are non-transgenic grains obtained from hemizygous transgenic plants. The NCs are the best counterpart to the TGs for transgene performance assays.

Two parent lines (P898012 and Tx430) were used for the different transformations. P898012, is a purple plant, type II tannin (low tannin) sorghum and Tx430, is a non-tannin, tan plant, inbred line. Macia (normal non-tannin line, tan plant improved variety popular in southern Africa), was used for backcrossing with a type II tannin TG line, P898012 with ABS032 gene construct (TG-P89801 (ABS032)), with the aim of breeding out the tannin trait and improving grain endosperm quality. Grain from both greenhouse trials and a summer confined field trial were obtained. All TG grain samples were tested and verified by Pioneer Hi-Bred for kafirin suppression. Methods used included sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot analysis using kafirin sub-class specific antibodies developed by Pioneer Hi-Bred.

The grain samples analysed in this study were in the form of single sectioned kernels (±8 kernels from greenhouse trials) or crushed whole grains (±500g, from a summer confined field trial). All samples were milled into flour using a hand-held mill (IKA A11 Basic, Staufen, Germany) until all flour passed through a 500 µm opening screen. All milled flours were stored at ±8°C until analysis.

2.2 Grain endosperm texture

Endosperm texture, defined as the proportion of corneous endosperm relative to floury endosperm in the grain, was determined subjectively by viewing sectioned kernels (±8

kernels, with or without germ) using a stereomicroscope, and comparing them to sorghum standards. The kernels were classified as corneous, intermediate or floury (ICC, 2008). For crushed grain samples, the meal was sieved through a 1500 µm mesh screen, 8 partially crushed kernels (with or without germ) showing large sections of endosperm were selected and viewed. Light micrographs were taken of all kernel sections using a stereomicroscope (Nikon Optiphot, Tokyo, Japan) fitted with a digital camera (Nikon SMZ800, Tokyo, Japan).

2.3 Chemical characterization

Condensed tannin content was determined on all bulk samples using the modified Vanillin HCl assay (1% conc. HCl in methanol extraction) according to Maxson and Rooney (1972), with subtraction of sample blanks. Catechin hydrate (Sigma, St Louis, MO) was used as a standard, and tannin content was expressed as mg catechin equivalents per 100 g flour (mg CE/100 g). Total protein content (N x 6.25), was determined by Dumas combustion, AACC standard method 46-30 (AACC International, 2000). Total amino acid composition was determined using reverse phase high performance liquid chromatography, Pico-Tag method (Bidlingmeyer et al., 1984). Due to small sample size, 10% moisture content was assumed and all data were expressed on a dry weight basis.

2.4 In vitro protein digestibility

In vitro protein digestibility (IVPD) using pepsin digestion was determined on whole grain flour under raw and wet cooked conditions, using either 200 mg or 20 mg flour scale using the pepsin digestibility method of Hamaker et al. (1986), suitably modified for small-scale assay. In brief, the method involved incubating the flour with pepsin (pepsin from porcine gastric mucosa, 800-2.500 units/mg protein, P7000-100G, Sigma-Aldrich) at pH, 37°C for 2 h. Protein digestibility is defined as the percentage nitrogen solubilized under the

conditions of the assay relative to flour total nitrogen. This was measured in terms of insoluble residue by the above Dumas method.

2.5 Protein nutritional quality

The protein nutritional quality was evaluated by both the Amino Acid Score (AAS) and Protein Digestibility Corrected Amino Acid Score (PDCAAS). The AAS was calculated as g lysine (limiting essential amino acid (EAA))/100 g protein of the sorghum sample / 4.8. Where 4.8 g lysine/100 g protein, is the recommendation for quality protein for 4-18 year olds (WHO/FAO/UNU Expert Consultation, 2007). The PDCAAS was determined by correcting the AAS by multiplying the wet cooked IVPD values obtained.

2.6 Transmission Electron Microscopy

Sections of cleaned peripheral endosperm (1 to 2 mm thick) were fixed in glutaraldehyde in pH 7.4 phosphate buffer (18 h) before staining with osmium tetroxide. Samples were dehydrated sequentially in acetone. Samples were infiltrated with Quetol resin and polymerised at 60°C. Ultra thin sections were stained with uranyl acetate, and Reynold's lead citrate, and viewed with JEOL JEM 2100F field emission electron microscope (Tokyo, Japan). All images shown depict subaleurone layer endosperm cells.

2.7. Statistical analysis

IVPD data were analysed by one-way analysis of variance (ANOVA) using each sample as the independent variable and the measured parameters as the dependent variables. The means were compared by Fisher's least significant differences (LSD). The calculations were performed using Statsgraphics Centurion XV (Stat Point, Herndon, VA).

3. Results and discussion

3.1 Tannin content

Backcrossing the tannin-containing TG-P898012 (ABS032) line into Macia (type I, non-tannin sorghum line) was effective in breeding out the tannin trait as a non-tannin TG line (TG-P898012xMacia (ABS032)-3) was obtained from the summer confined field trial. The non-tannin trait was confirmed by the absence of a pigmented testa layer (Figure 1g) and low levels of phenols and no tannins (Table 2). All other TG lines obtained from backcrossing TG-P898012 (ABS032) with Macia were type II tannin lines, confirmed by the presence of a pigmented testa layer (Figure 1d, e and f, white arrow) and significant levels of tannin (2.34±0.21 mg CE/100 mg flour) was found in TG-P898012xMacia (ABS032)-2 (Table 2). The presence of a pigmented testa in tannin-containing sorghum lines is genetically controlled, requiring both B_1 and B_2 dominant genes (reviewed by Dykes and Rooney, 2006).

3.2 Endosperm Texture

Visual examination of kernels from the different TG lines revealed considerable variation in endosperm texture modification compared to their respective NCs and the parent lines. This ranged from completely modified for TG-P898012 (ABS032) (Figure 1d) and backcrosses (TG-P898012xMacia (ABS032) -1, -2 and -3, Figure 1e, f and g), to little (or no) modification for TG-Tx430 (ABS149) (Figure 1m). P898012, TG-P898012 (ABS032) and some of the backcrosses showed a distinct lumen (small hole) in the centre of the grain (Figure 1, black arrows). The modified endosperm phenotype of TG-P898012 (ABS032) and backcrosses appears to be a direct consequence of the co-suppression of synthesis of several kafirin sub-classes (namely alpha-kafirin A1, B1 and B2; gamma-kafirin 1 and 2 and delta-kafirin 2, Table 1) within the endosperm. As stated, in sorghum, certain nutritional quality

traits such as high essential amino acid content and improved protein digestibility tend to be associated with soft endosperm (Tesso et al., 2006). Similarly, in high-lysine mutant maize (*opaque*-2 and *floury*-2), soft, starchy endosperm texture has been observed (reviewed by Shewry, 2007). For *floury*-2 mutant, the mutation resulted in a decrease in the synthesis of all sub-classes of zeins, modified endosperm texture and the zein protein bodies being smaller than normal and asymmetrical or misshapen (Lending and Larkins, 1992). Also, the native Ethiopian high-lysine sorghum landrace identified among the world germplasm collection are found to have soft endosperms (Singh and Axtell, 1973).

In contrast the above types, TG-Tx430 (ABS166) kernels had a modified endosperm texture with a large central floury portion with faint bands or patches of corneous-like endosperm (Figure 1l, black dashed arrow), unlike the corneous endosperm texture of the NC and normal parent line Tx430 (Figure 1n and c, respectively). It is likely that suppression of synthesis of only alpha-kafirin A1 and gamma-kafirin 1 (Table 1) did not disrupt protein body synthesis substantially. Hence, the co-expression of alpha-kafirin B1 and B2, gamma-kafirin 2 and beta-kafirin and delta-kafirin sub-classes, may have been sufficient to ensure normal protein body synthesis, and retain some of the corneous phenotype of the parent line Tx430. Studies with transgenic tobacco plants in which genes encoding one or more zein sub-classes indicates that beta-zein or gamma-zein must be co-expressed with alpha-zein (and delta-zein), to promote alpha-zein stability and retention in the endoplasmic reticulum, in order to form normal protein bodies (Coleman et al., 1996; Coleman et al., 2004). Further, the hard endosperm texture of TG-Tx430 (ABS149) suggests that the co-suppression of delta-kafirin 2 and gamma-kafirin 1 and 2 sub-classes synthesis alone, did not have a dramatic effect on the endosperm phenotype.

3.3 Protein content and amino acid profile

The different TG lines had protein contents ranging from 10.6% (TG-P898012xMacia (ABS032)-1) to 13.7% (TG-Tx430 (ABS149)) (Table 2). The TG lines and their respective null controls were all generally very similar or even identical in protein content. The protein contents reported here are well within the range for native Ethiopian high-lysine sorghums (10.0-17.2%) (Guiragossian et al., 1978; Singh and Axtell, 1973), high protein digestibility-high lysine mutant sorghums (10.2-14.7%) (Tesso et al., 2006; Weaver et al., 1998) and normal sorghums (8.1-16.8%) (reviewed by Rooney and Serna-Saldivar, 1990).

In contrast, amino acid profile differed substantially between the different TG lines, specifically with regard to lysine content (Table 3). TG-P898012 (ABS032) and the backcrosses (TG-P898012xMacia (ABS032)-1, -2 and -3) had the highest lysine content, ranging from 3.65 to 4.10 g lysine/100 g protein, compared to TG-Tx430 (ABS166), TG-Tx430 (ABS149) and the parent lines, which had lysine contents of 2.18, 2.43 and 2.08 g lysine/100 g protein, respectively, within normal ranges for sorghum (reviewed by Rooney and Serna-Saldivar, 1990). The amino acid profiles including lysine contents of the NCs were essentially the same as the parent lines (data not shown).

The lysine content of TG-P898012 (ABS032) and the backcrosses were slightly higher than that reported for native Ethiopian high-lysine genotypes (3.2-3.7 g/100 g protein) (Guiragossian et al., 1978; Singh and Axtell, 1973), and high protein digestibility-high lysine mutant sorghums (2.9-3.4 g/100 g protein) (Tesso et al., 2006; Weaver et al., 1998). But lysine values were in the range reported for quality protein maize (QPM) cultivars (3.43-4.56 g lysine/100 g protein) (Zarkadas et al., 2000). The high lysine TG lines also showed considerable reduction in proline (24% lower), alanine (17%), glutamic acid/glutamine (15%), leucine (15%) and phenylalanine (15%), and increases in arginine (76% higher), aspartic acid (48%), methionine (42%) and glycine (32%) compared to the normal sorghums (Table 3). Similar changes in the amino acid profile of high-lysine cereals have been reported

(Guiragossian et al., 1978; Singh and Axtell, 1973; Vendemiatti et al., 2008; Zarkadas et al., 2000). The altered amino acid profile observed for TG-P898012 (ABS032) and the backcrosses is probably a direct consequence of co-suppression of the synthesis of the major kafirin sub-classes. Because the kafirins contain essentially no lysine, these results indicate a substantial increase in the proportion of non-kafirin storage proteins and non-storage proteins in the grain, which are higher in lysine content. In other high-lysine sorghum genotypes, the proportion of lysine rich non-kafirin storage proteins (albumins, globulins and glutelins) was considerably higher compared to normal sorghums (Guiragossian et al., 1978; Vendemiatti et al., 2008). While in high-lysine *opaque*-2 mutant maize genotypes, the increase in lysine content is attributed to elongation factor-1α (EF-1α), a lysine rich (11% lysine) non-zein protein (Habben et al., 1995). Other non-zein proteins found to be over expressed in *opaque-2* mutants, include catalase-2 (7% lysine) and trypsin inhibitor (1% lysine).

The reduced suppression of LKR in TG lines with the ABS032 gene construct, may also have contributed to the increased lysine content. This is similar to the situation in *opaque-2* maize, where the activity of lysine LKR is reduced, resulting in increased levels of free lysine (reviewed by Gibbons and Larkins, 2005).

3.4 In Vitro Protein Digestibility

Different TG lines gave a wide range of raw, 61.5 % (TG-TX430 (ABS149) and 91.4% (TG-P898012xMacia (ABS032)-3) and wet cooked, 41.1% (TG-Tx430 (ABS149)) and 79.8% (TG-P898012xMacia (ABS032)-3) protein digestibilities. In all cases, the transgenic sorghums had substantially higher protein digestibility than their respective null controls. The differences observed IVPD were due to different kafirin sub-classes being suppressed, as well as the presence or absence of tannins within the grains. Generally, TG-P898012xMacia (ABS032)-3 the non-tannin TG line in which the major kafirin sub-classes was suppressed

showed substantially higher raw (91.4%) and wet cooked protein digestibilities (79.8%) compared to the tannin-containing TG lines, with similar kafirin suppression, ranging from 69.4 to 80.2%, for raw and 50.0 to 58.3% for cooked flours, respectively (Table 2). According to Taylor et al. (2007) kafirin proteins, specifically the gamma-kafirin, bind considerable quantities of tannins, ranging from 35 to 77%, depending on kafirin composition, forming very large molecular weight (>200k) aggregates of kafirin polymers and tannin molecules with reduced protein digestibility compared to unbound kafirin. In TG lines where co-suppression of fewer kafirin sub-classes occurred (TG-Tx430 (ABS166) and TG-Tx430 (ABS149)) lower protein digestibility improvement seemed to occur, even in the absence of tannins (Table 2). However, it should be noted that all three different constructs, ABS 032, 149 and 166 gave improvement in protein digestibility in the TGs when compared to their NCs.

Wet cooking reduced the protein digestibility of all sorghum lines. However, the reduction in digestibility was notably less in the TG lines compared to the parent lines and NCs (Table 2). The non-tannin line, TG-P898012xMacia (ABS032)-3, showed the least reduction in protein digestibility with cooking, only 12%, while the parent and NCs showed the highest reduction, at least 40%. A possible explanation is that suppression of synthesis of the major kafirins sub-classes, as in the case of TG lines with ABS032 gene construct or cysteine-rich gamma-kafirins in the case of TG lines with ABS166, and ABS149 gene construct, could result in lower levels of crosslinked polymers in the TG lines, compared to the NCs and parent lines.

Within the TG events there was a wide range of improvement in protein digestibility. For example, for TG-P898012xMacia (ABS032)-1 the raw digestibility ranged from 61.3 to 82.8% and the cooked digestibility ranged from 43.0 to 59.2% (Table 2). The reason for this was that different TG events from the same vector usually give different levels of transgene

expression and plant performance, due to factors such as transgene insertion site, pattern and copy number. Therefore the best potential TG event was ABS032-3 (TG-P898012xMacia(ABS032-3)) with a cooked protein digestibility of up to 80.8%. This compares to its respective NC (NC-P898012xMacia-3), where the highest cooked protein digestibility was 57.6% and the best normal sorghum, Macia with a cooked protein digestibility of 59.2%.

3.5 Protein Nutritional Quality

AAS and the PDCAAS varied considerably between the different TG lines, as a result of the differences in both the lysine contents (Table 3) and cooked IVPDs (Table 2). Generally, the TG lines with co-suppression of the major kafirin sub-classes (ABS032 gene construct) had significantly higher AAS (0.68-0.80) and PDCAAS (0.38-0.70) compared to TG lines where co-suppression of fewer kafirin sub-classes (0.45 and 0.28; TG-Tx430 (ABS166) and 0.51 and 0.21; TG-Tx430 (ABS149), respectively) and the parent lines (0.41-0.45 and 0.16-0.24, respectively) (Table 3). However, the presence of tannins in some of the TG lines (TG-P898012 (ABS032) and TG-P898012xMacia (ABS032)-1 and -2) reduced the PDCAAS, by at least 40%, because the IVPD of the tannin-containing TG lines was considerably lower (Table 3).

3.6 Endosperm ultrastructure

The peripheral endosperm texture and protein body structure of the different transgenic sorghum lines showed variable modification compared to the parent lines (Figure 2) and respective NCs (data not shown). Protein bodies of parent lines (P898012, Macia and TX430, Figure 3a, b and c) were typical, with tightly packed, round protein bodies, ±2 µm in diameter, with internal concentric ring structures (Figure 2a, black arrow) as described for

normal sorghum protein bodies (Adams et al., 1976). The protein bodies of all the NCs were typical, and were similar to the parent lines (data not shown). In normal sorghum protein bodies, highly crosslinked kafirin proteins (gamma- and beta-kafirins) are found at the protein body periphery and seen as dark-staining inclusions in the form of concentric rings within the protein body interior (Oria et al., 1995; Shull et al., 1992).

TG-P898012 (ABS032) (Figure 2d-e) and the backcrosses (TG-P898012xMacia (ABS032)-1, -2 and -3 (Figure 2f-g, h-i, j-k, respectively) showed substantial modification in peripheral endosperm texture and protein body structure. The protein bodies were sparsely packed, 2 µm in diameter, and were often surrounded by a dense continuous dark protein matrix (Figure 2d-k, white dashed arrow). Protein body margins were slightly folded (invaginated) (Figure 2d-k, black dashed arrow) with a proportion of the protein bodies showing occasional irregular, thick dark-staining inclusions (Figure 2e, i and k, white arrows). The characteristic internal concentric ring structure of normal protein bodies was absent. However, a number of atypical concentric rings were observed around the protein body periphery of TG-P898012xMacia (ABS032)-1 (Figure 2f-g, black arrows). The modified protein body structure of the TG lines with the ABS032 gene construct is dissimilar to that of high-lysine, high-protein digestibility mutants, where the protein bodies are described as being highly invaginated (with deep folds) (Oria et al.; 2000).

It appears that the co-suppression of synthesis of the major kafirin sub-classes in TG lines with ABS032 gene construct had a major effect on the peripheral endosperm texture and protein body structure, which may in part be responsible for the improved cooked IVPD of these TG lines. As explained, the improved protein digestibility of mutant sorghum lines, having highly invaginated protein bodies, is believed to be due to increased protein body surface area and easy accessibility of digestive enzymes to the more digestible alpha-kafirin proteins (Oria et al., 2000). In addition, the dense protein matrix observed in the TG lines, is

probably composed of more digestible lysine-rich endogenous proteins, which would further improve the cooked IVPD of these lines.

In contrast, TG-Tx430 (ABS166) and TG-Tx430 (ABS149) (Figure 2l-m and n-o, respectively); had protein bodies with structure more typical of the parent line Tx430 (Figure 2c). Protein bodies of TG-Tx430 (ABS166) were generally round, ±2µm in diameter, and the internal concentric ring structure was also present in some of the protein bodies (Figure 2l-m, black arrow). However, there were considerable levels of dark staining inclusions around the periphery of the protein bodies, giving the appearance of slight invagination of the periphery (Figure 2m, white arrow). The protein body packing density of TG-Tx430 (ABS166) also appeared to be less dense than that of the parent and patches of dark protein matrix were observed between some of the protein bodies (Figure 2l-m, white dashed arrow). The protein bodies of TG-Tx430 (ABS149) did not appear to be modified, showing typical protein body structure, packing density and size (±2 µm) (Figure 2n-o), as that of the parent Tx430 (Figure 2c). However, dark staining inclusions around the protein body periphery would indicate that these protein bodies are also slightly invaginated (Figure 2n-o, white arrow). The observed endosperm ultrastructure of the different TG Tx430 lines was not surprising as these TG lines had IVPD and protein nutritional quality within normal ranges. This is probably due to differences in the levels and types of cysteine-rich kafirin, available to form crosslinked polymers.

According to a number of studies conducted on the interaction of different zein proteins to form storage protein bodies, stable accumulation and aggregation of alpha-zein into protein bodies requires the interaction of either gamma-zein (Coleman et al., 1996) or beta-zein (Coleman et al, 2004). This is due to gamma- and beta-zein being structurally related, having similar roles in the initiation of protein bodies in the developing maize endosperm (Kim et al., 2002). Both gamma- and beta-zeins are rich in cysteine resulting in their aggregation via

intermolecular disulphide bonds to form an insoluble protein body core, required for the interaction and integration of alpha-zein (Coleman et al., 2004).

4. Conclusions

Co-suppression of the major kafirin sub-classes, alpha-, gamma- and delta, results in sorghum lines with substantially improved cooked protein digestibility, improved Amino Acid Score and hence greatly improved Protein Digestibility Corrected Amino Acid Score. The substantially improved protein digestibility appears to be associated with floury endosperm texture. In turn, this seems to be related to modified protein body structure due to the suppression of kafirin synthesis, When fewer kafirin sub-classes are suppressed, i.e. gamma 1 and delta 2, the endosperm is corneous, with apparently normal protein body structure but the improvement in cooked protein digestibility seems to be less.

Acknowledgements

This work was supported by the Bill and Melinda Gates Grand Challenges 9, Africa Biofortified Sorghum (ABS) Project. Additional financial support was obtained from the South African National Research Foundation, and the Faculty Research and Innovation Committee, Faculty of Science, Tshwane University of Technology. Mr Alan Hall and Mr Chris van der Merwe, Laboratory for Microscopy and Microanalysis, University of Pretoria for microscopy training and guidance.

5. References

AACC International, 2000. Crude protein-combustion, Standard Method 46-30, Approved Methods of the AACC, tenth ed. The Association, St Paul, MN.

Adams, C.A., Novellie, L., Liebenberg, N.v.d.W. 1976. Biochemical properties and ultrastructure of protein bodies isolated from selected cereals. Cereal Chemistry 53, 1-12.

ABS (African Biofortified Sorghum Project). 2009. http://biosorghum.org, accessed 12 January 2009

Bidlingmeyer, B.A., Cohen, S.A., Tarvin, T.L. 1984. Rapid analysis of amino acids using pre-column derivatization. Journal of Chromatography 366, 93-104.

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

Coleman, C.E., Yoho, P.R., Escobar, S., Ogawa, M. 2004. The accumulation of α-zein in transgenic tobacco endosperm is stabilized by co-expression of β-zein. Plant Cell Physiology 45(7), 864-871.

Duodu, K.G., Taylor, J.R.N., Belton, P.S. Hamaker, B.R. 2003. Factors affecting sorghum protein digestibility. Journal of Cereal Science 38, 117-131.

Dykes, L., Rooney, L.W. 2006. Sorghum and millets phenols and antioxidants. Journal of Cereal Science 44, 236-251.

FAO (Food and Agricultural Organisation). 1995. Sorghum and Millets in Human Nutrition. FAO Food and Nutrition Series No. 27, Rome: Food and Agricultrual Organisation of the United Nations.

Gibbon, B.C, Larkins, B.A. 2005. Molecular genetic approaches to developing quality protein maize. Trends in Genetics 21(4), 227-233.

Guiragossian, V., Chibber, B.A.K., Van Scoyoc, S. Jambunathan, R. Mertz, E.T., Axtell, J.D. 1978. Characteristics of proteins from normal, high lysine, and high tannin sorghums. Journal of Agricultural and Food Chemistry 26, 219-223.

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

Hamaker, B.R., Kirleis, A.W., Butler, L.G., Axtell, J.D., Mertz, E.T. 1987. Improving the *in vitro* protein digestibility of sorghum with reducing agents. Proceedings of the National Academy of Sciences of the United States of America 84, 626-628.

ICC (International Association for Cereal Science and Technology). 2008. Estimation of sorghum grain endosperm texture. ICC Standard 176. ICC, Vienna.

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

Lending C.R., Larkins, B.A. 1992. Effect of the *floury-2* locus on protein body formation during maize endosperm development. Protoplasma 171, 123-133.

Maxson, E.D., Rooney, L.W. 1972. Evaluation of methods for tannin analysis in sorghum grain. Cereal Chemistry 49, 719-729.

Mertz, E.T., Axtell, J.D., Ejeta, G., Hamaker, B.R. 1993. Development and recent impact of quality protein maize and sorghum. In: Taylor, J.R.N, Randall, P.G. and Viljoen. J.H.

(eds) Cereal Science and Technology. Impact on a changing Africa. Selected papers from the ICC international symposium. The CSIR, Pretoria, South Africa. Pp 115-131

O'Kennedy, M.M., Grootboom, A., Shewry, P.R. 2006. Harnessing sorghum and millet biotechnology for food and health. Journal of Cereal Science 44, 224-235.

Oria. M.P., Hamaker., B.R., Axtell, J.D., Huang, C.P. 2000. A highly digestible sorghum mutant cultivar exhibits a unique folded structure of endosperm protein bodies. Proceedings of the National Academy of Sciences of the United States of America 97, 5065-5070.

Oria, M.P., Hamaker, B.R., Shull, J.M. 1995. In vitro protein digestibility of developing and mature sorghum grain in relation to α-, β-, and γ-kafirin disulfide crosslinking. Journal of Cereal Science 22, 85-93.

Rooney, L.W., Serna-Saldivar, S.O. 1990. Sorghum. In: Lorenz, K.J., Kulp, K. (Eds.). Handbook of Cereal Science and Technology, Markel Dekker, New York. Chapter V, pp. 233-270.

Shewry, P.R. 2007. Improving the protein content and composition of cereal grain. Journal of Cereal Science 46:239-250.

Shull, J.M., Watterson, J.J., Kirleis, A.W. 1992. Purification and immunocytochemical location of kafirins in *Sorghum bicolor* (L.) Moench endosperm. Protoplasma 171, 64-74.

Singh, R., Axtell., L.D. 1973. High lysine mutant gene (hl) that improves protein quality and biochemical value of grain sorghum. Crop Science 13, 535-539.

Tesso T., Ejeta, G., Chandrashekar, A., Huang, C.P., Tandjung, A., Lewamy, M., Axtell, J.D., Hamaker, B.R. 2006. A novel modified endosperm texture in a mutant high-protein digestibility/high-lysine grain sorghum (*Sorghum bicolor* (L.) Moench). Cereal Chemistry 83, 194-201.

Taylor, J., Bean, S.B., Ioerger, B.P., Taylor, J.R.N. 2007. Preferential binding of sorghum tannins with γ-kafirin and the influence of tannin binding on kafirin digestibility and biodegradation. Journal of Cereal Science 46, 22-31.

Vendemiatti, A., Ferreira, R.R., Gomes, L.H., Medici, L.O., Azevedo, R.A. 2008. Nutritional quality of sorghum seeds: storage proteins and amino acids. Food Biotechnology 22, 377-397.

Weaver C.A., Hamaker, B.R., Axtell, J.D. 1998. Discovery of grain sorghum germplasm with high uncooked and cooked *in vitro* protein digestibilities. Cereal Chemistry 75, 665-670.

WHO/FAO/UNU Expert Consultation. 2007. Protein and Amino Acid Requirements in Human Nutrition, Report of a Joint WHO/FAO/UNU Expert Consultation. World Health Organization Technical Report No. 935. World Health Organization, Geneva.

Zarkadas, C.G., Hamilton, R.I., Yu, Z.R., Choi, V.K., Khanizadeh, S., Rose, N.G.W., Pattison, P.L. 2000. Assessment of Protein Quality of 15 new northern adapted cultivars of quality protein maize using amino acid analysis. Journal of Agricultural and Food Chemistry 48, 5351-5361.

Zhao, Z-Y., Cai, T., Tagliani, L., Miller, M., Wang, N., Pang, H., Rudert, M., Schroeder, S., Hondred, D., Seltzer, J., Pierce, D. 2000. *Agrobacterium*-mediated sorghum transformation. Plant Molecular Biology 44, 798-798.

Zhao, Z-Y., Glassman, K., Sewalt, V., Wang, N., Miller, M., Chang, S., Thompson, T., Catron, S., Wu, E., Bidney, D., Kedebe, Y., Jung, R. 2003. Nutritionally improved transgenic sorghum. In: Vasil, I.K (Ed). Plant Biotechnology 2002 and Beyond. Kluwer Academic Publishers, Dordrecht, pp. 413-416.

Legends to figures.

Figure 1. Sectioned sorghum kernels (with or without germ).

Parent lines: (a) P898012, (b) Macia, (c) Tx430 used for transformation or backcrossing. Transgenic lines: (d) TG-P898012 (ABS032), (e) TG-P898012xMacia (ABS032)-1, (f) TG-P898012xMacia (ABS032)-2, (g) TG-P898012xMacia (ABS032)-3 and their respective null controls (h, i, j and k). Transgenic lines: (l) TG-Tx430 (ABS166)), (m) TG-Tx430 (ABS149), (n, o) their respective null controls.

Black arrow indicates lumen in centre of endosperm, white arrow indicates pigmented testa layer, black dashed arrow indicates patches of vitreous endosperm. Bar = 2mm.

Figure 2. TEM of protein bodies in the peripheral endosperm of different sorghum lines.

Parent lines: (a) P898012, (b) Macia, (c) Tx430 used for transformation or backcrossing. Transgenic lines: (d, e) TG-P898012 (ABS032), (f, g) TG-P898012xMacia (ABS032)-1, (h, i); TG-P898012xMacia (ABS032)-2, (j, k) TG-P898012xMacia (ABS032)-3, (l, m) TG-Tx430 (ABS166), (n, o) TG-Tx430 (ABS149).

C, cell wall, P, protein body, S, starch granule. Black arrow indicates concentric ring structure, black dashed arrow indicates irregular shaped protein body, black arrow head indicates atypical concentric ring structure, white arrow indicates dark staining inclusions, white dashed arrow indicates dark protein matrix.

Table 1. Transgenic sorghum lines, null controls (no kafirin suppression detected) and normal lines studied.

Line	Sample code	n	Supplied by
Macia	Macia	1	BTS, 2004
P898012	P898012 Bulk	1	PHB, 2007
P898012	P898012 H/CG	1	PHB, 2009
Tx430	Tx430	1	PHB, 2009*
Transgenic P898012 (ABS032 gene construct), greenhouse trial, T1 seed, alpha-, gamma- and delta-kafirin suppression.	TG-P898012(ABS032)	1	PHB, 2007
P898012, T1 seed, no kafirin suppression.	NC-P898012	1	PHB, 2007
Transgenic P898012 (ABS032 gene construct) backcrossed into Macia, greenhouse trial, F3 seed, alpha-, gamma- and delta-kafirin suppression..	TG-P898012xMacia (ABS032)-1	4	PHB, 2009
P898012 backcrossed into Macia, greenhouse trial, F3 seed, no kafirin suppression.	NC-P898012xMacia-1	2	PHB, 2009
Transgenic P898012 (ABS032 gene construct) backcrossed into Macia, summer confined field trial, F3 seed, alpha-, gamma- and delta-kafirin suppression.	TG-P898012xMacia (ABS032)-2	3	PHB, 2009*
P898012 backcrossed into Macia, summer confined field trial, summer confined field trial, F3 seed, no kafirin suppression.	NC-P898012xMacia-2	1	PHB, 2009*
Transgenic P898012 (ABS032 gene construct) backcrossed into Macia, summer confined field trial, F3 seed, alpha-, gamma- and delta-kafirin suppression.	TG-P898012xMacia (ABS032)-3	3	PHB, 2009*
P898012 backcrossed into Macia, summer confined field trial, F3 seed, no kafirin suppression.	NC-P898012xMacia-3[3]	2	PHB, 2009*
Transgenic Tx430 (ABS166 gene construct), greenhouse trial, T1 seed, alpha- and gamma-kafirin suppression.	TG-Tx430 (ABS166)	6	PHB, 2009
Tx430, greenhouse trial, T1 seed, no kafirin suppression	NC-Tx430-1	6	PHB, 2009
Transgenic Tx430 (ABS149 gene construct), greenhouse trial, T1 seed, delta- and gamma-kafirin suppression.	TG-Tx430 (ABS149)	2	PHB, 2009
Tx430, greenhouse trial, T1 seed, no kafirin suppression.	NC-Tx430-2	2	PHB, 2009

BTS – Botswana (University of Pretoria, sorghum collection). PHB -Pioneer Hi-Bred, Des Moines, Iowa. TG- transgenic grain, NC - Non-transgenic null control. n= number of samples or transgenic events received, *samples receive as crushed grain (±500 g). All others received as sectioned kernels (±8 kernels).

Table 2. Presence of pigmented testa, tannin and protein content, and *in vitro* protein digestibility (IVPD) of raw and cooked whole grain flour for different transgenic (TG, with gene constructs in bold) sorghum lines compared to non-trangenic null controls (NC) and normal sorghum lines.

Line	n	Pigmented testa (Yes/No)	Tannin content (mg CE/100 mg flour)	Protein (g/100 g flour)	IVPD Raw (%)	IVPD Cooked (%)
Macia	1	No	0.02^{a}±0.02	10.1±0.7	75.2^{ghi}±1.6	59.2^{fg}±0.7
P898012 Bulk	1	Yes	3.26^{c}±0.12	10.9±0.5	41.1^{bc}±2.0	25.7^{b}±1.5
P898012 H/CG	1	Yes	ND	13.9±1.2	48.9^{cd}±0.3	30.7^{bc}±6.2
Tx430	1	No	ND	10.2±0.2	65.0^{ef}±0.8	48.2^{e}±0.3
TG-P898012**(ABS032)**[1]	1	Yes	ND	12.8±0.2	ND	73.7^{h}±2.5
NC-P898012	1	Yes	ND	10.1±0.2	32.3^{a}±1.7	22.2^{a}±3.8
TG-P898012xMacia**(ABS032)**-1[1]	4	Yes	ND	10.6±1.4 (8.8-12.4)	69.4^{fg} ±8.6 (61.3-82.8)	50.0^{e}±6.2 (43.0-59.2)
NC-P898012xMacia-1[1]	2	Yes	ND	9.0±0.2 (8.9-9.1)	47.3^{cd}±5.4 (42.8-51.8)	28.7^{b}±1.4 (27.8-29.5)
TG-P898012xMacia**(ABS032)-2**[2]	3	Yes	2.34^{b}±0.21 (1.90-2.73)	11.8±0.7 (11.3-12.8)	80.2^{i}±3.1 (77.2-82.8)	58.3^{fg}±3.0 (54.1-61.0)
NC-P898012xMacia-2[2]	1	Yes	2.65^{b}±0.02	11.8±0.2	68.0^{efgh}±0.1	34.6^{bcd}±0.1
TG-P898012xMacia**(ABS032)-3**[3]	3	No	0.02^{a}±0.01 (0.02-0.03)	12.4±0.3 (12.1-12.8)	91.4^{j}±1.8 (90.4-93.0)	79.8^{i}±1.4 (78.4-80.8)
NC-P898012xMacia-3[3]	2	No	0.02^{a}±0.01	12.3±0.2 (12.4-12.4)	75.5^{hi}±2.2 (73.8-77.2)	56.4^{f}±1.9 (55.1-57.6)
TG-Tx430**(ABS166)**[1]	6	No	ND	12.6±2.3 (8.9-15.5)	77.9^{i}±7.3 (66.0-85.2)	61.3^{g}±7.3 (50.9-71.6)
NC-Tx430-1[1]	6	No	ND	12.1±2.5 (8.5-15.5)	49.9^{d}±6.9 (43.5-63.2)	40.0^{d}±4.9 (33.6-48.7)
TG-Tx430**(ABS149)**[1]	2	No	ND	13.7±2.2 (11.7-15.6)	61.5^{e}±2.3 (60.0-63.0)	41.1^{d}±0.2 (40.9-41.2)
NC-Tx430-2[1]	2	No	ND	13.6±2.0 (11.8-15.5)	40.4^{b}±2.1 (39.7-41.0)	34.4^{c}±1.3 (33.8-35.0)

Values are means ±standard deviations, values in parentheses are the range.

Values of a parameter in the same column with different superscript letters were significantly different ($p \leq 0.05$).

n= number of samples analysed, samples were analysed in duplicate and the analysis was repeated at least once.

H/CG= Half crushed grain.

[1]Greenhouse trial. [2]Summer confined field trial, tannin type. [3]Summer confined field trial, non-tannin type. ND= Not determined.

Table 3. Amino acid composition (g/100 g protein), recovery, Amino Acid Score (AAS) and Protein Digestibility Corrected Amino Acid Score (PDCAAS) for different transgenic (TG) and normal sorghum lines.

	Transgenic lines with different ABS gene constructs (in bold)						Parent lines used for transformation or backcrossing		
Amino Acid	TG-P898012 **(ABS032)**[1]	TG-P898012x Macia**(ABS032)** -1[1]	TG-P898012x Macia**(ABS032)** -2[2, 4]	TG-P898012x Macia**(ABS032)** -3[3, 4]	TG-Tx430 (**ABS166**)[1]	TG-Tx430 (**ABS149**)[1]	P898012	Macia	Tx430
Non-EAA									
Glu	15.72±0.76	16.26±1.18	-	-	21.12±1.58	20.91±1.82	19.34±0.29	16.95±0.67	21.32±0.47
Asp	6.74±0.00	7.43±0.78	-	-	6.60±0.54	5.14±0.40	5.24±0.13	4.99±0.13	4.54±0.35
Ala	7.46±0.25	6.93±0.50	-	-	7.86±0.53	9.20±1.03	8.77±0.23	7.77±0.20	9.08±0.03
Pro	6.29±0.13	6.15±0.46	-	-	5.94±0.40	7.70±0.59	8.26±0.23	7.58±0.20	8.70±0.24
Arg	5.75±0.25	6.21±0.96	-	-	3.91±0.42	4.35±0.04	3.46±0.00	2.87±0.33	4.06±0.47
Ser	3.73±0.19	3.76±0.33	-	-	3.60±0.28	4.35±0.36	4.16±0.16	3.63±0.07	4.46±0.10
Gly	3.68±0.13	3.71±0.43	-	-	2.84±0.38	3.07±0.03	2.85±0.05	2.50±0.07	3.09±0.16
His*	1.71±0.13	1.80±0.24	-	-	2.38±0.32	1.76±0.13	2.10±0.05	1.55±0.07	2.21±0.19
EAA									
Ile	3.86±0.13	3.50±0.18	-	-	3.46±0.26	4.00±0.27	3.95±0.05	3.39±0.13	3.90±0.02
Leu	10.74±0.32	13.11±0.94	-	-	10.74±0.73	18.41±2.05	13.70±0.19	12.10±0.60	18.60±0.25
Met	2.20±0.19	2.38±0.22	-	-	1.57±0.21	2.14±0.07	1.61±0.14	1.51±0.13	1.82±0.03
Cys	0.00±0.00	0.00±0.00	-	-	0.00±0.00	0.00±0.00	0.09±0.11	0.00±0.00	0.04±0.05
Phe	4.72±0.32	3.86±0.27	-	-	4.18±0.30	4.82±0.35	5.29±0.05	4.24±0.27	4.70±0.03
Tyr	3.59±0.02	3.68±0.29	-	-	3.63±0.26	4.41±0.38	3.81±0.09	3.58±0.40	4.33±0.13
Thr	2.61±0.76	2.52±0.33	-	-	2.89±0.29	2.50±0.17	2.92±0.12	2.82±0.13	2.46±0.08
Val	5.21±0.13	4.92±0.39	-	-	4.48±0.35	4.93±0.24	4.96±0.08	4.47±0.20	4.95±0.07
Lys	3.28±0.06	3.65±0.47	4.1±0.2	4.1±0.5	2.18±0.37	2.43±0.07	2.10±0.05	1.95±0.20	2.18±0.47
Recovery	87.3	89.9			87.4	100.0	81.9	92.6	100.0
AAS	0.68	0.76	0.8	0.8	0.45	0.51	0.41	0.44	0.45
PDCAAS	0.50	0.38	0.5	0.7	0.28	0.21	0.24	0.21	0.16

Values are means ±standard deviations. - = no data available. [1]Greenhouse trial. [2]Summer confined field trial, tannin type. [3]Summer confined field trial, non-tannin type. [4]Lysine data supplied by Pioneer Hi-Bred, 2009. EAA- Essential amino acids. His*- EAA for infants. AAS-((g lys/100 g protein)/4.8), 4.8 is WHO recommendation for lysine quality protein for 4-18 year olds, and PDCAAS- (AAS x Cooked IVPD) (WHO/FAO/UNU Expert Consultation, 2007).

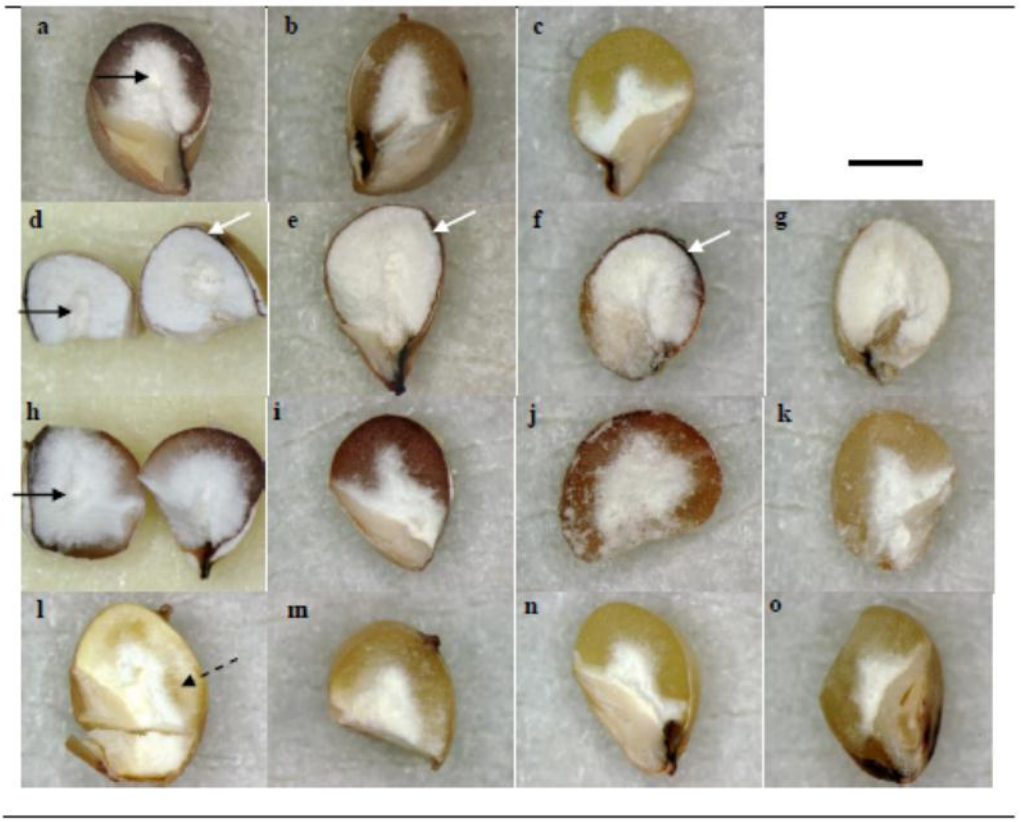

Figure 1

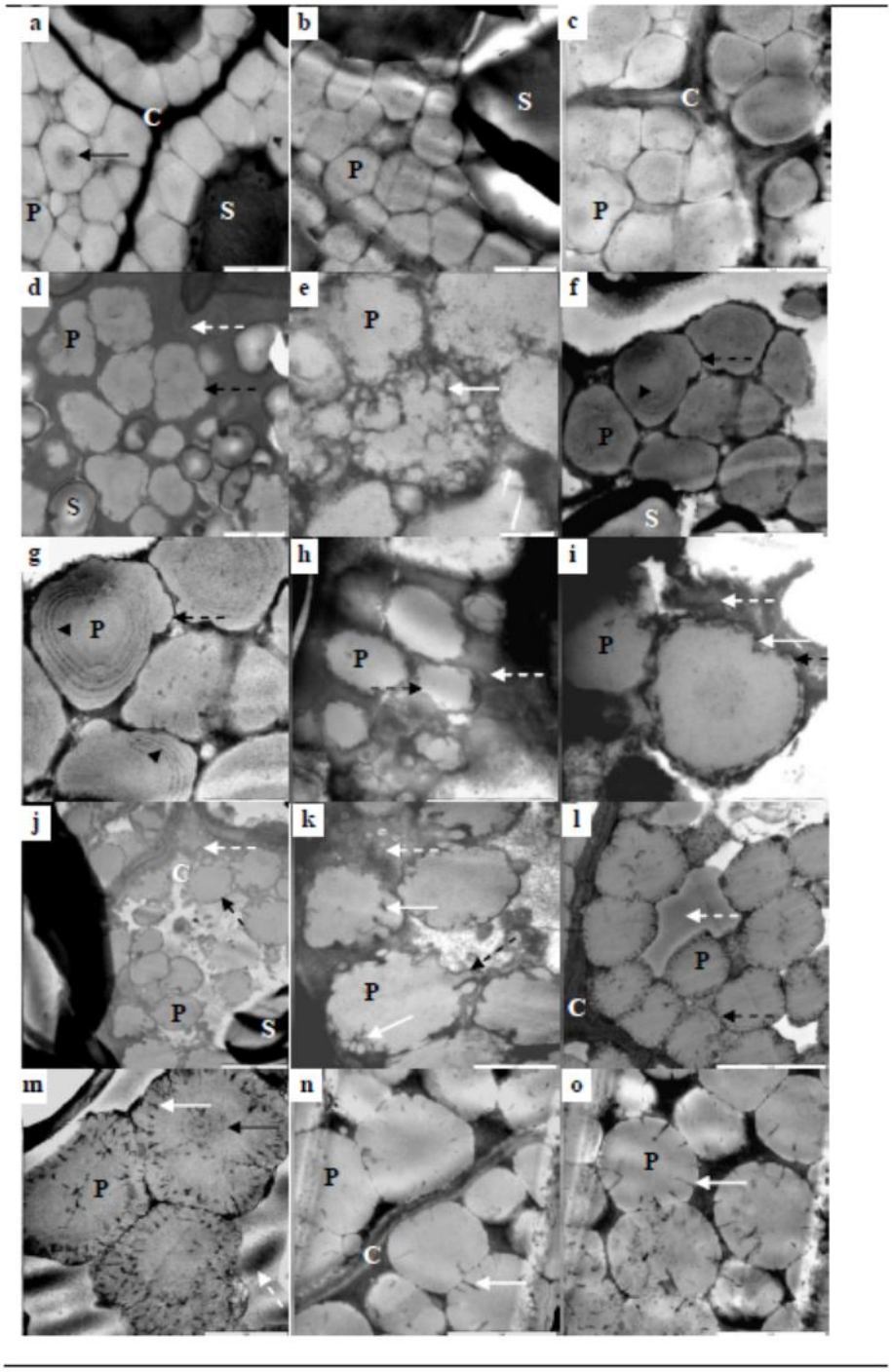

Figure 2