# THE PROLAMIN STORAGE PROTEINS OF SORGHUM AND MILLETS

P R Shewry[1] and N G. Halford[2]

[1]Long Ashton Research Station, Long Ashton, Bristol BS41 9AF, UK,
E-mail: peter.shewry@bbsrc.ac.uk

[2]Rothamsted Research, Harpenden, Herts AL5 2JQ, UK

Most millets are classified with maize, sorghum and *Coix* (Job's tears) in the grass sub-family Panicoideae, but two species, tef and ragi (finger millet) are placed in a separate sub-family, the Chloridoideae. In all species the major storage proteins are alcohol-soluble prolamins, but their properties vary. For example, although the kafirins of sorghum are closely related to the zeins of maize, they differ in their solubility properties, being more efficiently extracted with 60% (v/v) tertiary butanol, a less polar solvent than propan-2-ol. Maize prolamins are classified into four groups, termed alpha-, beta-, gamma- and delta-zeins. Of these, the alpha-zeins are the major fraction but have low nutritional quality, with low contents of cysteine and methionine and no lysine or tryptophan. In contrast, the beta-zeins and delta-zeins contain about 11 mol% and over 20 mol% methionine, respectively.

Alpha-type prolamins appear to be the major components in sorghum, *Coix* and all millet species which have been analysed in detail (pearl millet, foxtail millet, kodo millet, tef, ragi), with the exception of fonio (*Digitaria exilis*). Sorghum and *Coix* also contain homologues of the beta-and gamma-zeins of maize and sorghum homologues of delta-zein. Less well-characterized methionine-rich components have also been reported in foxtail millet and fonio, forming the major prolamin group in the latter species.

## INTRODUCTION

The name millet is applied to a number of small seeded cereals, most of which are native to the tropics or sub-tropics. Most of these are classified with maize, sorghum and *Coix* (Job's tears) in the grass sub-family Panicoideae, but two species are placed in a separate sub-family, the Chloridoideae. These are ragi (Indian finger millet, *Eleusine coracana*) and tef (*Eragrostis tef*).

Maize is one of the three most important crops in the world, the others being rice and wheat. Hence it has been the subject of a massive volume of research which provides a framework for more restricted studies of millets and sorghum.

### Prolamins

The grain storage proteins of maize, millets and sorghum can be defined as prolamins in that they are soluble in alcohol/water mixtures. Although 60-70% (v/v) ethanol was initially used for extraction, it is now more usual to use 50% (v/v) propan-1-ol.

However, 60% (v/v) tertiary butanol (2-methyl-2-propanol) is a more efficient solvent for the prolamins (kafirins) of sorghum, which may indicate that they have a more hydrophobic nature.

Furthermore, in all species some prolamins are only efficiently extracted if the solvent also contains a reducing agent, either dithiothreitol or 2-mercaptoethanol. This is because they form polymers stabilized by inter-chain disulphide bonds, with the reduced subunits but not the polymers being alcohol-soluble. Also, the alcohol-soluble fractions may contain disulphide-stabilized dimers and oligomers.

Prolamins account for about half of the total grain proteins in all these species, the precise amounts being determined by the availability of nutrients (N and S).

## Prolamin groups

SDS-PAGE of the prolamins (zeins) of maize shows a number of bands with apparent molecular masses ranging from about 27,000 to 10,000 (Figure 1). Detailed studies of purified proteins and cloned cDNAs and genes have allowed these to be classified into four groups of proteins, called alpha-, beta-, gamma- and delta-zeins. These are discussed in detail below. Similar analyses of prolamin fractions from sorghum, *Coix* and millets show components with similar mobilities (Figure 1), which suggests that related proteins are present.

## Alpha-prolamins

The alpha-zeins are the major components in maize, accounting for about 70% of the total prolamins. SDS-PAGE shows two broad bands of mass about 19,000 (Z19) and 22,000 (Z22) but isoelectric focusing shows that each of these comprises a number of components. In fact, up to 40 components can be resolved by 2-D electrophoresis[1] with 70 to 100 alpha-zein genes being present[2,3]. Determination of amino acid sequences demonstrates that the true masses range from 23,000-24,000 for Z19 zeins and 26,500-27,000 for Z22 zeins[4]. Alpha-zeins contain 0 or 1 cysteine residues and so are present as monomers or alcohol-soluble dimers. Their amino acid sequences contain peptide repeat motifs of about 20 residues, which are poorly conserved but tend to be rich in leucine, alanine and glutamine residues. Nine repeats are present in the Z19 and ten in the Z22 zeins[5].

Alpha-type prolamins also form the major prolamin components in *Coix* and sorghum, and these appear to be more closely related in sequence to the Z22 than to the Z19 zeins[6,7]. It has been proposed that the repeat units present in these proteins form alpha-helices which are arranged in antiparallel fashion to give a compact circular or hexagonal structure[5,8] but direct evidence for this structure is lacking.

The low proportions of lysine and tryptophan in alpha-type prolamins contribute to the poor nutritional quality of the whole grain.

*N*-terminal amino acid sequences indicate that the major prolamins in pearl millet (pennisetin), tef and ragi are also alpha-type[9,10] but detailed information is not available on these components.

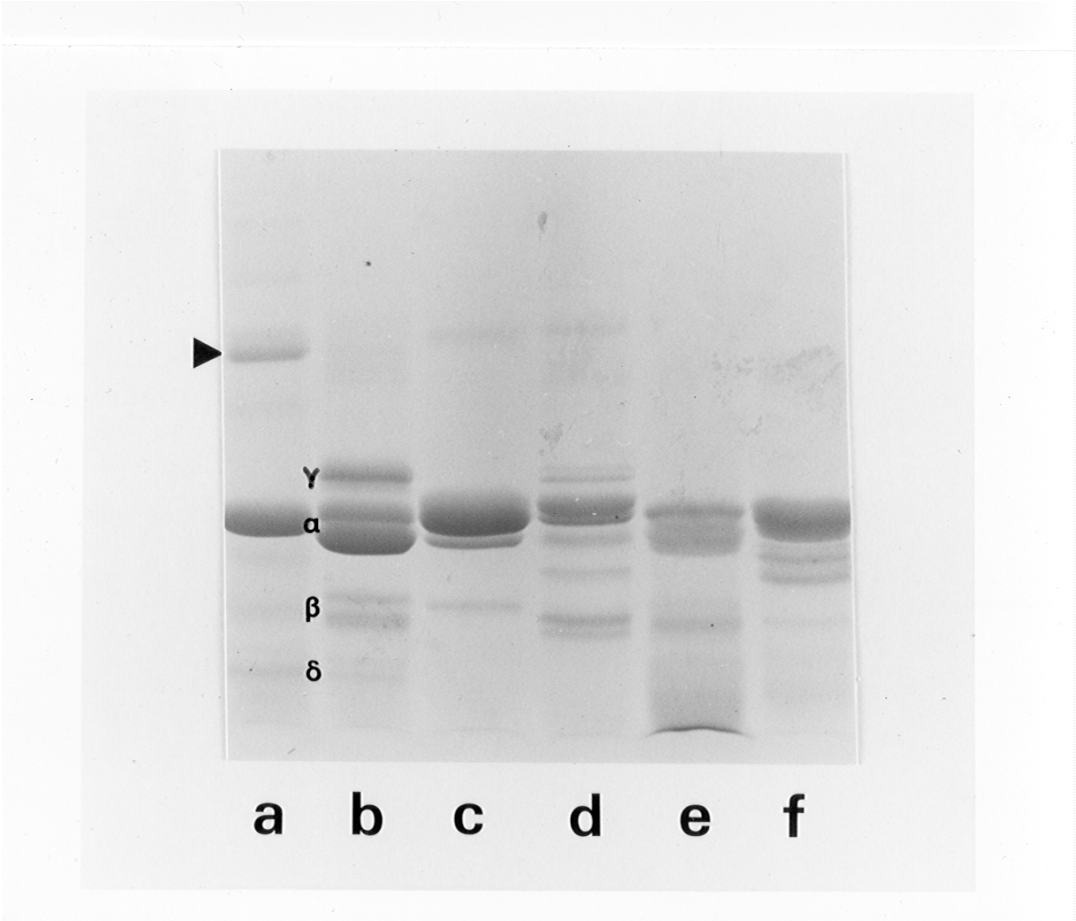


**Figure 1. SDS-PAGE of prolamin fractions from a pearl millet; b, maize; c, sorghum; d, *Coix*; e, tef; f, finger millet. α, β, γ and δ refer to the groups of zeins in track b. Taken from ref[23]**

## Beta-, gamma- and delta-prolamins

The minor prolamins of maize and related cereals all belong to a major family of plant proteins called the "prolamin superfamily". This includes the prolamins of the temperate cereals (wheat, barley, rye, oats) and a number of small sulphur-rich proteins, many of which are restricted to seeds. They include 2S storage albumins of dicotyledonous plants, non-specific lipid transfer proteins, and puroindolines, grain softness proteins and alpha-amylase/trypsin inhibitors of cereal seeds. The structural and evolutionary relationships of these proteins have been discussed in detail by Shewry *et al.*[11]. The beta-, gamma- and delta-prolamins all form polymers stabilized

by inter-chain disulphide bonds and are predominantly, if not completely, insoluble in aqueous alcohols unless a reducing agent is included.

The beta-zeins of maize have molecular masses of about 15,000 by SDS-PAGE with true masses of about 17,500. They comprise about 160 amino acids with 18 residues of methionine and 7 residues of cysteine[12]. Shull *et al.*[13] proposed that $M_r$ 16,000, 18,000 and 20,000 kafirin bands corresponded to beta-kafirins, although only the $M_r$ 20,000 band reacted with an antiserum to beta-zein. This conclusion was supported by amino acid analysis of a fraction containing the $M_r$ 16,000 and 20,000 bands which showed 6 mol% methionine and 5 mol% cysteine. It is rather surprising that cloned cDNAs or genes for beta-kafirin have not yet been reported, but a cloned cDNA for beta-coixin from *Coix lachryma-jobi* showed high similarity to beta-zein[14].

Gamma-zeins are the second most abundant group of proteins in maize, with two bands of $M_r$ about 27,000 (true mass about 22,000) and 16,000. They differ from all other zeins in being soluble in water as reduced subunits. Both sorghum and *Coix* contain prolamin bands of $M_r$ about 27,000 - 28,000 which react with gamma-zein antiserum and cloned cDNAs and genes from both species encode proteins with high homology to gamma-zeins (Figure 2)[15-19]. All of the proteins have clear domain structures, with an *N*-terminal domain comprising repeats of the hexapeptide motif Pro.Pro.Pro.Val.His.Leu. Eight almost perfectly conserved repeats are present in the $M_r$ 27,000 gamma-zein, four in gamma-kafirin, two in gamma-coixin but only two incomplete copies in the $M_r$ 16,000 gamma-zein. The high histidine content may account for the solubility of the reduced subunits in water. All gamma-prolamins also contain ten conserved cysteine residues and are rich in proline, glutamine and non-polar amino acids.

The delta-zeins of maize comprise two minor components of $M_r$ about 10,000 and 18,000 with true masses of about 14,400 and 21,100[4]. Both are methionine-rich, with the $M_r$ 18,000 component (26.9 mol% Met) having apparently been derived from the $M_r$ 10,000 protein (22.8 mol% Met) by duplication of part of the methionine-rich central region. A DNA-derived sequence for a delta-kafirin of sorghum shows high homology with the $M_r$ 10,000 delta-zein, except for the absence of part of the methionine-rich region (Figure 3). Consequently the methionine content is lower than that of the delta-zeins, with 18 residues out of a total of 114 (15.8 mol%).

The minor prolamins of maize are encoded by small gene families, with one beta-zein gene, two or three gamma-zein genes and two delta-zein genes[21].

## Minor prolamins in millets

It is likely that millets also contain prolamins related to the beta-, gamma- and delta-prolamins of maize and sorghum, but the only component which has been conclusively identified is from fonio (*Digitaria exilis*). The whole grain of fonio contains about 4.8 wt% methionine and this derives from the presence of two methionine-rich proteins which account for about 35% of the total prolamin fraction[20]. These components have $M_r$ of about 19,000 (6.4 mol% Met, 6.4 mol% Cys) and

17,500 (7.8 mol% Met, 5.3 mol% Cys) and the former shows *N*-terminal amino acid sequence homology with the $M_r$ 10,000 delta-zein of maize. Naren and Vivupaksha[21] also prepared two methionine-rich prolamins from Italian or foxtail millet (*Setaria italica*), which had masses by SDS-PAGE of about 7,900 and 9,100. These proteins, called alpha- and beta-setarins, could also be related to the beta- or delta-zeins but the same authors showed that antibody to alpha-setarin did not react with prolamins from maize or sorghum[22].

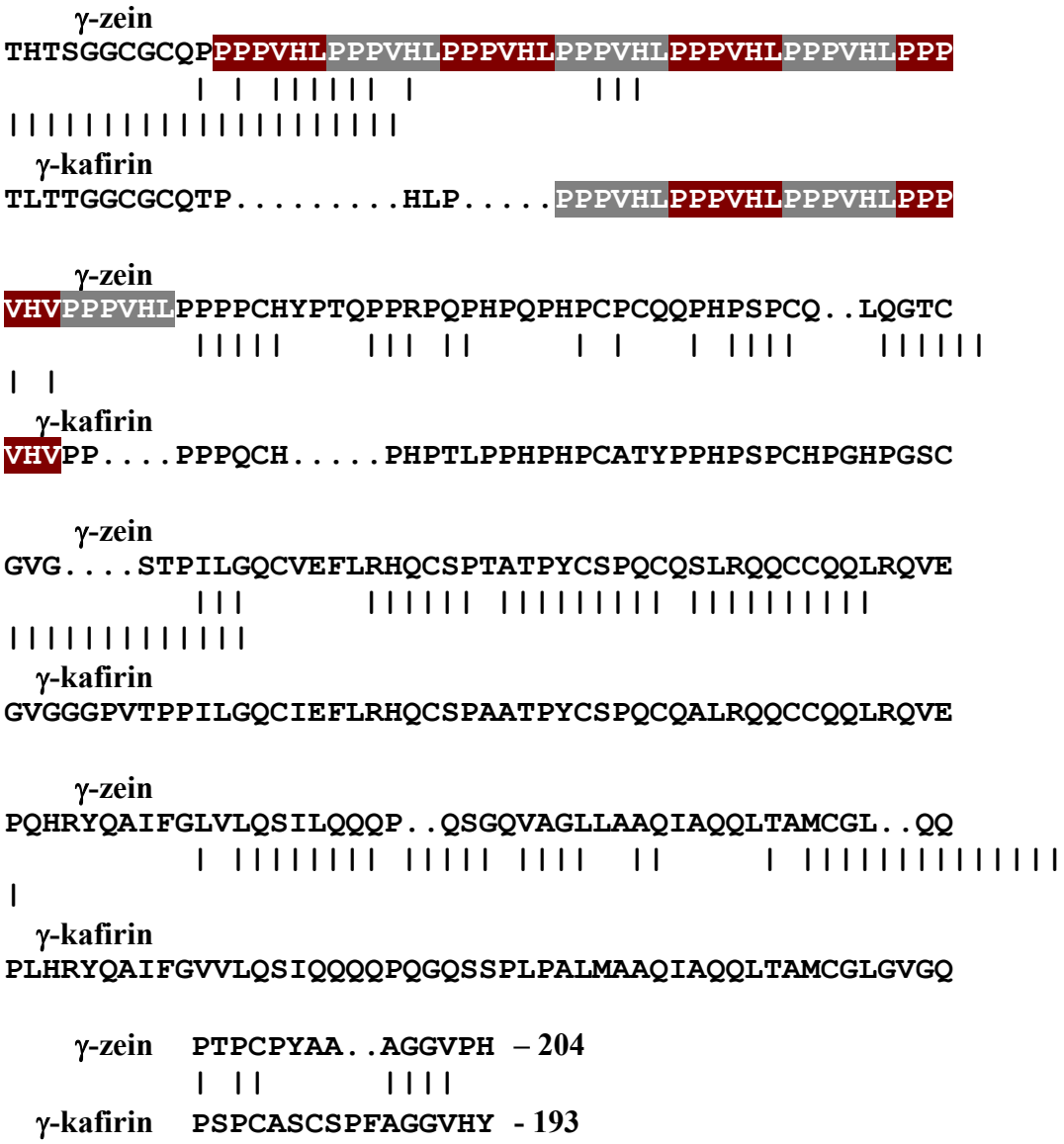


**Figure 2. Alignment of the amino acid sequences of the $M_r$ 27,000 gamma-zein of maize and gamma-kafirin of sorghum. The boxes indicate repeats of the conserved motif Pro.Pro.Pro.Val.His.Leu**

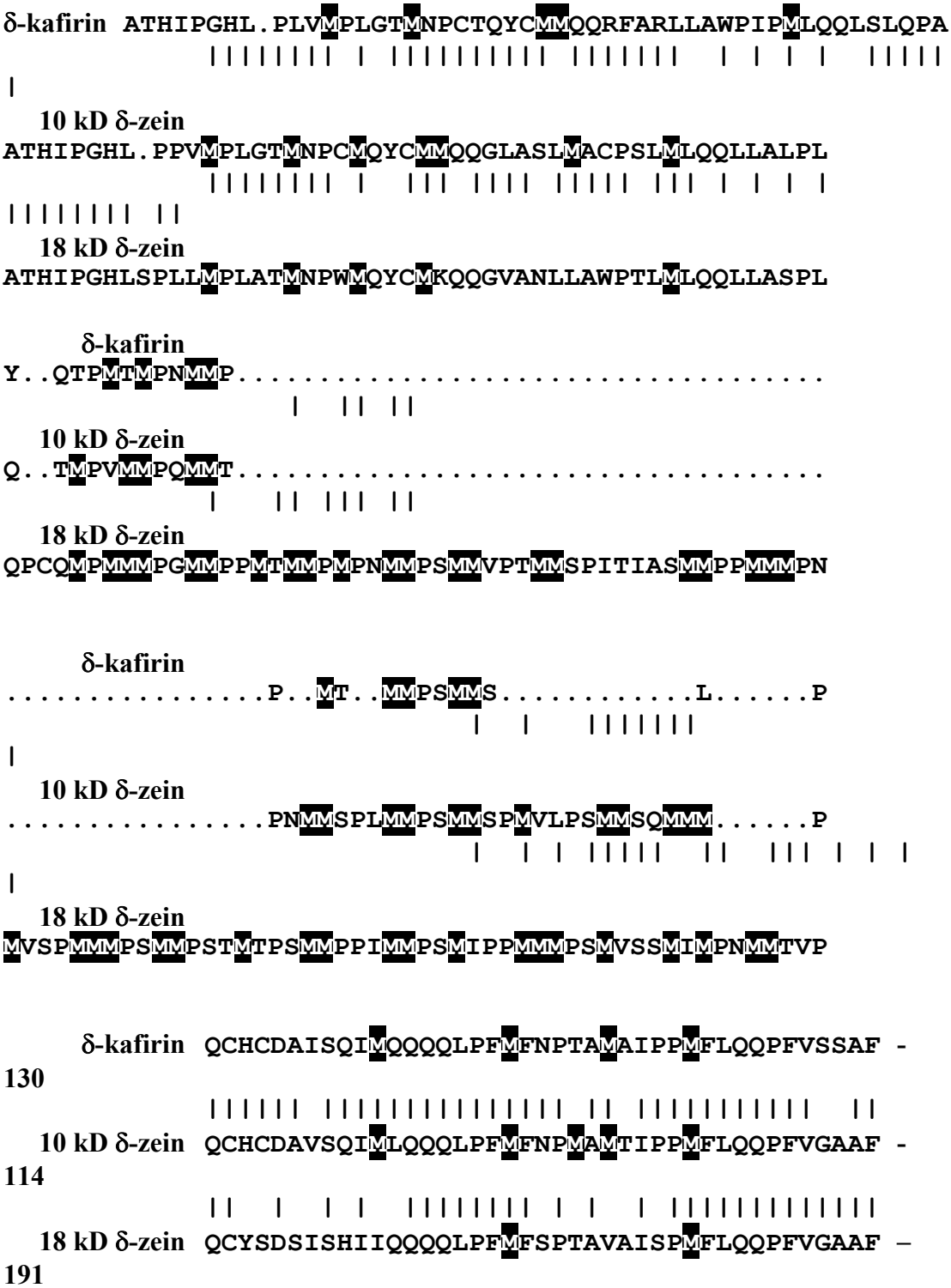


**Figure 3. Alignment of the amino acid sequences of $M_r$ 18,000 and $M_r$ 10,000 delta-zeins of maize and delta-kafirin of sorghum. Lines indicate residues in the delta-kafirin and $M_r$ 18,000 delta-zein that are identical to those in the $M_r$ 10,000 delta-zein. Methionine residues are shown in black boxes.**

## CONCLUSIONS

Although sorghum and millets appear to contain similar prolamin storage proteins to maize the amount of available information on these species is limited. Further detailed studies are clearly required, particularly of the minor millets. Of particular interest is the presence in most, if not all, of methionine-rich prolamins related to the beta- and delta-zeins of maize. The selection of genotypes containing high proportions of these proteins or the manipulation of their amounts by genetic engineering could lead to improved nutritional quality.

## REFERENCES


1. Wall, J.S., Fey, D.A. and Paulis, J.W. Improved two-dimensional electrophoretic separation of zein proteins: Application to study of zein inheritance in corn genotypes. *Cereal Chemistry* **33** (1984) 141-146.

2. Hagen, G. and Rubenstein, I. Complex organization of zein genes in maize. *Gene* **13** (1981) 239-249.

3. Wilson, D.R. and Larkins, B.A. Zein gene organization in maize and related grasses. *Journal of Molecular Evolution* **29** (1984) 330-340.

4. Coleman, C.E. and Larkins, B.A. Prolamins of maize. In 'Seed proteins', (P.R. Shewry and R. Casey, eds), Kluwer Academic Publishers, The Netherlands (1999) pp. 109-139.

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

6. Ottiboni, L.M.M., Leite, A., Yunes, J.A., Targon, M.L.P.N., Souza Filho, G.A. and Arruda, P. Sequence analysis of 22 kDa-like alpha-coixin genes and their comparison with homologous zein and kafirin genes reveals highly conserved protein structure and regulatory elements. *Plant Molecular Biology* **21** (1993) 765.

7. DeRose, R., Ma, D-P., Kwon, I-S., Hasnain, S.E., Klassy, R.C. and Hall, T. Characterization of the kafirin gene family from sorghum reveals extensive homology with zein from maize. *Plant Molecular Biology* **12** (1989) 245-256.

8. Argos, P., Pedersen, K., Marks, M.S. and Larkins, B.A. A structural model for maize zein proteins. *Journal of Biological Chemistry* **257** (1982) 9984-9990.

9. Bietz, J.A. Cereal prolamin evolution and homology revealed by sequence analysis. *Biochemical Genetics* **20** (1982) 1039-1053.

10. Tatham, A.S., Fido, R.J., Moore, C.M., Kasarda, D.D., Kuzmicky, D.D., Keen, J.N. and Shewry, P.R. Characterisation of the major prolamin of Tef (*Eragrostis tef*) and finger millet (*Eleusine coracana*). *Journal of Cereal Science* **24** (1996) 65-71.

11. Shewry, P.R., Jenkins, J., Beaudoin, F. and Mills, E.N.C. The classification, functions and evolutionary relationships of plant proteins in relation to food allergens. In 'Plant Food Protein Allergens' (E.N.C. Mills and P.R. Shewry, eds) Blackwell Science Ltd. (2003) In press.

12. Pedersen, K., Argos, P., Naravana, S.V.L. and Larkins, B.A. Sequence analysis and characterization of a maize gene encoding a high-sulfur zein protein of $M_r$ 15,000. *Journal of Biological Chemistry* **261** (1986) 6279-6284.

13. Shull, J.M., Watterson, J.J. and Kirleis, A.W. Proposed nomenclature for the alcohol-soluble proteins (kafirins) of *Sorghum bicolor* (L. Moench) based on molecular weight, solubility and structure. *Journal of Agricultural and Food Chemistry* **39** (1991) 83-87.

14. Leite, A., Yunes, J.A., Turcinelli, S.R. and Arruda, P. Cloning and characterization of a cDNA encoding a sulfur-rich coixin. *Plant Molecular Biology* **18** (1992) 171-174.

15. Leite, A., Ottoboni, L.M.M., Targon, M.L.P.N., Silva, M.J., Turcinelli, S.R. and Arruda, P. Phylogenetic relationship of zeins and coixins as determined by immunological cross-reactivity and Southern blot analysis. *Plant Molecular Biology* **14** (1990) 743.

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

17. Watterson, J.J., Shull, J.M., Mohamed, A.A. and Kirleis, A.W. Isolation of a high-cysteine kafirin protein and its cross-reactivity with gamma-zein antiserum. *Journal of Cereal Science* **12** (1990) 137-144.

18. Barros, E.G., Takasaki, K., Kirleis, A.W. and Larkins, B.A. Nucleotide sequence of a cDNA clone encoding gamma-kafirin protein from *Sorghum biocolor*. *Plant Physiology* **97** (1991) 1606-1607.

19. Leite, A., Freitas, F.A., Yunes, J.A. and Arruda, P. Nucleotide sequence of a cDNA clone encoding gamma-coixin from *Coix lacryma-jobi* seeds. *Plant Physiology* **97** (1991) 1604-1605.

20. de Lumen, B.B., Thompson, S. and Odegard, W.J. Sulfur amino acid-rich proteins in aca (*Digitaria exilis*), a promising underutilized African cereal. *Journal of Agricultural and Food Chemistry* **41** (1993) 1045-1047.

21. Naren, A.P. and Virupaksha, T.K. Alpha- and beta-setarins: methionine-rich proteins of Italian millet (*Setaria italica* (L.) Beauv.). *Cereal Chemistry* **67** (1990a) 32-34.

22. Naren, A.P. and Virupaksha, T.K. Effect of sulfur deficiency on the synthesis of alpha-setarin, a methionine-rich protein of Italian millet. *Cereal Chemistry* **67** (1990b) 136-138.

23. Shewry, P.R. The seed storage proteins of pseudocereals and minor cereals. In '*Pseudocereals and Less Well-known Cereals'* (P. Belton and J. Taylor, eds) Springer-Verlag. (2002) pp. 1-24.