International Journal of Chemical Studies

P-ISSN: 2349–8528
E-ISSN: 2321–4902
www.chemijournal.com
IJCS 2021; 9(1): 3062-3068

Received: xx-11-2020
Accepted: xx-12-2020

Sonawane Chetan
Department of Plant Biotechnology, Centre for Plant Molecular Biology & Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

S Varanavasiappan
Department of Plant Biotechnology, Centre for Plant Molecular Biology & Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

A Muthuswamy
Department of Pulses, Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore 641003, India

KK Kumar
Department of Plant Biotechnology, Centre for Plant Molecular Biology & Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

Corresponding Author:
KK Kumar
Department of Plant Biotechnology, Centre for Plant Molecular Biology & Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

Isolation and Characterization of 21.7 kDa prolamin gene from Foxtail millet (*Setaria italica*)

Sonawane Chetan, S Varanavasiappan, A Muthuswamy and KK Kumar

DOI: https://doi.org/10.22271/chemi.2021.v9.i1aq.11697

Abstract
Prolamins are the most distributed seed storage proteins in cereals. Prolamins are the main seed storage protein of foxtail millet (*Setaria italica*). The genomic and cDNA clone of a prolamin gene was isolated from foxtail millet. The isolated prolamin gene of foxtail millet lacks an intron based on a sequence contrast between the genomic and the cDNA clones. Semi-quantitative RT-PCR analysis showed a higher level of expression of the prolamin gene in an immature panicle. Physiochemical properties analysis based on the deduced amino acid sequence of foxtail millet prolamin gene showed that it codes for 204 amino acids, having a molecular weight of 21.7 kDa, pI 8.98 and GRAVY value 0.467. Homology search done with 21.7 kDa prolamin protein showed the highest similarity with α-zein of *Setaria italic, Setaria viridis* and *Panicum miliaceum.* The secondary structure predicted for the 21.7 kDa prolamin protein of foxtail millet indicates 14 coils and 15 helices. Like other prolamins, the 21.7 kDa prolamin protein is also deficient in the essential amino acid lysine.

Keywords: Foxtail millet, 22kDa prolamin, seed storage protein

Introduction
Seeds of cereals and legumes are the primary sources of protein all over the world, which are consumed directly as food and feed by humans and animals, providing more than 70% of the world's caloric intake (Mandal and Mandal, 2000) [5]. Still, these crops are commonly deficient in quality proteins and nutritionally balanced amino acids. So there is a need for isolation and characterization of genes for nutritionally important proteins that can subsequently be transferred to cereals and legumes by genetic engineering for biofortification purposes.

Seed storage proteins are proteins that accumulate significantly within the developing seed, the main purpose of which is to serve as a storage reservoir for nitrogen, carbon, and sulfur. These proteins are easily mobilized during the seed germination process and act as the primary source of reduced nitrogen for growing seedlings. Also, seed storage proteins did not take part on carry enzymatic reactions. Seed storage proteins of different plants are structurally different, but they share similar characteristics. One of the key features of seed protein storage is that it accumulates at a high level in specific tissues at a specific stage of development (Shewry *et al.*, 2003) [9]. Storage proteins in seeds are the products of several specific genes whose expression is tightly regulated during development. Thus, SSP genes provide a powerful system to investigate the regulation of their expression (Xi and Zheng, 2011) [11].

Seed storage proteins are usually classified according to their solubility as albumins (water-soluble), globulins (salt soluble), prolamins (alcohol soluble) and glutelin's (soluble in diluted acid or base) (Osborne, 1924) [6]. Variations in aqueous solubility, capacity to form disulfide associations, molecular weight, and gene sequence were subsequently used to classify the prolamin sub-family. For example, the zeins and the kafirins are grouped into α, β, γ, and δ types (Holding, 2014) [3]. Prolamin is the dominant seed storage protein in cereals, being high in amino acids proline and glutamine (Shewry and Halford, 2002) [10]. These proteins are distinguished based on of 6-8 cysteine residue and 3-4 disulfide bonds that are strongly soluble in 40-70% ethanol. Prolamins are further classified into three different types' *viz*., sulfur-rich prolamins (α, β, γ gliadins), sulfur-poor prolamins (ω-gliadins) and high molecular weight prolamins. The prolamin protein is named based on the cereal source in which it was derived, setarin for foxtail millet, zein for corn, kafirin for sorghum, and hordein for barley. Foxtail millet is one of the most important millet crops.

It is better adapted to dry, adverse soils than most other grain crops, such as maize and sorghum (Panaud, 2006) [7]. It can be grown in both tropical as well as temperate regions under low average rainfall. It has several characteristics like drought tolerance, wide adaptation, and High genetic diversity. Mature foxtail seeds comprise primarily of proline-rich, alcohol-soluble prolamin (setarin) proteins, containing approximately 60% of the total protein, with less disulfide cross-linked protein content than other cereal and millet seeds (Sachdev *et al.*, 2021) [8]. In this study, we characterize the isolated 21.7 kDa prolamin gene and predict its physical and chemical properties using a bioinformatics tool.

Material and methodology

Isolation of genomic clone of foxtail millet prolamin gene

Genomic DNA was isolated from the young leaf of foxtail millet (Co.7 variety) using the CTAB method (Doyle and Doyle, 1990) [2]. The 21.7kDa prolamin genomic sequence of foxtail millet was retrieved from the NCBI database and primers were designed for amplifying the full-length gene of the 21.7 kDa prolamin gene. The 21.7 kDa prolamin gene was amplified using the designed forward (CCGACAACCAAC CTCTTAGC) and reverse primer (TCTAGAAGAGGGCAG AACCA). About 50 ng of genomic DNA was used as the template for PCR amplification. PCR conditions were initial denaturation at 94 °C for 5 min accompanied by 35 cycles at 94 °C for 1 min, 58 °C for 1 min, 72 °C for 1 min, and 72 °C for 10 min. PCR amplified product has been sequenced (Scigenom Pvt. Ltd., Cochin).

Isolation of cDNA clone of foxtail millet prolamin gene

Total RNA was isolated from young leaves and immature panicles of foxtail millet using TRIzol reagent (Bio-basic, Canada). Qualitative analysis of the total isolated RNA was done using 1.2% agarose gel electrophoresis and quantified spectrophotometrically. Isolated RNA further used for cDNA synthesis (iScript™ cDNA Synthesis kit from Bio-rad, USA) following manufacturer instructions. To ensure the absence of genomic DNA contamination in the cDNA preparation, foxtail millet actin gene was amplified using the forward (TGTGATGTTGATATCAGGAAGGA) and (GGG ACCGGTTTCGTCATACT) reverse primer, That amplifies 200 bp with cDNA and 650 bp with genomic DNA as a template. PCR cycling parameters is as follows, 94 °C for 5 min, followed by 35 cycles of 94 °C for 45s, 56 °C for 45s, 72 °C for 1min and 72 °C for 10 min. For RT-PCR amplification, 591 bp fragment of the 21.7 kDa prolamin gene was amplified using the designed forward (CTCTTTCAGTGAGCGCTGC) and reverse primer (TCTAGAAGAGGGCAGAACCA). Semi-quantitative RT- PCR cycling parameters is as follows,94 °C for 5 min, followed by 35 cycles of 94 °C for 45s, 56 °C for 45s, 72 °C for 1min and 72 °C for 10 min.

Sequence analysis of 21.7 kDa prolamin gene

The nucleotide sequences of isolated prolamin gene from genomic DNA and cDNA sequence were compared using ClustalW2 online tool to find the presence of intron. Deduced amino acid sequence from the 21.7 kDa prolamin gene was used for protein parameter analysis (https://web.expasy.org/protparam/) as well as a query to identify homologous genes in foxtail millet and different plant species using BLAST p search tool against protein sequences in the NCBI database. Multiple sequence alignment of isolated 21.7 kDa prolamin protein and other closely related prolamin protein from foxtail millet and related crop species*viz. Setaria italica* (XP_004979704.1), *Setaria viridis* (XP_034606929.1), *Cenchrus americanus* (ABR09541.1), *Panicum miliaceum* (RLM68922.1), *Panicum hallii* (XP_025827903.1), *Sacchrum officinarum* (ABP64786.1), *Sorghum bicolour* (XP_002451115.1), *Zea mays* (AQK50540.1) was carried out using CLUSTALW tool in Bio Edit software in order to find conserved domains. This amino acid alignment was used for phylogenetic analysis by MEGA6 software using the maximum likelihood method with 1,000 bootstrap replications by the Poisson model. The domains conserved in this sequence were predicted using the conserved domain database (CDD) search tool.

Sequence characterization of 21.7 kDa prolamin gene of foxtail millet

The secondary structure was predicted using the PRISPRED tool (http://bioinf.cs. ucl.ac.uk/psipred/). The physicochemical properties like molecular weight (MW), hydropathicity (GRAVY), and isoelectric point were determined using the PROTOPARAM tool (https://web.expasy.org/protparam/). The hydrophobicity analysis was done by using Kyte and Doolittle scale (https://web.expasy.org/protscale/) (Kyte and Doolittle, 1982). Transmembrane domain was predicted using PROTTER server (http://wlab.ethz.ch/protter/#) and CELLO2GO web for protein subcellular localization prediction with gene ontology annotation server (http://cello.life.nctu.edu.tw/cello2go/).

Result and Discussion

Isolation of 21.7 kDa prolamin gene of foxtail millet

Using the designed primers, the full-length prolamin gene was amplified using foxtail millet genomic DNA and cDNA. DNA sequence of both genomic clone and cDNA clone showed 615 bp gene and sequence alignment between them demonstrated the absence of introns (Fig. 1).

Expression analysis of 21.7 kDa prolamin gene of foxtail millet

The cDNA from leaf and immature panicle of foxtail millet was used for semi-quantitative RT-PCR analysis. RT-PCR amplification of actin gene confirmed the absence of genomic DNA contamination in the cDNA preparations (Fig.2). Semi-quantitative-PCR analysis shows that the expression of 21.7 kDa prolamin gene is high in the immature panicle than in the leaf (Fig. 3).

In silico analysis of 21.7 kDa prolamin protein of foxtail millet

The conserved domain search done with 21.7 kDa prolamin protein predicted alignment with the Pfam 01559 zein superfamily (Pssm-ID: 366705). Protein sequence alignment between the isolated prolamin gene and prolamin gene deposited in NCBI database showed variation at two amino acid positions (L46V and V151L) (Fig. 4). The PROTOPARAM tool predicted the physiochemical properties of protein sequences such as molecular weight, isoelectric point, aliphatic index, and GRAVY value. The molecular weight of the isolated prolamin protein is 21.7 kDa. The isoelectric point (pI) value of 8.15 indicates that this protein is basic in nature. The Instability Index (II) is estimated to be 65.56 and demonstrates that 21.7 kDa protein is unstable in nature, i.e. short-lived protein. Aliphatic index of 120.34 reveals that this 21.7 kDa protein is stable at a wide range of temperatures. The Grand Average of hydropathicity (GRAVY) is 0.467, showing that 21.7 kDa protein is a

hydrophobic protein. The major amino acids in the protein are alanine (19.6 per cent), leucine (18.6 per cent), glutamine (16.7 per cent), and proline (8.3 per cent).

Multiple sequence alignment of isolated 21.7 kDa prolamin protein and other closely related prolamin protein from foxtail millet and related crop species *viz. Setaria italica* (XP_004979704.1), *Setaria viridis* (XP_034606929.1), *Panicum miliaceum* (RLM68922.1), *Panicum hallii* (XP_025827903.1) showed high level of sequence homology with conserved motif (Fig. 5). Phylogenetic analysis was done between the isolated 21.7 kDa prolamin protein, deposited S. Italic prolamin and other closely related prolamin protein from related crop species *viz. Setaria italica* (XP_004979704.1), *Setaria viridis* (XP_034606929.1), *Cenchrus americanus* (ABR09541.1), *Panicum miliaceum* (RLM68922.1), *Panicum hallii* (XP_025827903.1), *Sacchrum officinarum* (ABP64786.1), *Sorghum bicolour* (XP_002451115.1), Zea mays (AQK50540.1). The phylogenetic tree constructed show close linage with *Setaria italic* (XP_004979704.1), *Setaria viridis* (XP_034606929.1), *Panicum miliaceum* (RLM68922.1) with distant linage with other species (Fig. 6).

Prediction of the secondary structure from the 21.7 kDa genomic DNA protein sequence using the PSIPHRED tool revealed that 21.7 kDa protein contains 14 coils and 15 helices (Fig. 7). The hydrophobic nature of the protein sequence is predicted to be Kyte-Doolittle hydropathy plots (Fig.8). This peptide was found to be located without any transmembrane domains (Fig.9).CELLO server results predict that this 21.7 kDa prolamin protein is mostly found in the plasma membrane and nucleus.

The analysis of spatial and temporal expression of tissue-specific genes is of great value to investigate the molecular mechanism of plant development. Systemic transcriptomic analysis of seed development has been performed in arabidopsis (Day *et al.*, 2008) [1], maize (Xin *et al.*, 2011) [12], and rice (Xu *et al.*, 2012) [13]. Although the expression analyses in endosperm-related genes were involved, these studies mainly focused on the metabolic pathways of endosperm such as nutrient partitioning. Thus, the understandings of the regulatory mechanism of the endosperm development remain limited. Analysis of endosperm-specific genes can help to clone and identify novel genes involved in endosperm development and nutrient synthesis. The characterization of 21.7 kDa prolamin protein of foxtail millet will be useful for assessing the nutritional aspects.

```
22kda_genomic_clone     ATGGCAGCCAAGATATTTGCCTTCCTTGCGCTCCTTGCTCTTTCAGTGAGCGCTGCTACC
22kda_cDNA_Clone        ATGGCAGCCAAGATATTTGCCTTCCTTGCGCTCCTTGCTCTTTCAGTGAGCGCTGCTACC
                        ************************************************************

22kda_genomic_clone     GCGGTCCTTGTACCACAGTGCTCAGTAGCCGCCGCCGCAGCCACAATTCCCCAGTACCTC
22kda_cDNA_Clone        GCGGTCCTTGTACCACAGTGCTCAGTAGCCGCCGCCGCAGCCACAATTCCCCAGTACCTC
                        ************************************************************

22kda_genomic_clone     TCACCTTACACAGCTGTTGGGTATGAACACCCAATTGTGCAATCCTACAGGCTACAGCAG
22kda_cDNA_Clone        TCACCTTACACAGCTGTTGGGTATGAACACCCAATTGTGCAATCCTACAGGCTACAGCAG
                        ************************************************************

22kda_genomic_clone     GCACTTGCAGCAAGCATCCTACCATCATCGGCCATGTTCCTACCACAACAGTCGGCCTTA
22kda_cDNA_Clone        GCACTTGCAGCAAGCATCCTACCATCATCGGCCATGTTCCTACCACAACAGTCGGCCTTA
                        ************************************************************

22kda_genomic_clone     TTGCAGCAGCAATCCCTGTCTCATCTGACAGTACAGAGCATCACGGCACAGCAACAACGC
22kda_cDNA_Clone        TTGCAGCAGCAATCCCTGTCTCATCTGACAGTACAGAGCATCACGGCACAGCAACAACGC
                        ************************************************************

22kda_genomic_clone     ATTCTATCACCGTTGAACCAACTAGCCTTGGCGAACCCCGCCGCATACTTGCAACAACAG
22kda_cDNA_Clone        ATTCTATCACCGTTGAACCAACTAGCCTTGGCGAACCCCGCCGCATACTTGCAACAACAG
                        ************************************************************

22kda_genomic_clone     ACGCTACTCCCGTTCAACCAGCTGGCCCTGGCGAACCCCGCTGCCTTCTTGCAGCAACAA
22kda_cDNA_Clone        ACGCTACTCCCGTTCAACCAGCTGGCCCTGGCGAACCCCGCTGCCTTCTTGCAGCAACAA
                        ************************************************************

22kda_genomic_clone     CAGCTGCTTCCGTTCAACCAACTGGCTGCATTGAACCCAGCCGCCATCTTGCAGCAGCAA
22kda_cDNA_Clone        CAGCTGCTTCCGTTCAACCAACTGGCTGCATTGAACCCAGCCGCCATCTTGCAGCAGCAA
                        ************************************************************

22kda_genomic_clone     CTGTCACCATTGAACCCACTTGCATTGGCAAACCCCGCCGCCTTCTGGCAGCAGCAACAG
22kda_cDNA_Clone        CTGTCACCATTGAACCCACTTGCATTGGCAAACCCCGCCGCCTTCTGGCAGCAGCAACAG
                        ************************************************************

22kda_genomic_clone     CTAGTCAACCAACTAGCTTTGACGAGTCCCGCCGCCTTCTTGCAGCAACCCATCGTTGGT
22kda_cDNA_Clone        CTAGTCAACCAACTAGCTTTGACGAGTCCCGCCGCCTTCTTGCAGCAACCCATCGTTGGT
                        ************************************************************

22kda_genomic_clone     TCTGCCCTCTTCTAG
22kda_cDNA_Clone        TCTGCCCTCTTCTAG
                        ***************
```

Fig 1: Sequence alignment of genomic clone and cDNA clone of isolated prolamin gene.

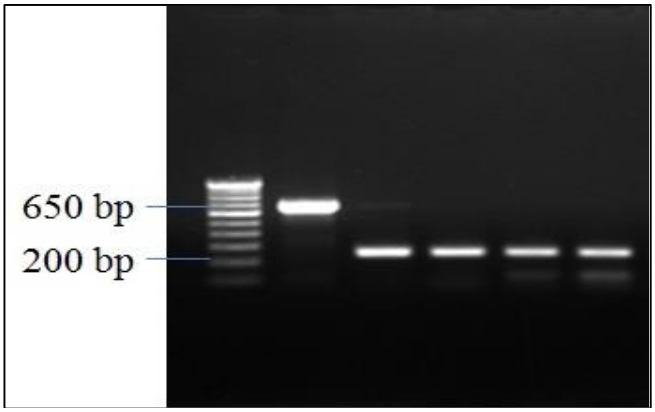

Fig 2: RT-PCR amplification of foxtail millet actin gene. Lane 1- 100 bp ladder, Lane 2-Genomic DNA (Control), Lane 3&4- Leaf cDNA, Lane 5&6- immature panicle cDNA

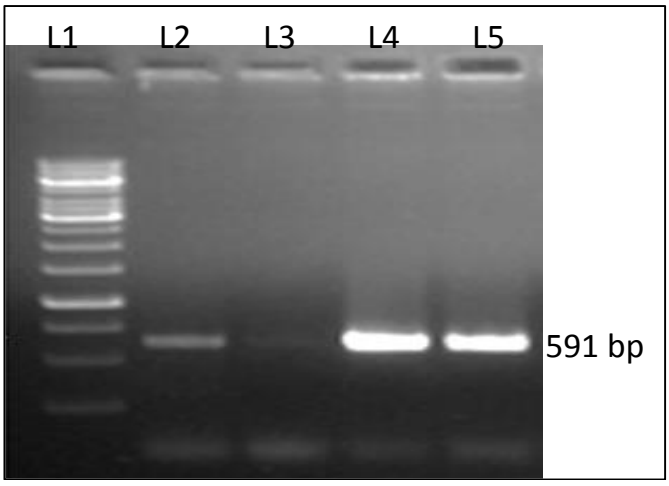

Fig 3: RT- PCR Analysis for 21.7 kDa prolamin gene. Lane 1- 1 kb ladder, Lane 2 &3 –Leaf cDNA, Lane 4&5- Immature panicle cDNA

```
21.7kDa_Prolamin_Cloned      MAAKIFAFLALLALSVSAATAVLVPQCSVAAAAATIPQYLSPYTAVGYEHPIVQSYRLQQ   60
XP_004979704.1_Setaria       MAAKIFAFLALLALSVSAATAVLVPQCSVAAAAATIPQYLSPYTALGYEHPIVQSYRLQQ   60
                             ************************************************:***********

21.7kDa_Prolamin_Cloned      ALAASILPSSAMFLPQQSALLQQQSLSHLTVQSITAQQQRILSPLNQLALANPAAYLQQQ  120
XP_004979704.1_Setaria       ALAASILPSSAMFLPQQSALLQQQSLSHLTVQSITAQQQRILSPLNQLALANPAAYLQQQ  120
                             ************************************************************

21.7kDa_Prolamin_Cloned      TLLPFNQLALANPAAFLQQQQLLPFNQLAALNPAAILQQQLSPLNPLALANPAAFWQQQQ  180
XP_004979704.1_Setaria       TLLPFNQLALANPAAFLQQQQLLPFNQLAAVNPAAILQQQLSPLNPLALANPAAFWQQQQ  180
                             ******************************:*****************************

21.7kDa_Prolamin_Cloned      LVNQLALTSPAAFLQQPIVGSALF    204
XP_004979704.1_Setaria       LVNQLALTSPAAFLQQPIVGSALF    204
*************************
```

Fig 4: Protein sequence alignment between the isolated prolamin gene and prolamin gene deposited in NCBI

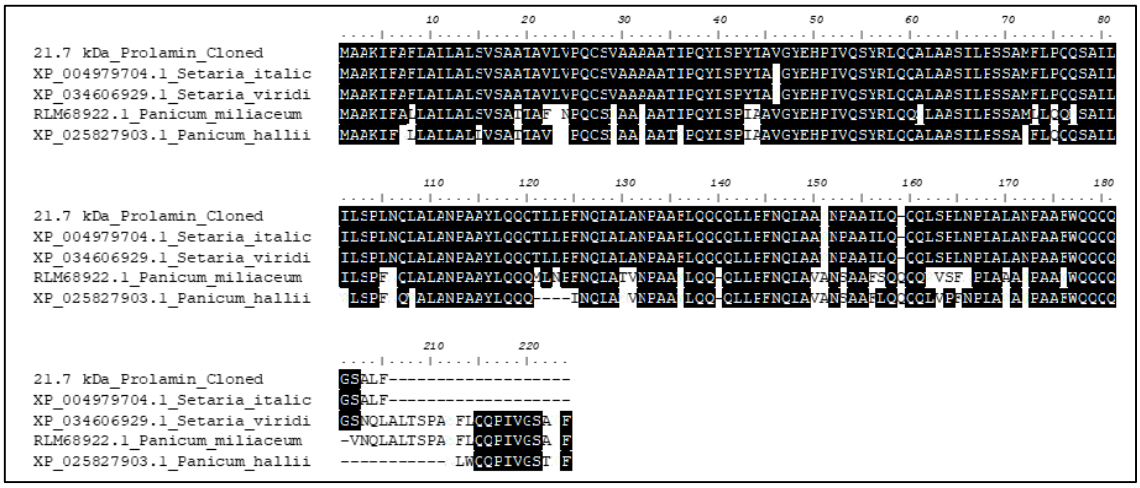

Fig 5: Multiple sequence alignment of 21.7 kDa prolamin and closely related prolamins of other species

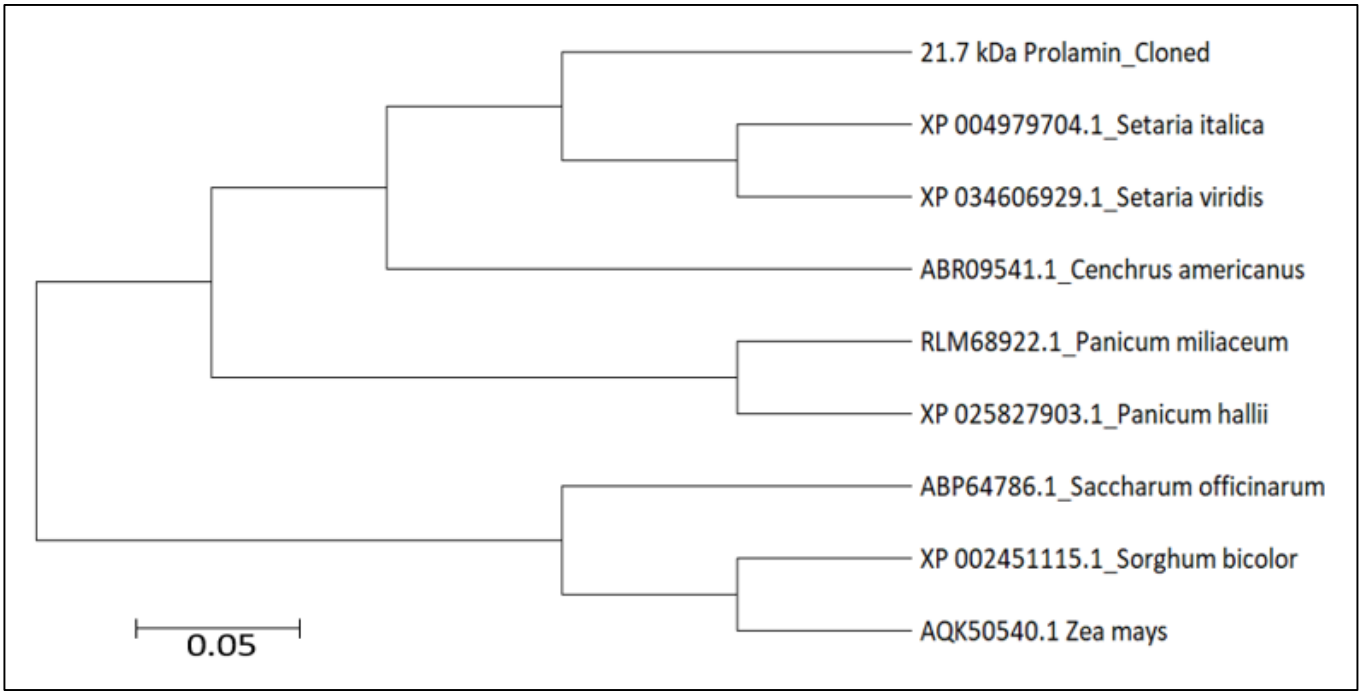

Fig 6: Phylogenetic analysis of 21.7 kDa prolamin protein

Fig 7: Secondary structure prediction of 21.7 kDa prolamin protein of foxtail millet

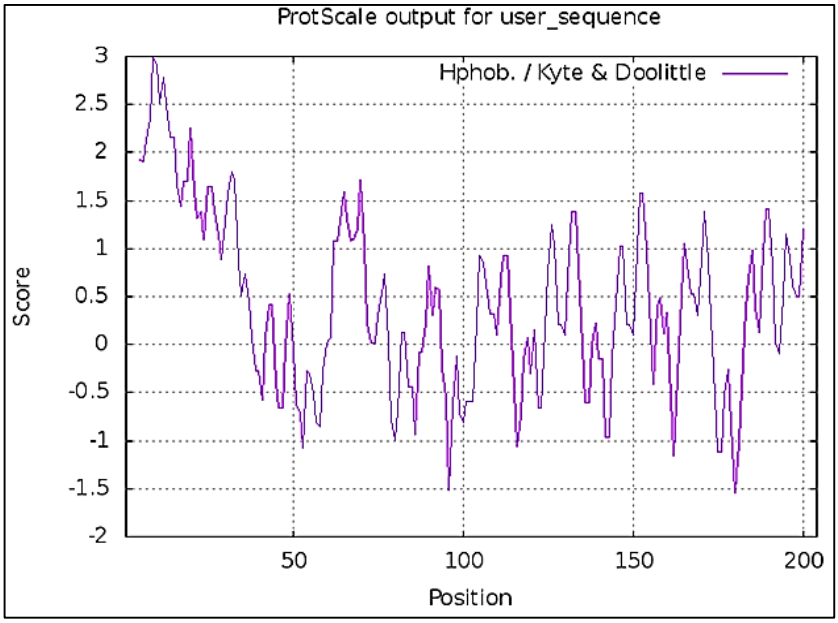

Fig 8: Kyte and Doolittle hydropathy analysis of 21.7 kDa prolamin protein. Above zero donates hydrophobicity and region below zero denotes hydrophilicity

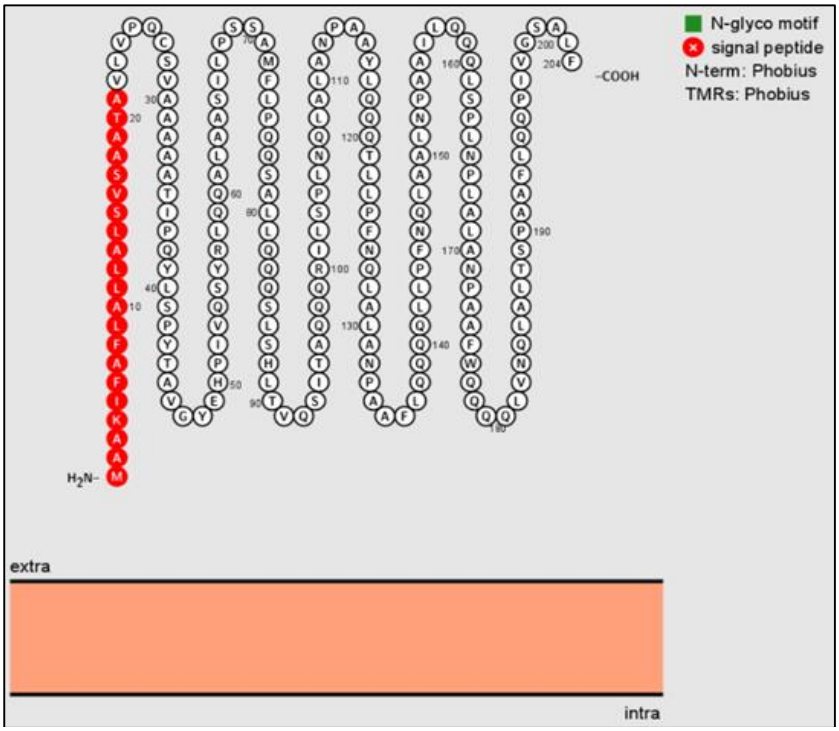

Fig 9: Protein localization of 21.7 kDa prolamin protein of foxtail millet using PROTTER server

Conclusion

Foxtail millet has been known for centuries as a nutritious cereal crop. This research highlights the presence of a unique 21.7 kDa prolamin seed storage protein gene, having high level of sequence homology withα-zein class of prolamin of *Setaria italic, Setaria viridis* and *Panicum miliaceum.* The amino acid composition of 21.7 kDa prolamin protein indicates the deficiency in essential amino acid lysine. Physiological properties studied through *in silico* analysis can provide new insight into foxtail millet seed storage proteins and utilization for nutritional improvement.

Acknowledgement

The financial support provided by Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi is thankfully acknowledged.

References

1. Day RC, Herridge RP, Ambrose BA, Macknight RC. Transcriptome Analysis of Proliferating Arabidopsis Endosperm Reveals Biological Implications for the Control of Syncytial Division, Cytokinin Signaling, and Gene Expression Regulation. Plant Physiology 2008;148(4):1964-1984. doi: 10.1104/pp.108.128108
2. Doyle J, Doyle J. Isolation of plant DNA from fresh tissue. Focus (San Francisco, Calif.) 1990;12(1):13-15.
3. Holding DR. Recent advances in the study of prolamin storage protein organization and function. Front Plant Sci 2014;5:276. doi: 10.3389/fpls.2014.00276
4. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol 1982:157(1):105-132. doi: 10.1016/0022-2836(82)90515-0

5. Mandal S, Mandal R. Seed storage proteins and approaches for improvement of their nutritional quality by genetic engineering. Current Science 2000, 576-589.
6. Osborne T. Classification of vegetable proteins. The Vegetable Proteins. Longmans, Green and Co.: New York 1924, 25-35.
7. Panaud O. Foxtail millet *Cereals and millets* (pp. 325-332): Springer 2006.
8. Sachdev N, Goomer S, Singh LR. Foxtail millet: a potential crop to meet future demand scenario for alternative sustainable protein. Journal of the Science of Food and Agriculture 2021;101(3):831-842.
9. Shewry PR, Halford NG. Cereal seed storage proteins: structures, properties and role in grain utilization. Journal of experimental botany, 2002;53(370):947-958.
10. Shewry PR, Halford NG, Lafiandra D. Genetics of wheat gluten proteins. Advances in genetics 2003;49:111-184.
11. Xi DM, Zheng CC. Transcriptional regulation of seed storage protein genes in Arabidopsis and cereals. Seed Science Research 2011;21(4):247.
12. Xin X, Lin XH, Zhou YC, Chen XL, Liu X, Lu XX *et al*. Proteome analysis of maize seeds: the effect of artificial ageing. Physiol Plant 2011;143(2):126-138. doi: 10.1111/j.1399-3054.2011.01497.x
13. Xu H, Gao Y, Wang J. Transcriptomic analysis of rice (Oryza sativa) developing embryos using the RNA-Seq technique. PloS one 2012;7(2):e30646.