Journal of Proteomics & Bioinformatics - Open Access

Research Article
OPEN ACCESS Freely available online
doi:10.4172/jpb.1000154
JPB/Vol.3 Issue 10

# Molecular Cloning and Comparative *In silico* Analysis of Calmodulin Genes from Cereals and Millets for Understanding the Mechanism of Differential Calcium Accumulation

**Manoj Nath, Anshita Goel , Gohar Taj and Anil Kumar***

*Department of Molecular Biology and Genetic Engineering, College of Basic Sciences and Humanities, G.B. Pant University of Agriculture and Technology, Pantnagar (Uttarakhand)*

## Abstract

Calmodulin gene (CaM) - a calcium sensor was cloned by the technique of polymerase chain reaction (PCR) from cereals and millets using single set of primer designed from Eleusine coracana ESTs available in dbEST database. The PCR amplicons of 613bp were consistently observed in all cereals and millets except Setaria italica. The amplicons from these cereals were further gel eluted, cloned, sequenced and subjected to homology search, multiple sequence alignment, phylogenetic tree construction and motif analysis. The homology search confirmed their identity to CaM genes except for Triticum aestivum. Multiple sequence alignment of the translated CaM amino acid sequences confirmed the presence of conserved consensus sequences of 110 amino acids, which was uniformly observed across different cereals and millets except Sorghum bicolor. Phylogenetic tree constructed based on protein sequences of cloned CaM genes revealed closed evolutionary relationship between three varieties of Eleusine coracana (PRM-1, PRM-701 and PRM-801) and Hordeum vulgare as compared to other cereals and millets which also possess higher concentration of calcium in seeds. Thus indicates that structural variation in CaM has some role in the differential accumulation of calcium. Three EF-hands associated with a CaM gene were present consistently except Sorghum bicolor, having only two EF-hands. The in silico 3D-structural analysis of cloned sequences showed similar pattern and reveals high degree of conservation in CaM in terms of structure and interaction with calcium ions, thus reflecting to further investigate the role of CaM isoforms, their expression patterns and downstream interaction with transport machinery involved in calcium accumulation.

**Keywords:** CaM; Cereals; Millets; Multiple sequence alignment; Motif; EF-hand

## Introduction

Plant cells are equipped with highly efficient mechanisms to perceive, transduce and respond to a wide variety of internal and external signals during their growth and development. Perception of signals via receptors results in generation or synthesis of non-proteinaceous molecules often termed messengers such as calcium which control diverse cellular processes through calcium sensors which is also known as calcium binding proteins (Bowler and Fluhr, 2000; Reddy, 2001; Snedden and Fromm, 2001; White and Broadley, 2003).The intracellular calcium sensor protein calmodulin (CaM) interacts with a large number of proteins to regulate their biological functions in response to calcium stimulus. This molecular recognition process is diverse in its mechanism, but can be grouped into several classes based on structural and sequence information (Yap et al., 2000). CaM is a small (148 residues) multifunctional acidic protein arranged in two globular domains connected with a long flexible helix. Each globular domain contains a pair of intimately linked EF hands. CaM undergoes conformational changes on binding of four $Ca^{2+}$ ions to their four EF-hand domains. On binding $Ca^{2+}$, hydrophobic surfaces are exposed and allow the $Ca^{2+}$-CaM complex to interact with, and to regulate the activity of, an array of target proteins. The activities of these proteins affect physiological responses to the vast array of specific stimuli received by plant cells (Yang and Poovaiah, 2003). It is a key calcium sensor in all eukaryotes, regulates diverse cellular processes by interacting with other proteins (Reddy et al., 2000).

In plants, CaM regulates a variety of targets including kinases, metabolic proteins, cytoskeletal proteins, ion channels and pumps, and transcription factors (Bouche et al., 2005). CaM is found in the apoplast and in the cytosol, ER and nucleus in plant cells. Within the cytosol, the estimated CaM concentration is 5 to 40 $\mu$M (Zielinski, 1998). CaM is the paradigm for EF-hand proteins that are involved in signal transduction (Kawasaki et al., 1998). In addition to the evolutionarily conserved form of CaM, plants possess an extended family of CaM isoforms and CaM-like of the most interesting differences between $Ca^{2+}$ signaling in plants proteins (Snedden and Fromm, 1998). Primary structures of CaM are generally conserved throughout evolution among various organisms. However, in plants, one striking difference is that numerous isoforms of CaM may occur within a single plant species. A large family of genes encoding CaM isoforms from several plants has been identified including Arabidopsis (*Arabidopsis thaliana*) (Zielinski, 2001), potato (*Solanum tuberosum*) (Takezawa et al., 1995), soybean (*Glycine max*) (Lee et al., 1995) and petunia (*Petunia hybrida*) (Rodriguez-Concepcion et al., 1999). The representative structure of EF-hand of sequenced *Eleusine coracana* CaM clone is shown in Figure 1.

Cereals and millets constituted the staple food of the world since their domestication~10,000 years ago which were most important group of cultivated plants in terms of food production and acreage

***Corresponding author:** Dr Anil Kumar, Professor & Head, Department of Molecular Biology & Genetic Engineering, College of Basic Sciences & Humanities G.B. Pant University of Agriculture & Technology Pantnagar, U.S. Nagar, India, Fax: 05944-233287; E-mail: ak_gupta2k@rediffmail.com

**Received** September 15, 2010; **Accepted** October 11, 2010; **Published** October 13, 2010

**Citation:** Nath M, Goel A, Taj G, Kumar A (2010) Molecular Cloning and Comparative *In silico* Analysis of Calmodulin Genes from Cereals and Millets for Understanding the Mechanism of Differential Calcium Accumulation. J Proteomics Bioinform 3: 294-301. doi:10.4172/jpb.1000154

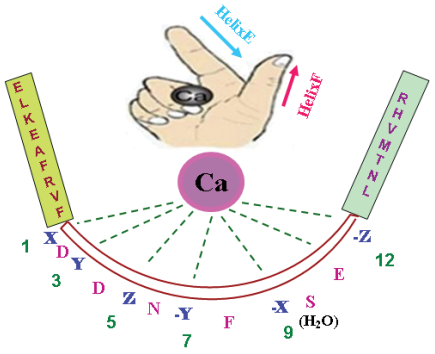


**Figure 1:** EF-hand structure generated from cloned CaM of Eleusine *coracana* based on its amino acid sequence.

covered, providing >60% of the calories and proteins requirement of our daily diet (O'Kennedy et al., 2006; Varshney et al., 2006). The genomes of cereals and millets are very different in terms of size, ploidy level and chromosome number. Despite these significant differences, comparative mapping analysis reveals colinearity at the great level of genetic markers while genes tend to be highly conserved at the DNA sequence level between different cereals as well as millet genomes (Devos and Gale, 2000; Feuillet and Keller, 2002). Several attempts have been made to reveal the existing synteny and colinearity on the basis of comparative genomics (Kellogg, 1998; Paterson et al., 2004; Paterson et al., 2005; Caetano-Anolles, 2005).

*Eleusine coracana*, commonly called finger millet or ragi is one of the important minor cereals cultivated mainly in East Africa and Southern India where it makes an important contribution to food security due to the high nutritional value and good storage ability of its grain. Its improvement has, however, lagged behind that of other crops (Srinivasachary et al., 2007). It is an excellent source of calcium (376-515mg/100g) which is far above the ground then the other cereals and millets (Barbeau and Hilu, 1993). The high calcium level in finger millet prompted us to investigate the factors or molecular mechanism responsible for it accumulation. Both environmental and genetic factors influence calcium accumulation in bean seed (Quenzer et al., 1978). However, little is known about factors influencing distribution of calcium and variation of calcium levels in different cereals and millets.

Therefore an attempt has been made in the present study to characterize the CaM of finger millet along with cereals and millets to identify the structural similarity of CaM genes with their possible role in calcium signaling and calcium accumulation in cereals. This paper reports *in silico* analysis of PCR amplified, cloned and sequenced CaM from finger millet varieties viz. brown, golden and white along with other cereals and millets which were amplified with the same set of primer designed from finger millet EST, of five cereals namely rice (*Oryza sativa*), maize (*Zea mays*), sorghum (*Sorghum bicolor*), barley (*Hordeum vulgare*) and oat (*Avena sativa*), and five millets viz. finger millet (*Eleusine coracana*), barnyard millet (*Echinochloa frumentacea*), proso millet (*Panicum milliaceum* Linn.), little millet (*Panicum antidotale*) and kodo millet (*Paspalum scrobiculatum* Linn.). All nucleotide sequences of cloned CaM were submitted to GenBank database and have received the accession numbers (Table 1).

## Materials and Methods

### Calcium estimation

Calcium content of different cereals and millets was estimated by Atomic Absorption Spectrophotometer (AAS) method. The samples were prepared as described by Barbeau and Hilu (1993). Samples were wet ashed to avoid mineral decomposition and destruction (volatilization) that may occur while ashing in a muffle furnace. Ashed samples were analyzed for calcium by atomic absorption spectrophotometry (AAS) using a Perkin-Elmer 2100 AA Spectrophotometer. Each analysis was done in triplicate.

### PCR amplification and cloning

The finger millet EST sequence (CX264726) was downloaded from finger millet EST database of NCBI (http://www.ncbi.nlm.nih.gov/nucest) and CaM specific primer set EF03F/ Ca0R1 (F-AACTGTCATGCGTTCATTGG, R- GGAGTCGAGCTGATTTGATGA) was designed using Primer-3 online tool (Untergasser et al., 2007). The genomic DNA of different cereals and millets were isolated by standard method (Murray and Thompson, 1980), quantified and analyzed on agarose gel electrophoresis (Maniatis et al., 1989). PCR amplification was performed as per the standard protocol using 50-100 ng of template DNA with 60°C annealing temperature based on Tm value of the primers. The amplified products were analyzed on agarose gel and the expected size amplicons were gel eluted using QIAquick Gel Extraction Kit (Qiagen, USA) and cloned in pGEM-Teasy vector (Promega, USA) as per the kit instructions. Putative cloned *CaM* genes were sequenced using M13 universal primer present in pGEM-T easy vector.

| Sl. No. | Name of Crop(s) (Scientific name) | Cultivar | Sub-family | Sequence length (bp) | Accession Number |
|---|---|---|---|---|---|
| 1 | Finger millet (*Eleusine coracana*) | PRM-1 (Brown) | Chloridoideae | 609 | GQ926851 |
| 2 | Finger millet (*Eleusine coracana*) | PRM-701 (Golden) | Chloridoideae | 615 | GQ926852 |
| 3 | Finger millet (*Eleusine coracana*) | PRM-801 (White) | Chloridoideae | 615 | GQ926853 |
| 4 | Rice (*Oryza sativa*) | Saathi | Oryzoideae | 613 | GQ926854 |
| 5 | Maize (*Zea mays*) | DGPMC4 | Panicoideae | 616 | GQ926855 |
| 6 | Oat (*Avena sativa*) | UP0212 | Pooideae | 614 | GQ926856 |
| 7 | Barley (*Hordeum vulgare*) | DOLMA | Pooideae | 615 | GQ926857 |
| 8 | Sorghum (*Sorghum bicolor*) | Pant chari 5 | Panicoideae | 579 | GQ926858 |
| 9 | Little millet (*Panicum antidotale*) | Local | Panicoideae | 614 | GQ926859 |
| 10 | Kodo millet (*Paspalum scrobiculatum* Linn.) | Local | Panicoideae | 614 | GQ926860 |
| 11 | Proso millet(*Panicum milliaceum* Linn.) | 405 | Panicoideae | 623 | GQ926861 |
| 12 | Barnyard millet (*Echinochloa frumentacea*) | PRB401 | Panicoideae | 613 | GQ926862 |

**Table 1:** List of assigned accession number from Gen Bank of sequenced CaM genes from cereals and millets.

Journal of Proteomics & Bioinformatics - Open Access

www.omicsonline.com

Research Article
OPEN ACCESS Freely available online
doi:10.4172/jpb.1000154
JPB/Vol.3 Issue 10

### *In silico* analysis

The vector sequence was wiped off by VecScreen software. The sequences were further subjected to bioinformatics softwares GENESCAN (Stormo, 2000) and FGENESH (Solovyev et al., 2006) for fishing out the probable genes and the putative CDS and were translated to protein sequence using translation tool (http://ca.expasy.org/tools/dna.html). The translated CaM protein sequences were subjected to protein functional analysis using PFAM version 23.0 (Finn et al., 2006), PROSITE version 20.37 (Castro et al., 2006) and INTERPROSCAN version 4.4 (Quevillon et al., 2005) and SMART version 5.1(a Simple Modular Architecture Research Tool) (Schultz et al., 1998). These sequences of *CaM* proteins from different cereals and millets were aligned using ClustalW (Thompson et al., 1994) and phylogenetic tree was constructed using UPGMA method. A phylogenetic tree was constructed by MEGA 3.1 software (Tamura et al., 2007). The conserved motifs present in these sequences were analyzed using BLOCKS and MEME (Multiple EM for Motif Elicitation) software version 3.5.7 (Bailey et al., 2006). For motif analysis of CaM, the selection of maximum number of motifs was set to 10 with minimum width of 15 amino acids for cloned CaM sequences and 3 with minimum width of 20 amino acids for the comparative analysis with the published CaM sequences which were retrieved from NCBI database while other factors were of default selections. These selections were made in order to minimize the 'Evalue' of the given parameter based on the probability of finding an equally well conserved pattern in set of sequences. Motifs involved with regulatory region of putative CaM genes were analysed using PLACE tool (Higo et al., 1999) and PLANTCARE tool (Lescot et al., 2002).

### Three dimensional structural analyses

For constructing the structures of calmodulin, a template for homology modeling was searched with BLAST program on the Protein Data Bank (www.rcsb.org/pdb/; Berman et al., 2000). Template structure was selected with a cut-off sequence identity of < 40%. Three dimensional structures of calmodulin were modeled using MOE software version 2008.10 (Chemical Computing Group's Molecular Operating Environment) and energies were also minimized using parameters such as Force field= MMFF94X, Gradient= 0.05, Cutoff: On=8, off=10, Solvation: Dielectric=1, Exterior=80. Secondary structure components of the calmodulin sequences were analyzed using SWISS-PDB viewer (http://spdbv.vital-it.ch) and Geno3D (http://geno3d-pbil.ibcp.fr) tools. Binding of Calcium with different EF hands was performed by MOE and HEX Docking software. Model consistency and viability were appraised by PROCHECK software available online (http://www.ebi.ac.uk./Thorton/software.html) for protein structure verification (Laskowski et al., 1993). Analysis of the modeled structures was performed using the RasMol (Raster Display of Molecules) version 2.7.3.1 (Sayle and Milner-White, 1995).

| Sl. No. | Name of Crop(s) | Cultivar | Calcium(mg/100g) |
|---|---|---|---|
| 1 | Finger millet | PRM-1(Brown) | 314.5 |
| 2 | Finger millet | PRM-701(Golden) | 232 |
| 3 | Finger millet | PRM-801(White) | 144 |
| 4 | Rice | Saathi | 12.4 |
| 5 | Wheat | HD2329 | 38.3 |
| 6 | Maize | DGPMC4 | 11.7 |
| 7 | Oat | UP0212 | 14 |
| 8 | Barley | DOLMA | 24.2 |
| 9 | Sorghum | Pant chari 5 | 14.9 |
| 10 | Litle millet | Local | 11.4 |
| 11 | Kodo millet | Local | 23.8 |
| 12 | Proso millet | 405 | 13.6 |
| 13 | Barnyard millet | PRB401 | 12.6 |
| 14 | Foxtail millet | PQK-1 | 27.5 |

**Table 2:** Total calcium content of different cereals and millets.

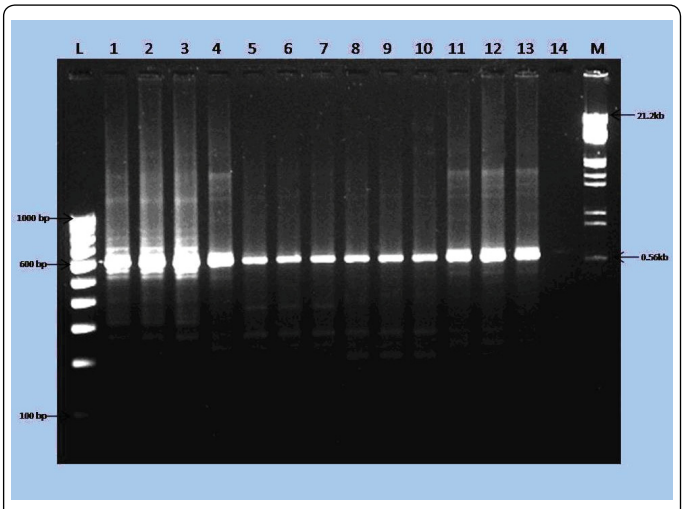


**Figure 2: PCR amplification profiling of cereals and millets using EF-03F/Ca01 R1 primer set** L- 100 bp ladder, Lane 1- Brown, 2-Golden, 3-White, 4-Rice, 5-Wheat, 6- Maize, 7- Oat, 8- Barley, 9- Sorghum, 10-Little millet, 11- Kodo millet, 12- Proso millet, 13- Barnyard millet, 14- Foxtail millet, M-λ DNA / EcoR1 / HindIII Double Digest ladder.

## Results and Discussion

### Determination of calcium content of Cereals and millets

Calcium content was estimated using AAS. The calcium content was highest in finger millet as compared to other cereals and millets. The calcium content in PRM-1 (Brown), PRM-701(golden) and PRM-801 (white) varieties of finger millet was found to be 314.5, 232 and 144 mg/100g respectively, having high calcium as compare to other cereals and millets (Table 2). The results are in general agreement with previous investigators who have reported that finger millet containing much higher calcium than other cereal grains

### PCR amplification of CaM from different cereals and millets

The CaM from different cereals and millets were PCR amplified using same CaM specific primer designed from finger millet ESTs, gel eluted, cloned in pGEMTeasy vector and sequenced. The PCR amplicon of expected size i.e. 613 bp was uniformly observed with all the templates except foxtail millet as shown in Figure 2.

### Comparative *in silico* analysis of *CaM* genes

The *CaM* sequences were subjected to BLAST search for deducing similarity with sequences available in different databases. CaM sequences, range from 579 to 623bp, were obtained for different cereals and millets and this variation in sequence length might be due to amplification of variable length of genomic sequence by same CaM specific primer. These sequences were subjected to homology search to confirm their identity for CaM genes except in *T. aestivum*

Homology search confirm the identity as CaM genes in *E. coracana*, *Z. mays*, *H. vulgare*, *A. sativa*, *E. frumentacea*, *P. antidotale* and *P. scrobiculatum* while as CaM like gene in *O. sativa*, *S. bicolor* *P. milliaceum*. Further, the nucleotide sequences were translated to respective protein sequences using translation tool and subjected to protein Blast to reveal the similarity at protein level with other existing CaM proteins to further confirm their identity as CaM. Beside this the sequences were also confirmed by subjecting the nucleotide sequences to gene finding software GENESCAN and FGENESH. All the

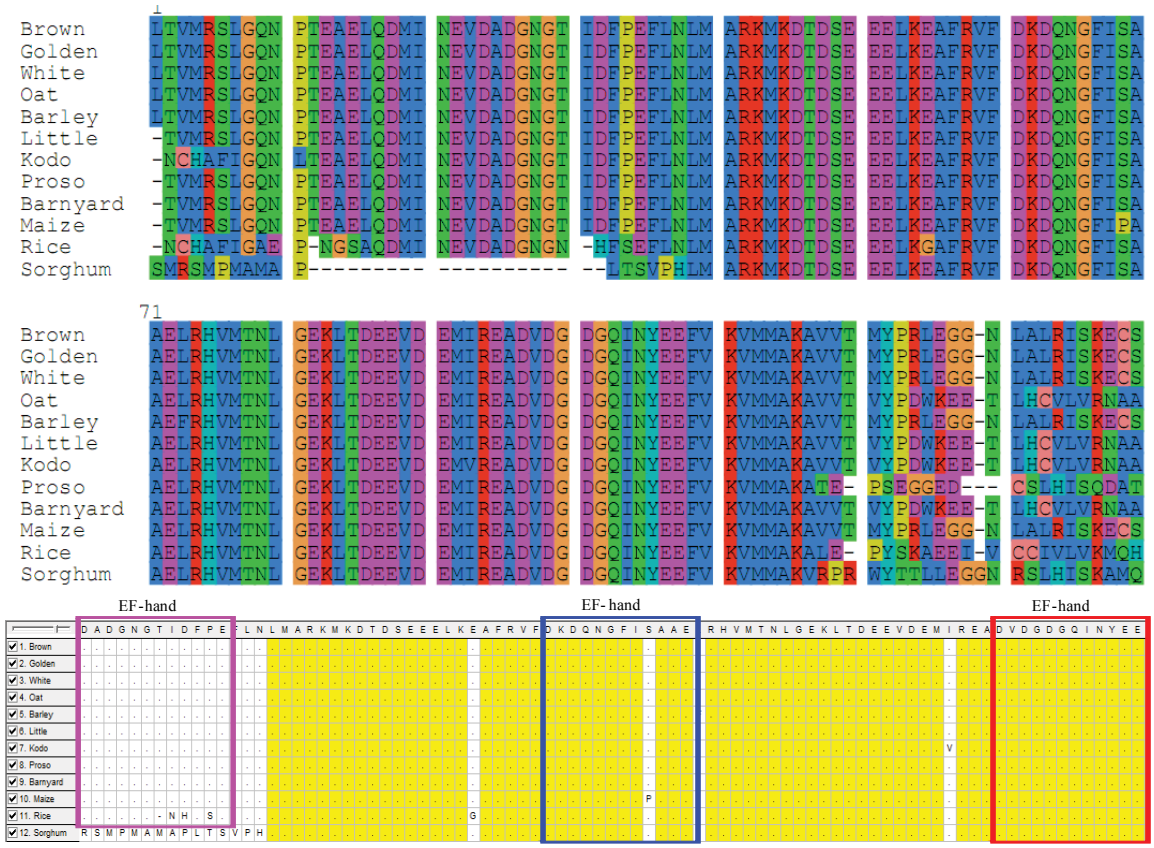


**Figure 3:** Multiple Sequence alignment showing conserved amino acids of the CaM protein sequences from cereals and millets. Conserved EF – hand shown in rectangles and yellow dotted portion indicates the conserved amino acid residues among different CaM sequences.

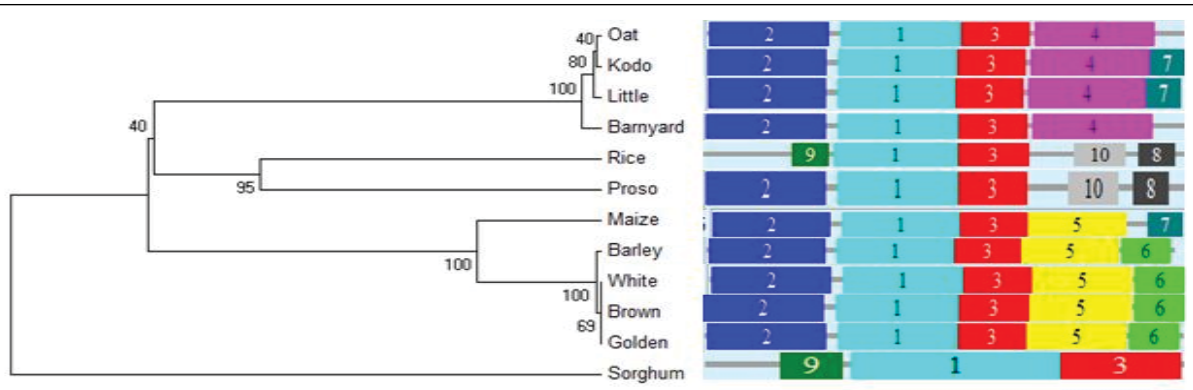


**Figure 4:** Phylogenetic tree based on 12 CaM sequences and schematic distribution of respective conserved motifs using MEME software. Multilevel consensus sequences for the MEME defined motifs are listed in Table 3.

CaM sequences were aligned using Clustal W to find out the extent of similarity amongst the sequences which showed highly conserved (110 aa) sequences associated with three EF-hand except sorghum which is containing only two EF-hand, though the few variations were also observed in sequences of CaM as shown in Figure 3a.

The translated CaM protein sequences subjected to phylogenetic tree construction using a neighbor-joining method, based on amino acid sequence similarity, formed two distinct clusters reveals all cloned CaMs grouped in one cluster except sorghum which is separated from the others (Figure3b).

In all CaMs, motif-1 was most commonly observed which is functionally related to its calcium binding properties while Motif-2 also having similar function as Motif-1 was present in all CaMs except Rice and Sorghum which might be due to the CaM like gene or single set of primer amplify the different region. Beside this Motif-3 also most frequently present in all CaMs which is functionally unknown. Motif-6, which is having Protein kinase C phosphorylation site, observed only in barley and finger millet varieties (Brown, golden and white) while Motif 4 having Kinase binding site is present in Oat, Kodo millet, Little millet, Barnyard millet as shown in Figure 4. The sequences of all motif shown in Table 3.

Research Article
OPEN ACCESS Freely available online
doi:10.4172/jpb.1000154
JPB/Vol.3 Issue 10

On the other hand these cloned CaM sequences were compared with the published CaM sequences of dicots and monocots. CaM sequences retrieved from NCBI and construct Phylogenetic tree using UPGMA method. These sequences were further subjected to MEME program for motif analysis (Figure 5).

The translated protein sequence of cloned CaM gene formed cluster with existing sequences of CaM genes of *O. sativa, Z. mays, A. thaliana, B. napus and B. oleracea*. The phylogenetic tree analysis reveals the different CaMs of dicots and monocots grouped in one cluster and may be more similar during evolution or it may be highly conserved. Motif-1 and Motif-2, functionally related to its calcium binding properties, was most frequently observed in all CaM sequences except sorghum.

Further *in silico* analysis of CaM sequences reveals variation in EF-hand position when these sequences were subjected to various online bioinformatics tool viz. Pfam (Finn et al., 2006) and SMART (Schultz et al., 1998) as shown in Table 4. The results reveal that EF-hand positions were similar in finger millet genotypes, Oat and Barley though distinctly different from the other cereals and other millets. In all *CaM* sequences three EF-hand copies were present except Sorghum where only two EF-hand copies observed which might be due to the amplification of the different region by single set of primer.

The *CaM* nucleotide sequences were further subjected to online bioinformatics tools PLACE and PLANTCARE software for analyzing different motifs present. In all the *CaM* sequences, seed specific *Skn-1* motif, CAAT-box present while the GCN-4 motif present only in sorghum and proso millet. Apart from this, the TATA-Box observed in all the *CaM* sequences of cereals and millets except barley while 5UTR Py-rich stretch region present only in proso millet. The G-Box, I-box, GATA-motif observed in all *CaM* sequences except rice and proso millet. The GAG-motif absent only in maize while present in remaining *CaM* sequences. The results shows structural and functional similarity amongst *CaM* sequences in various cereals and millets in terms of presence and absence of various motifs.

## Comparative three dimensional structure analysis of cloned CaMs of cereals and millets

All CaM protein sequences from cereals and millets were subjected to the MOE software to generate three dimensional structures (Figure 6). The protein sequences of CaMs of all the cereals except *Sorghum bicolor* possess three EF-hands for binding three calcium ions. Presence of two EF hands in *Sorghum bicolor* might be due to the partial CaM sequence obtained by PCR amplification using same set of primers. Structures of all twelve cloned CaMs were modeled using MOE employing the NMR structure of CaM (PDB accession no. 2F3Y.A) as a template and obtained a refined model

| Motif | No. of amino acid | E-value | Multilevel consensus sequence |
|---|---|---|---|
| 1 | 50 | 5.6e-530 | FRVFDKDQNGFISAAELRHVMTNLGEKLTDEEVDEMIREADVDGDGQINY |
| 2 | 50 | 1.3e-368 | TVMRSLGQNPTEAELQDMINEVDADGNGTIDFPEFLNLMARKMKDTDSEE |
| 3 | 29 | 3.4e-170 | EEFVKVMMAKAVVTWYPRWKGGTLHCRIS |
| 4 | 50 | 6.1e-117 | AALVSFVFPEEPVLAVAAAEHLKSPISSVSSNLLLPVICKALLVLDADCG |
| 5 | 42 | 3.1e-119 | KECSSCFLCLPRTRTSCSCCTFEVPYLQCVVEPLASCVNLGF |
| 6 | 21 | 2.8e-035 | CSGRGLRLDPFVVLFLHQISS |
| 7 | 15 | 4.1e-014 | WIRLFYFFIKSARLQ |
| 8 | 15 | 8.4e-002 | VCPGCSWACMLYIFF |
| 9 | 15 | 9.8e+000 | MARKMKDTDSEEELK |
| 10 | 21 | 1.2e+001 | QCFSCYVPPKCVFRDLIFYVN |

**Table 3:** Multilevel consensus sequences for the MEME defined motifs.

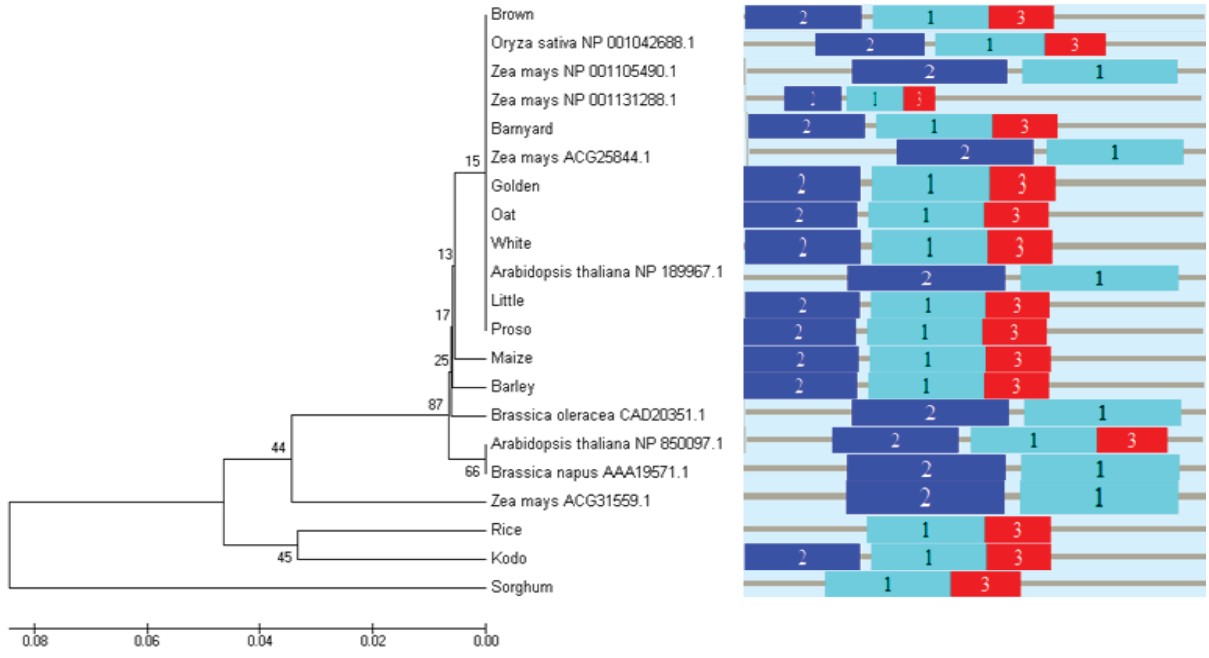


**Figure 5:** Phylogenetic tree generated from cloned CaM sequences and nine available sequences of CaM retrieved through BLAST search. Multilevel consensus sequences for the MEME defined motifs are listed in Table 3.

| Crop | EF- Hand | Begin | End | Sequence | E-value |
|---|---|---|---|---|---|
| Finger millet (Brown) | EFh1 | 15 | 43 | ELQDMINEVDADGNGTIDFPEFLNLMARK | 1.81e-08 |
| | EFh 2 | 52 | 80 | ELKEAFRVFDKDQNGFISAAELRHVMTNL | 2.23e-08 |
| | EFh3 | 88 | 116 | EVDEMIREADVDGDGQINYEEFVKVMMAK | 2.26e-09 |
| Finger millet (Golden) | EFh1 | 15 | 43 | ELQDMINEVDADGNGTIDFPEFLNLMARK | 1.81e-08 |
| | EFh 2 | 52 | 80 | ELKEAFRVFDKDQNGFISAAELRHVMTNL | 2.23e-08 |
| | EFh3 | 88 | 116 | EVDEMIREADVDGDGQINYEEFVKVMMAK | 2.26e-09 |
| Finger millet (White) | EFh1 | 15 | 43 | ELQDMINEVDADGNGTIDFPEFLNLMARK | 1.81e-08 |
| | EFh 2 | 52 | 80 | ELKEAFRVFDKDQNGFISAAELRHVMTNL | 2.23e-08 |
| | EFh3 | 88 | 116 | EVDEMIREADVDGDGQINYEEFVKVMMAK | 2.26e-09 |
| Rice | EFh1 | 13 | 40 | SAQDMINEVDADGNGNHFSEFLNLMARK | 2.35e+00 |
| | EFh2 | 49 | 77 | ELKGAFRVFDKDQNGFISAAELRHVMTNL | 8.18e-07 |
| | EFh3 | 85 | 113 | EVDEMIREADVDGDGQINYEEFVKVMMAK | 2.26e-09 |
| Maize | EFh1 | 14 | 42 | ELQDMINEVDADGNGTIDFPEFLNLMARK | 1.81e-08 |
| | EFh2 | 51 | 79 | ELKEAFRVFDKDQNGFIPAAELRHVMTNL | 3.56e-07 |
| | EFh3 | 87 | 115 | EVDEMIREADVDGDGQINYEEFVKVMMAK | 2.26e-09 |
| Oat | EFh1 | 15 | 43 | ELQDMINEVDADGNGTIDFPEFLNLMARK | 1.81e-08 |
| | EFh 2 | 52 | 80 | ELKEAFRVFDKDQNGFISAAELRHVMTNL | 2.23e-08 |
| | EFh3 | 88 | 116 | EVDEMIREADVDGDGQINYEEFVKVMMAK | 2.26e-09 |
| Barley | EFh1 | 15 | 43 | ELQDMINEVDADGNGTIDFPEFLNLMARK | 1.81e-08 |
| | EFh 2 | 52 | 80 | ELKEAFRVFDKDQNGFISAAELRHVMTNL | 2.23e-08 |
| | EFh3 | 88 | 116 | EVDEMIREADVDGDGQINYEEFVKVMMAK | 2.26e-09 |
| Sorghum | EFh1 | 31 | 59 | ELKEAFRVFDKDQNGFISAAELRHVMTNL | 2.23e-08 |
| | EFh2 | 67 | 95 | EVDEMIREADVDGDGQINYEEFVKVMMAK | 2.26e-09 |
| Little | EFh1 | 14 | 42 | ELQDMINEVDADGNGTIDFPEFLNLMARK | 1.81e-08 |
| | EFh2 | 51 | 79 | ELKEAFRVFDKDQNGFIPAAELRHVMTNL | 3.56e-07 |
| | EFh3 | 87 | 115 | EVDEMIREADVDGDGQINYEEFVKVMMAK | 2.26e-09 |
| Kodo | EFh1 | 14 | 42 | ELQDMINEVDADGNGTIDFPEFLNLMARK | 1.81e-08 |
| | EFh2 | 51 | 79 | ELKEAFRVFDKDQNGFIPAAELRHVMTNL | 2.23e-08 |
| | EFh3 | 87 | 115 | EVDEMIREADVDGDGQINYEEFVKVMMAK | 2.70e-07 |
| Proso | EFh1 | 14 | 42 | ELQDMINEVDADGNGTIDFPEFLNLMARK | 1.81e-08 |
| | EFh2 | 51 | 79 | ELKEAFRVFDKDQNGFIPAAELRHVMTNL | 3.56e-07 |
| | EFh3 | 87 | 115 | EVDEMIREADVDGDGQINYEEFVKVMMAK | 2.26e-09 |
| Barnyard | EFh1 | 14 | 42 | ELQDMINEVDADGNGTIDFPEFLNLMARK | 1.81e-08 |
| | EFh2 | 51 | 79 | ELKEAFRVFDKDQNGFIPAAELRHVMTNL | 2.23e-08 |
| | EFh3 | 87 | 115 | EVDEMIREADVDGDGQINYEEFVKVMMAK | 2.26e-09 |

**Table 4**: *In silico* **comparative analysis of EF-hand domains in CaM from cereals and millets**

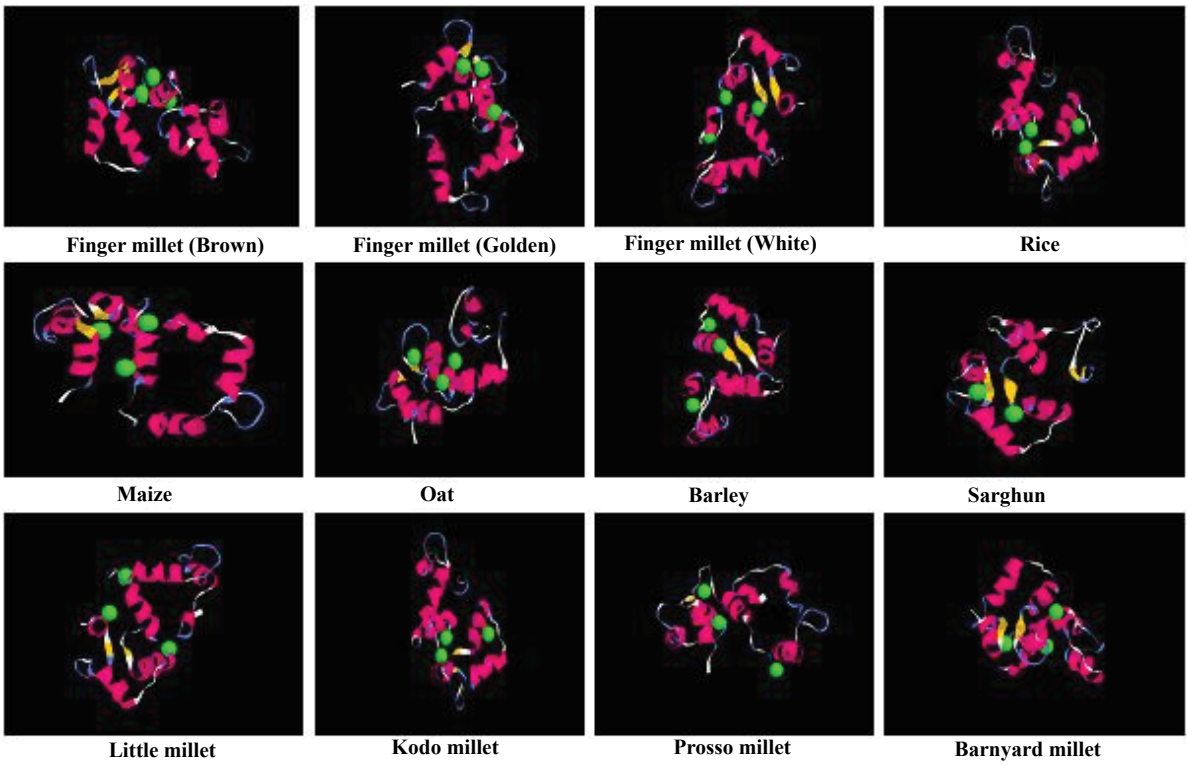


**Figure 6:** 3D structure of CaM from different cereals and millets using the MOE software. The bound calcium ions were shown in green color.

after energy minimization. For each CaM molecule eleven structures were generated in the database, out of these which possess maximum score, was selected. The initial energy of the protein was calculated (in kcal/mol) by GizMOE using MMFF94 X force field. The energies of the designed structures were minimized using the energy minimization tool of MOE.

Research Article
OPEN ACCESS Freely available online
doi:10.4172/jpb.1000154
JPB/Vol.3 Issue 10

The minimized structures were finally saved as **.pdb** files which were validated online by PROCHECK **(data not shown)**. MOE generated structures contained 114, 116, 116, 113, 115, 116, 116, 95, 115, 115, 115 and 115 amino acid residues for Brown, Golden, White, Rice, Maize, Oat, Barley, Sorghum, Little millet, Kodo millet, Proso millet and Barnyard millet respectively. In all cereals and millets the calcium binding loop containing same sequence of amino acid residues which binds with calcium ion except Maize and Rice. The distribution of amino acids in EF-hand, which is involved in binding with calcium, at 1,3,5,7, 9, 12 position are (EFh1 Asp, Asp, Asn, Thr, Asp, Glu, EFh2 – Asp, Asp, Asn, Phe, Ser, Glu, EFh3 – Asp, Asp, Asp, Gln, Asn, Glu) . But in Maize, Proline is present at –X ($9^{th}$) position instead of Serine in EFh2. In Rice calcium binding loop has 11 residues in which calcium binds with Asp, Asp, Asn, Asn, Phe and Glu at position 1, 3,5,7,9 and $11^{th}$ respectively in EF hand1. Sorghum having two EF-hands (EFh2 and EFh3) but the calcium binds with same residues as in others CaM. The interaction energy generated due to the calcium binding is almost similar in all cereals and millets. Brown cultivar of Finger millet has minimum interaction energy among the cereals and millets i.e. 34.2 kcal/mol while golden and white cultivar of finger millet have same interaction energy i.e. 32.0 kcal/mol. Other millets like little, kodo, proso and barnyard have -32.0 kcal/mol, -34.1 kcal/mol, -31.3 kcal/mol and -29.9 kcal/mol interaction energy respectively while other cereals like rice, maize, oat, barley and sorghum have -24.1 kcal/mol, -33.8 kcal/mol, -33.2 kcal/mol, -32.8 kcal/mol and -28.0 kcal/mol respectively. CaMs in cereals and millet were structurally and functionally similar due to the similar pattern of three dimensional structures.

The Arabidopsis genome is predicted to encode 50 CaM-like proteins (CMLs) with sequence identities ranging from about 20% to 75% at the protein level (Mc-Cormack and Braam 2003; Mc-Cormack et al., 2005). Of the 15 currently known CaM genes in *Arabidopsis thaliana*, five have 100% similarity to *CaM2*, and CaM3 and CaM5 calmodulins have 100% *sequence identity* with that of CaM2 (Zielinski, 2004). In the present report the 100% sequence identity was found in CaMs of finger millet varieties viz. PRM-1 (brown), PRM-701(golden) and PRM-801(white).

A maximum 37 CaMs and related potential calcium sensor proteins (five loci were defined as *CaM* genes and thirty two additional CaM-like (CML) genes) were identified in the rice genome. Multiple sequence alignment of the rice CaM amino acid sequences with those of typical CaMs from other species, indicates there is high degree of sequence conservation. They have reported that OsCaM1 amino acid sequences are identical to CaMs of barley and wheat (Boonburapong and Buaboocha, 2007). This paper reports, multiple sequence alignment of the finger millet CaM with the CaMs of other cereals and millets, indicates high degree of sequence conservation though few variations were also observed which might be due to the CaM sequences range from 579 to 623bp. The CaM of finger millet genotypes (brown, golden and white) were identical to CaM of barley which reflecting the close relationships among finger millet and barley as evident from the phylogenetic tree construction with other cereal. Phylogenetic tree result outlines the development of CaM in finger millet and other cereals and millets which indicate that CaM is strictly conserved and has same evolutionary pattern. The availability of genetic map of *E. coracana* (Dida et al., 2007) could be used to investigate the diversity of CaM genes and its possible locations to specific chromosomes in near future. In order to understand the molecular mechanisms associated with calcium accumulation in finger millet, there is also need to identify the CaM isoforms, gene copy number, expression analysis and downstream interaction with target proteins to investigate the genetic mechanism of differential calcium accumulation.

## Conclusion

The PCR based amplification of CaM of different cereals and millets using same primer set designed from finger millet was investigated *in silico* after sequencing of respective PCR amplicons. The homology search, multiple sequence alignment, EF-hand copies, phylogenetic tree construction and motif analysis has clearly revealed the identity of these sequences as CaM of respective cereals and millets. Multiple sequence alignment reveals high degree of sequence conservation in CaM of the cereals and millets though there were some alterations which might be due to the partial sequence ranged from 579 to 623 bp of the CaM obtained in the cereals and millets. The phylogenetic analysis reveals the CaM enjoy common phylogeny in different cereals and millets as well as in dicots viz. *Arabidopsis thaliana, Brassica* spp. The *in silico* 3D-structural analysis of cloned sequences showed similar structures and reveals high degree of conserved CaM in cereals and millets. The paper reports the evolutionary colinearity of CaM in the cereals and millets though the CaM of *Eleusine coracana* and *Hordeum vulgare* having closed evolutionary relationship as compare to others. The CaM gene of *E. coracana* is first report based on *in silico* studies though it has already been reported in *Oryza sativa, T. aestivum, H. vulgare and Z. mays.* Further research is in progress to find out the isoforms of CaM genes in brown, golden and white varieties of *E. coracana* to elucidate its role in signal transduction transport and physiological function including its interaction with target proteins.

### Acknowledgement

The authors acknowledge the facilities provided by DBT-funded SUB-DIC, Bioinformatics Centre, Pantnagar for carrying out in silico investigation. The support provided by Director, Experiment Station, G.B. Pant University of Agriculture and Technology, Pantnagar is also thankfully acknowledged.

### References

1. Bailey TL, Williams N, Misleh C, Li WW (2006) MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res 34: W369–W373.
2. Barbeau WE, Hilu KW (1993) Protein, calcium, iron, and amino acid content of selected wild and domesticated cultivars of finger millet. Plant Foods Hum Nutr 43: 97-104.
3. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The Protein Data Bank. Nucleic Acids Res 28: 235-242.
4. Boonburapong B, Buaboocha T (2007) Genome-wide identification and analyses of the rice calmodulin and related potential calcium sensor proteins. BMC Plant Biol 7: 4.
5. Bouche N, Yellin A, Snedden WA, Fromm H (2005) Plant-specific calmodulin-binding proteins. Annu Rev Plant Biol 56: 435-466.
6. Bowler C, Fluhr R (2000) The role of calcium and activated oxygens as signals for controlling cross-tolerance. Trends Plant Sci 5: 241-246.
7. Caetano-Anolles G (2005) Evolution of Genome Size in the Grasses. Crop Sci 45: 1809-1816.
8. Castro ED, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, et al. (2006) ScanProsite, detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res 34: W362-W365.
9. Devos KM, Gale MD (2000) Genome Relationships, The Grass Model in Current Research. Plant Cell, 12: 637-646.
10. Dida MM, Srinivasachary RS, Bennetzen JL, Gale MD, Devos KM (2007) The genetic map of finger millet, *Eleusine coracana*. Theor Appl Genet 114: 321-332.

11. Feuillet C, Keller B (2002) Comparative genomics in the Grass families, Molecular characterization in the Grass genome structure and evolution. Ann Bot 89: 3-10.

12. Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, et al. (2006). Pfam, clans, web tools and services. Nucleic Acids Res 34: D247-D251.

13. Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis acting regulatory DNA elements (PLACE) database, 1999. Nucleic acid res 27: 297-300.

14. Kawasaki H, Nakayama S, Kretsinger RH (1998) Classification and evolution of EF-hand proteins. Biometals 11: 277-295.

15. Kellogg EA (1998) Relationships of cereal crops and other grasses. Proc Natl Acad Sci USA 95: 2005-2010.

16. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK, a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26: 283-291.

17. Lee SH, Kim JC, Lee MS, Heo WD, Seo HY, et al. (1995) Identification of a novel divergent calmodulin isoform from soybean which has differential ability to activate calmodulin-dependent enzymes. J Biol Chem 270: 21806-21812.

18. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, et al. (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. Nucleic Acids Res 30: 325-327.

19. Maniatis T, Sambrook J, Fritsch EF (1989) Molecular cloning, a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

20. McCormack E , Braam J (2003) Calmodulins and related potential calcium sensors of Arabidopsis. New Phytol 159: 585-598.

21. McCormack E, Tsai YC, Braam J (2005) Handling calcium signaling, Arabidopsis CaMs and CMLs. Trends Plant Sci 10: 383-389.

22. Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8: 4321-4326.

23. O'Kennedy MM, Grootboom A, Shewry PR (2006) Harnessing sorghum and millet biotechnology for food and health. J Cereal Sci 44: 224-235.

24. Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc Natl Aca Sci USA 101: 9903-9908.

25. Paterson AH, Freeling M, Sasaki T (2005) Grains of knowledge, Genomics of model cereals. Genome res 15: 1643-1650.

26. Quenzer NM, Huffman VI, Burns EE (1978) Some factors affecting pinto bean quality. Journal of Food Sciences 43: 1059-1061.

27. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, et al. (2005) InterProScan, protein domains identifier. Nucleic Acids Res 33: W116-W120. PMID: 15980438

28. Reddy AS (2001) Calcium, silver bullet in signaling. Plant Sci 160: 381–404.

29. Reddy AS, Reddy VS, Golovkin M (2000) A calmodulin binding protein from Arabidopsis is induced by ethylene and contains a DNA-binding motif. Biochemical and Biophysical Research Communication 279: 762-769.

30. Rodriguez-Concepcion M, Yalovsky S, Zik M, Fromm H, Gruissem W (1999) The prenylation status of a novel plant calmodulin directs plasma membrane or nuclear localization of the protein. EMBO J 18: 1996-2007.

31. Sayle R, Milner-White EJ (1995) RasMol, biomolecular graphics for all. Trends Biochem Sci 20: 374.

32. Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool, Identification of signalling domains. Proc Natl Acad Sci USA 95: 5857-5864.

33. Snedden WA, Fromm H (1998) Calmodulin, calmodulin-related proteins and plant responses to the environment. Trends in plant Sciences 3: 298-304.

34. Snedden WA, Fromm H (2001) Calmodulin as a versatile calcium signal transducer in plants. New Phytol 151: 35–66.

35. Solovyev V, Kosarev P, Seledsov I, Vorobyev D (2006) Automatic annotation of eukaryotic genes, pseudogenes and promoters. Genome Biol 7: 1-12.

36. Srinivasachary RS, Dida MM, Gale MD, Devos KM (2007) Comparative analyses reveal high levels of conserved colinearity between the finger millet and rice genomes. Theor Appl Genet 115: 489-499.

37. Stormo GD (2000) Gene-Finding Approaches for Eukaryotes. Genome Res 10: 394-397.

38. Takezawa D, Liu ZH, An G, Poovaiah BW (1995) Calmodulin gene family in potato, developmental and touch-induced expression of the mRNA encoding a novel isoform. Plant Mol Biol 27: 93-703.

39. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4, Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. Mol Biol Evol 24: 1596-1599.

40. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W, improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680.

41. Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, et al. (2007) Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res 35: W71-W74.

42. Varshney RK, Hoisington AD, Tyagi KA (2006) Advances in cereal genomics and applications in crop breeding. Trends Biotechnol 24: 1-10.

43. White PJ, Broadley MR (2003) Calcium in plants. Ann Bot 92: 487–511.

44. Yang T, Poovaiah BW (2003) Calcium/calmodulin-mediated signal network in plants. Trends in Plant Sciences 8: 505-512.

45. Yap KL, Kim J, Truong K, Sherman M, Yuan T, et al. (2000) Calmodulin Target Database. J Struct Funct Genomics 1: 8-14.

46. Zielinski RE (1998) Calmodulin and calmodulin-binding proteins in plants. Annu Rev Plant Physiol Plant Mol Biol 49: 697–725.

47. Zielinski RE (2001) Characterization of three new members of the Arabidopsis thaliana calmodulin gene family, conserved and highly diverged members of the gene family functionally complement a yeast calmodulin null. Planta 214: 446-455.

48. Zielinski RE (2004) http://www.life.uiuc.edu/zielinski/Cam%20clones.htm.