Available online at www.sciencedirect.com
ScienceDirect
www.sciencedirect.com/science/journal/16720229
GENOMICS PROTEOMICS & BIOINFORMATICS

Article

Genome-Wide Comparative *in silico* Analysis of Calcium Transporters of Rice and Sorghum

Anshita Goel[1,2], Gohar Taj[1], Dinesh Pandey[1], Sanjay Gupta[3], and Anil Kumar[1*]

[1] *Department of Molecular Biology and Genetic Engineering, College of Basic Sciences and Humanities, G.B. Pant University of Agriculture and Technology, Pantnagar 263145, India;*
[2] *Department of Biotechnology, Uttrakhand Technical University, Dehradun 248007, India;*
[3] *Department of Biotechnology, SBS Biomedical P.G. Institute of Biomedical Science and Research, Dehradun 248161, India.*

Genomics Proteomics Bioinformatics 2011 Oct; 9(4-5): 138-150 DOI: 10.1016/S1672-0229(11)60017-X

Received: Feb 07, 2011 Accepted: Sep 04, 2011

Abstract

The mechanism of calcium uptake, translocation and accumulation in Poaceae has not yet been fully understood. To address this issue, we conducted genome-wide comparative *in silico* analysis of the calcium (Ca^{2+}) transporter gene family of two crop species, rice and sorghum. Gene annotation, identification of upstream *cis*-acting elements, phylogenetic tree construction and syntenic mapping of the gene family were performed using several bioinformatics tools. A total of 31 Ca^{2+} transporters, distributed on 9 out of 12 chromosomes, were predicted from rice genome, while 28 Ca^{2+} transporters predicted from sorghum are distributed on all the chromosomes except chromosome 10 (Chr 10). Interestingly, most of the genes on Chr 1 and Chr 3 show an inverse syntenic relationship between rice and sorghum. Multiple sequence alignment and motif analysis of these transporter proteins revealed high conservation between the two species. Phylogenetic tree could very well identify the subclasses of channels, ATPases and exchangers among the gene family. The *in silico cis*-regulatory element analysis suggested diverse functions associated with light, stress and hormone responsiveness as well as endosperm- and meristem-specific gene expression. Further experiments are warranted to validate the *in silico* analysis of the predicted transporter gene family and elucidate the functions of Ca^{2+} transporters in various biological processes.

Key words: *in silico* comparison, calcium transporter, rice, sorghum, genomic annotation, synteny

Introduction

Calcium plays a central role in plants in the regulation of growth and development by acting as a second messenger in the signal transduction pathways. The membrane-associated calcium (Ca^{2+}) transporter proteins are essential to maintain calcium homeostasis under normal physiological conditions and provide tolerance against various stresses. There are three major classes of Ca^{2+} transporter proteins: channels, ATPases (pumps) and exchangers (*1-6*). Members of these transporter proteins may differ in their cellular and tissue distribution and the regulation by other signaling pathways (*7, 8*). The spatial and temporal regulation of calcium concentration in plant cells depends on the coordinated activities of channels and active transporters located on different organelles and membranes (*9*). Calcium channels, pumps and ex-

*Corresponding author.
E-mail: anilkumar.mbge@gmail.com

changers, which differ in their cellular distribution and mechanism of transport, operate the complex and tight regulation of calcium homeostasis. Specific isoforms of these proteins are responsible for increasing or reducing free calcium in the cytosol (*10*).

Whereas diffusion of molecules across membranes (either intracellular or across the plasma membrane) is mediated by calcium channels, the calcium pump (ATPase) is a membrane-bound Ca^{2+} transporter that uses energy derived from ATP hydrolysis to transport Ca^{2+} across membranes against their concentration gradient (*11*). There are two major Ca^{2+} ATPase families (*12, 13*): P-type IIA and P-type IIB. The P-type IIA family lacks N-terminal auto-regulatory domain while the IIB family of plant is characterized by the presence of an auto-inhibitory N-terminal domain containing a Ca^{2+}/CaM-binding site and a serine phosphorylation site (*14*). Calcium exchanger is a secondary active transporter using energy from the flow of one ion (for example, Na^{+}) down its concentration gradient to transport Ca^{2+} against its concentration gradient (*10*). Ca^{2+} transporters on various membranes play an important role in orchestrating diverse biological processes. The results of electrophysiological studies and molecular analyses indicate the existence of many species of Ca^{2+} transporter proteins (*15, 16*). Although a burgeoning number of Ca^{2+} transporters have been identified, it is often difficult to associate functions with particular transporters. Thus, there is a dire need of in-depth understanding of structural and functional roles of such transporter family, which can be elucidated by comparing the Ca^{2+} transporter gene family in the grass species of rice and sorghum.

Grasses, covering 20% of the earth's land surface, are ecologically well adapted (*17*). This group of plants is especially important to agriculture, contributing a large portion of the calories consumed in the human diet (*18*). Both rice and sorghum belong to the Poaceae family and represent two types of carbon metabolism: C3 and C4, respectively. The evolution of C4 photosynthesis in the Sorghum lineage involved redirection of C3 progenitor genes as well as recruitment and functional divergence of both ancient and recent gene duplicates (*19*). Analysis of the sorghum sequence provides new insights into the recruitment of C3 genes to the C4 pathway, allowing us to identify more orthologs in cereals.

Though there exists great diversity among cereals in terms of genome size, ploidy level and chromosome numbers, attempts have been made to reveal the existing synteny and colinearity on the basis of comparative genomics (*19*). A total of 19,929 (57.8%) sorghum gene models were collinear with rice. About 60% of sorghum genes are located in syntenic regions to rice and orthologous relationships are well established by genetic markers as well as whole-genome comparisons (*19, 20*). The identification of specific genes and their function would help to understand the development, evolution and differences of crop plants. Additionally, the comparative analysis of sorghum and rice will also offer new insights into the role of different Ca^{2+} transporters like calcium channel, pump and exchanger. Furthermore, examination of comparative synteny mapping between two monocot model species will allow understanding how genes evolve within monocot species and thus allow better identification of target genes for Ca^{2+} signaling. In view of the un-availability of sufficient information on the molecular mechanism(s) associated with calcium nutrition, transport and accumulation in cereal grains, we focused on comparing the Ca^{2+} transporter gene family in the grass species rice and sorghum.

In the present study, we performed genome-wide *in silico* identification of Ca^{2+} transporter gene family of sorghum from its recently sequenced genome (*19*), with annotation for chromosomal location(s), gene structure and phylogenetic tree construction. The putative functions of the predicted Ca^{2+} transporter genes in sorghum were investigated by analyzing the *cis*-regulatory elements present in the promoter region of these genes. Moreover, the comparative phylogeny and syntenic mapping with rice transporter gene family were also analyzed.

Results and Discussion

Genome-wide annotation of Ca^{2+} transporter gene family

The availability of sequenced genome of sorghum (*19*) provided ample opportunity to annotate and predict the complete set of information of Ca^{2+}

transporter genes using various bioinformatics tools. For the identification of transporter genes, BLAST (*21, 22*) search was conducted to analyze the complete genome sequence of rice and sorghum, respectively. The genome-wide search revealed presence of 28 transporter genes from sorghum genome, including 1 channel, 15 ATPases and 12 exchangers (**Table 1**). In comparison, 31 transporter genes were found as annotated from the whole genome of rice, consisting of 1 channel, 14 ATPases and 16 exchangers (**Table 2**). It shows that fewer Ca^{2+} transporters were identified in sorghum than in rice. This could be due to the fewer chromosomes in sorghum (10 chromosomes) as compared to rice (12 chromosomes).

Multiple sequence alignment

To investigate the sequence features of Ca^{2+} transporter proteins in rice and sorghum, we performed multiple sequence alignment of amino acid sequences for the 59 identified Ca^{2+} transporter proteins (Figure S1). Sequence alignment of calcium channel, ATPase and exchanger by ClustalW indicated an overall homology amongst their own respective groups. High conservedness was observed in calcium channel between sorghum and rice. In all Ca^{2+} ATPases, alignment revealed one conserved motif, DKTGTLT (highlighted in Figure S1B). This conserved region locates on motif 2 in Ca^{2+} ATPase, which contains E1-E2 ATPase phosphorylation site as shown in **Table 3**.

Table 1 Calcium transporters in sorghum

Serial No.	Locus	Co-ordinate*	No. of conserved domains	Chromosome No.	Function
1	Sb03g031110	59449940 - 59431785	2	3	Two Pore Ca channel
2	Sb03g045370	72613513 - 72619278	6	3	Ca ATPase 3
3	Sb08g001260	1231390 - 1237532	6	8	Ca ATPase, IIB type
4	Sb05g002380	2551588 - 2556879	6	5	Ca ATPase, IIB type
5	Sb01g014620	13926528 - 13931352	5	1	Ca ATPase 2
6	Sb09g024300	53852476 - 53858301	6	9	Ca ATPase 6
7	Sb01g038990	62459145 - 62464238	5	1	Ca ATPase, IIB type
8	Sb07g026810	61947637 - 61958992	6	7	Ca ATPase, IIB type
9	Sb06g027770	56594227 - 56606829	6	6	Ca ATPase, IIB type
10	Sb09g001850	1851129 - 1855578	5	9	Ca ATPase, IIB type
11	Sb07g028160	63171206 - 63175489	6	7	Ca ATPase, IIB type
12	Sb01g043620	66764553 - 66770308	6	1	Ca ATPase, IIB type
13	Sb02g028935	64048135 - 64069090	6	2	Ca ATPase, IIB type
14	Sb01g021870	26371223 - 26374261	4	1	Ca ATPase, IIB type
15	Sb04g005130	4956445 - 4976039	3	4	Ca ATPase, IIB type
16	Sb06g029175	57842933 - 57844648	1	6	Ca ATPase1, IIA type
17	Sb04g010130	13357947 - 13369114	2	4	CAX 1c
18	Sb04g003135	2940231 - 2942365	2	4	Ca exchanger 2
19	Sb01g033220	56450631 - 56454742	2	1	Ca exchanger
20	Sb01g021270	24787055 - 24788792	1	1	Ca exchanger
21	Sb08g022240	54041794 - 54043632	1	8	Na/Ca exchanger
22	Sb06g031080	59454144 - 59458446	2	6	Ca exchanger
23	Sb03g008600	9201460 - 9203592	1	3	K-dependent Na/Ca exchanger
24	Sb03g013753	17913353 - 17913868	1	3	Na/Ca exchanger
25	Sb09g030750	59382798 - 59385885	1	9	CAX 1
26	Sb04g011451	17193273 - 17194308	1	4	Ca exchanger
27	Sb01g012910	11968709 - 11970895	1	1	Ca exchanger
28	Sb04g008850	10445502 - 10462382	1	4	Ca exchanger

*Note: "Co-ordinate" indicates the co-ordinates (5'-3') of coding sequences in a chromosome.

Table 2 Calcium transporters in rice

Serial No.	Locus	Co-ordinate	No. of conserved domains	Chromosome No.	Function
1	Os01g48680	27919763 - 27906603	2	1	Two Pore Ca Channel
2	Os01g71240	41219326 - 41225966	5	1	Ca ATPase 11,IIB type
3	Os02g08018	4208367 - 4203438	2	2	Ca ATPase 10, IIB type
4	Os03g17310	9640836 - 9635147	4	3	Ca ATPase 2, IIA type
5	Os03g42020	23340016 - 23345128	4	3	Ca ATPase 2, IIB type
6	Os03g52090	29903956 - 29882993	5	3	Ca ATPase 3, IIA type
7	Os03g10640	5433293 - 5438896	5	3	Ca ATPase 2, IIB type
8	Os04g51610	30399000 - 30387339	5	4	Ca ATPase 9, IIB type
9	Os05g41580	24284626 - 24291437	6	5	Ca ATPase 4, IIB type
10	Os05g02940	1083959 - 1085686	2	5	Ca ATPase 2, IIA type
11	Os08g40530	25656887 - 25650557	2	8	Ca ATPase 9, IIB type
12	Os10g28240	14613997 - 14610890	4	10	Ca ATPase 13, IIB type
13	Os11g04460	1858363 - 1864021	4	11	Ca ATPase 4, IIB type
14	Os12g04220	1782572 - 1788562	5	12	Ca ATPase 4, IIB type
15	Os12g39660	24452031 - 24457112	5	12	Ca ATPase 2, IIB type
16	Os01g11414	6133239 - 6137905	5	1	Ca exchanger
17	Os01g37690	21075484 - 21071420	1	1	vacuolar cation/proton exchanger 1a
18	Os02g04630	2070492 - 2075369	2	2	Na/Ca exchanger 4
19	Os02g21009	12432344 - 12448156	2	2	vacuolar cation/proton exchanger 1c
20	Os02g43110	25952315 - 25951872	2	2	Na/Ca exchanger 1
21	Os03g01330	235750 - 234794	1	3	Na/K/Ca exchanger 1
22	Os03g08230	4195311 - 4193392	0	3	Ca exchanger
23	Os03g27960	16060007 - 16063229	2	3	vacuolar cation/proton exchanger 2
24	Os03g45370	25612199 - 25613926	2	3	Ca exchanger
25	Os04g55940	33130222 - 33133575	2	4	vacuolar cation/proton exchanger 3
26	Os05g51610	29530679 - 29526503	2	5	vacuolar cation/proton exchanger 1b
27	Os10g30070	15542987 - 15544780	2	10	Na/K/Ca exchanger 6
28	Os11g01580	334289 - 333096	2	11	Na/K/Ca exchanger 6
29	Os11g05070	2216489 - 2213647	1	11	Na/K/Ca exchanger 6
30	Os11g43860	26019832 - 26017041	1	11	Ca exchanger
31	Os12g42910	26635589 - 26637352	2	12	Ca exchanger

However, other motifs in Ca^{2+} transporters show degeneracy in various regions of protein sequences. In calcium exchangers, only a few conserved amino acids were observed between sorghum and rice.

Phylogenetic tree

To examine the phylogenetic relationship among transporters in rice and sorghum, a rooted tree was constructed by aligning their amino acid sequences (**Figure 1**). As is attested by the bootstrap values quoted on the nodes, the phylogenetic analysis of transporters across both species clearly reveals four groups: (A) P-type IIB Ca^{2+} ATPases, (B) calcium exchangers, (C) calcium channels and (D) P-type IIA Ca^{2+} ATPases. The difference (sequence-wise and hence in domains) between Ca^{2+} exchangers and Ca^{2+} ATPases is evident from the segregation between their respective clusters. Ca^{2+} ATPases were further separated as P-type IIA and IIB on the basis of their amino

Table 3 Multilevel consensus sequences of the motifs defined by MEME software

Motif	Width	E value	Multilevel consensus sequences
1	50	9.3e-980	MGIQGTEVAKESSDMIIMDDNFSTIVNVARWGRSVYNNIQKFIQFQLTVN
2	50	4.5e-920	PLAVTLCLAFAMKKMMNDKALVRHLSACETMGSATCICSDKTGTLTTNHM
3	44	5.0e-753	TAVQLLWVNMIMDTLGALALATEPPNDNMMKRPPVGRREPFITN
4	50	4.1e-822	YTCIGIVGIKDPCRPGVKDAVETCMSAGIKVRMVTGDNINTAKAICRECG
5	50	7.5e-635	YKQSLQFKHLDKEKKKIQVQVTRDGYRQKVSIYDLVPGDIVHLKIGDQVP
6	28	8.5e-476	RKMFGHVVAVTGDGTNDAPALHEADIGL
7	36	4.9e-454	GTKVQDGYCKMLVTAVGMRTEWGKLMATISEDGDDE
8	39	1.4e-507	CATVSLYFCIATEGWPKGWYDPPGIIGSILLPVMNTAPS
9	50	2.4e-576	VFNEFNSREMEKINVFRGIFKNWIFMGIIAITVVFQFIIIEFLGKFANTV
10	39	6.7e-451	AIEGPEFREKSPEEMRELIPKIQVMARSSPNDKHTLVKH

acid composition. In spite of being very few in number, as compared to IIB Ca^{2+} ATPases (Cluster A) and calcium exchangers (Cluster B), IIA Ca^{2+} ATPases do form a separate cluster with significant bootstrap values (Cluster D). It is also noteworthy that the two calcium channels (Cluster C) are a subset of the calcium exchangers (Cluster B). This is expected since calcium channels and exchangers share much structural and functional similarity. In addition, most of the members belonging to the same cluster also share one or more conserved domains. Therefore, a majority of the Ca^{2+} transporters in rice are expected to be functional orthologs of the Ca^{2+} transporters in sorghum. Furthermore, all clusters contain members from the two different species, although rice is a C3 plant whereas sorghum has C4 metabolic pathway. The Ca^{2+} transporters of these two plants cluster together with respect to their orthology, which suggests that the structure and function of most of these genes might have remain conserved during evolutionary timescale across the plants employing C3 or C4 metabolic pathway.

Chromosomal distribution of Ca^{2+} transporters and the synteny between rice and sorghum

At the whole-genome level, there is some colinearity between rice and sorghum. The orthologs from Ca^{2+} transporter gene families in both rice and sorghum were mapped and their corresponding chromosome locations are summarized in Tables 1, 2 and **Figure 2**. The comparative synteny map also revealed the expansion and inversion of some chromosome regions in sorghum. The genome sizes of the grass family vary dramatically from 400 Mb in rice to 730 Mb in sorghum. In rice, the 31 identified Ca^{2+} transporter genes were found to be distributed on 9 out of 12 chromosomes, while in sorghum the 28 Ca^{2+} transporters were distributed on all the chromosomes except chromosome 10 (Chr 10). In the case of rice, maximum 8 (4 ATPases and 4 exchangers) transporters were found to be located on Chr 3, followed by 4 on Chr 1, 2 and 11, respectively, and 3 on each of Chr 5 and 12. Chr 10 contains 2 Ca^{2+} transporters (1 ATPase and 1 exchanger) while Chr 8 contains only 1 ATPase. No Ca^{2+} transporters were found on Chr 6, 7 and 9. In the case of sorghum, maximum 7 transporters (4 ATPases and 3 exchangers) were found to be located on Chr 1, followed by 4 on each of Chr 3 and 4, and 3 on each of Chr 6 and 9, while there is only 1 ATPase on Chr 2. Most of the Ca^{2+} ATPases and exchangers are distributed on Chr 1, 3 and 4 in sorghum.

Alignment of the sorghum genome map with those of other cereals has revealed extensive macro colinearity, especially between sorghum and rice (*23-26*). However, the distribution of the Ca^{2+} transporter genes on the 9 chromosomes in both sorghum and rice is not uniform, and some transporter genes are found in clusters in various regions on certain chromosomes. Most of the genes located on Chr 1 in sorghum show syntenic relationship with those located on Chr 3 in rice, while genes on Chr 3 in sorghum are syntenic with those on Chr 1 of rice. Similar trend was observed for Ca^{2+} transporter genes in this study. Calcium channel and ATPase genes located on Chr 3 in sorghum show synteny with those on Chr 1 in rice. Conversely, all the Ca^{2+} transporters on Chr 3 in rice show the syntenic relationship with those on Chr 1 in

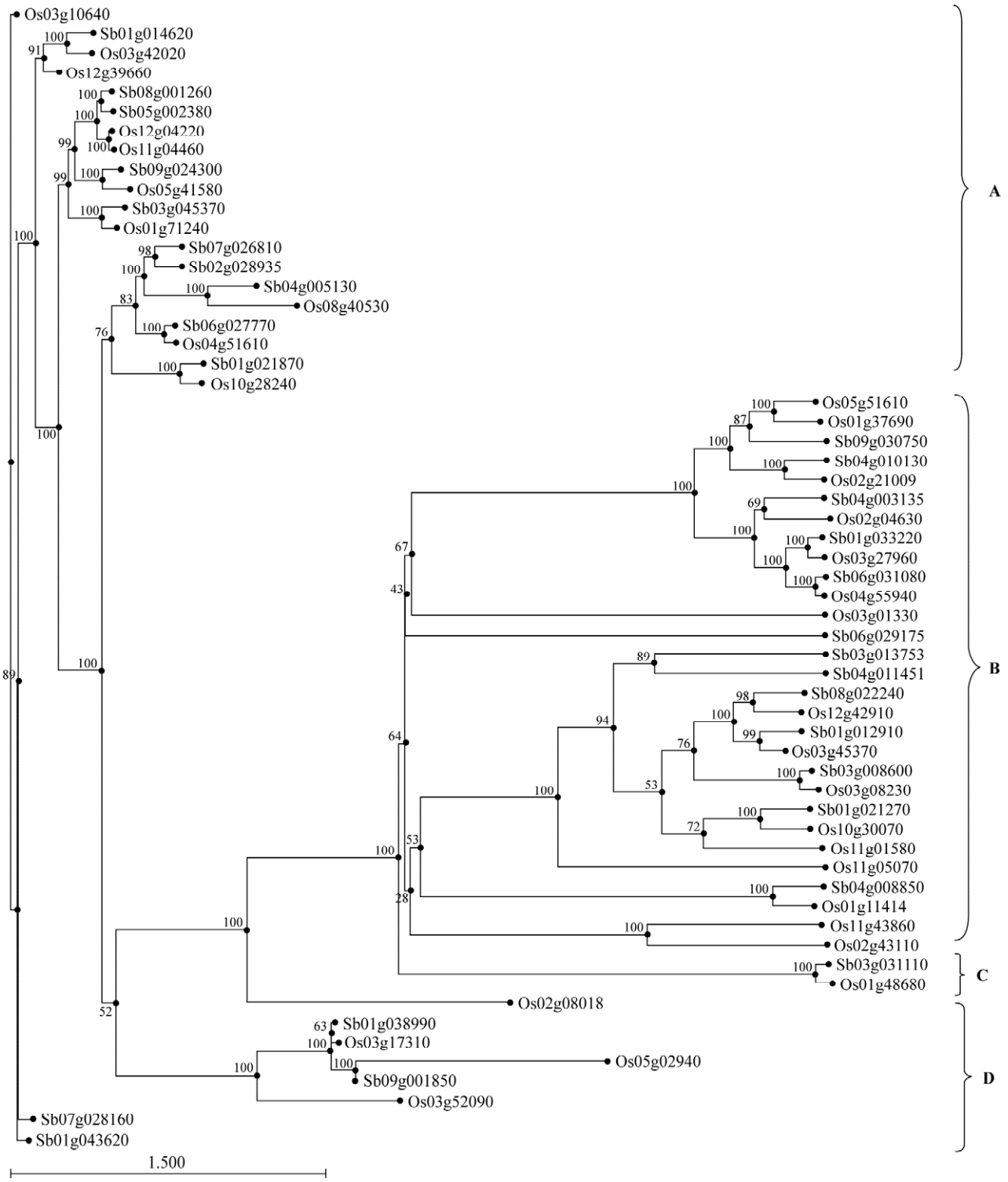

Figure 1 Phylogenetic tree of calcium transporter proteins from rice and sorghum. The phylogenetic tree was constructed by unweighted pair group method with arithmetic mean (UPGMA) using MEGA version 4.0.02. The calcium transporters were classified into four groups: A. IIB type calcium ATPases, B. calcium exchanger, C. calcium channel and D. IIA type calcium ATPases.

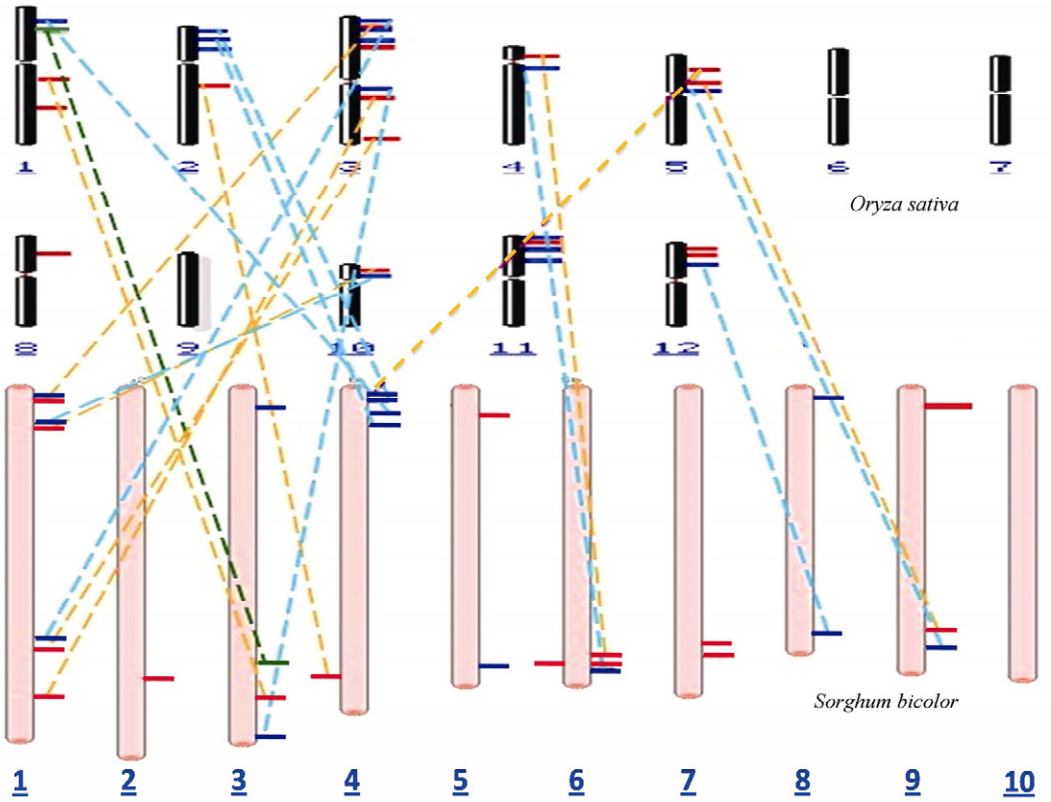

Figure 2 Genomic distribution and syntenic mapping of calcium transporter genes. Chromosomes from rice and sorghum were drawn in black and pink with length to approximate scale. The distribution of calcium channel (green), ATPase (red) and exchanger (blue) on different chromosomes in rice and sorghum was depicted. In addition, the synteny of calcium channel, ATPase and exchanger genes between rice and sorghum was indicated by the dashed lines in dark green, brown and cyan, respectively.

sorghum except one Ca^{2+} ATPase gene, which shows synteny with Ca^{2+} ATPase on Chr 3 in sorghum. In addition, syntenic relationship was also noticed between one calcium exchanger present on Chr 1 in rice and that on Chr 4 of sorghum, while one ATPase and one exchanger found on Chr 10 in rice show syntenic relationship with those on Chr 1 in sorghum. Furthermore, Ca^{2+} transporter genes on Chr 2 in rice show syntenic relationship with those on Chr 4 in sorghum, and one ATPase and one exchanger present on Chr 4 in rice show synteny with those on Chr 6 in sorghum, respectively. These Ca^{2+} transporters with colinearity also share some common motifs and much structural and functional similarity.

Therefore, the rice–sorghum syntenic map provides a reliable and well documented source of data that could be exploited in the future to locate the most conserved syntenic segments, which will also be well positioned for in-depth genome sequence analysis and comparative analysis of the role of Ca^{2+} transporters in other cereals.

Analysis of conserved motifs

Every calcium ATPase contains all 10 motifs while calcium exchangers contain only a few motifs (**Figure 3**), out of which two exchangers (1 in rice and 1 in sorghum) contain only one single motif that appears twice. Motif 8 is found in all the calcium transporters, which is functionally related to the N-myristoylation site. Motifs 1 and 2 are also most commonly observed in all transporters, containing N-glycosylation site and E1-E2 ATPase phosphorylation site, respectively. Motif 3 contains cation transporting ATPase motif, which is present in only a few exchangers but completely absent in calcium channels. Motif 4 containing both N-myristoylation site and protein kinase C phosphorylation site is present in a few exchangers from both species and in channels from sorghum, but is absent in channels from rice. On the other hand, Motif

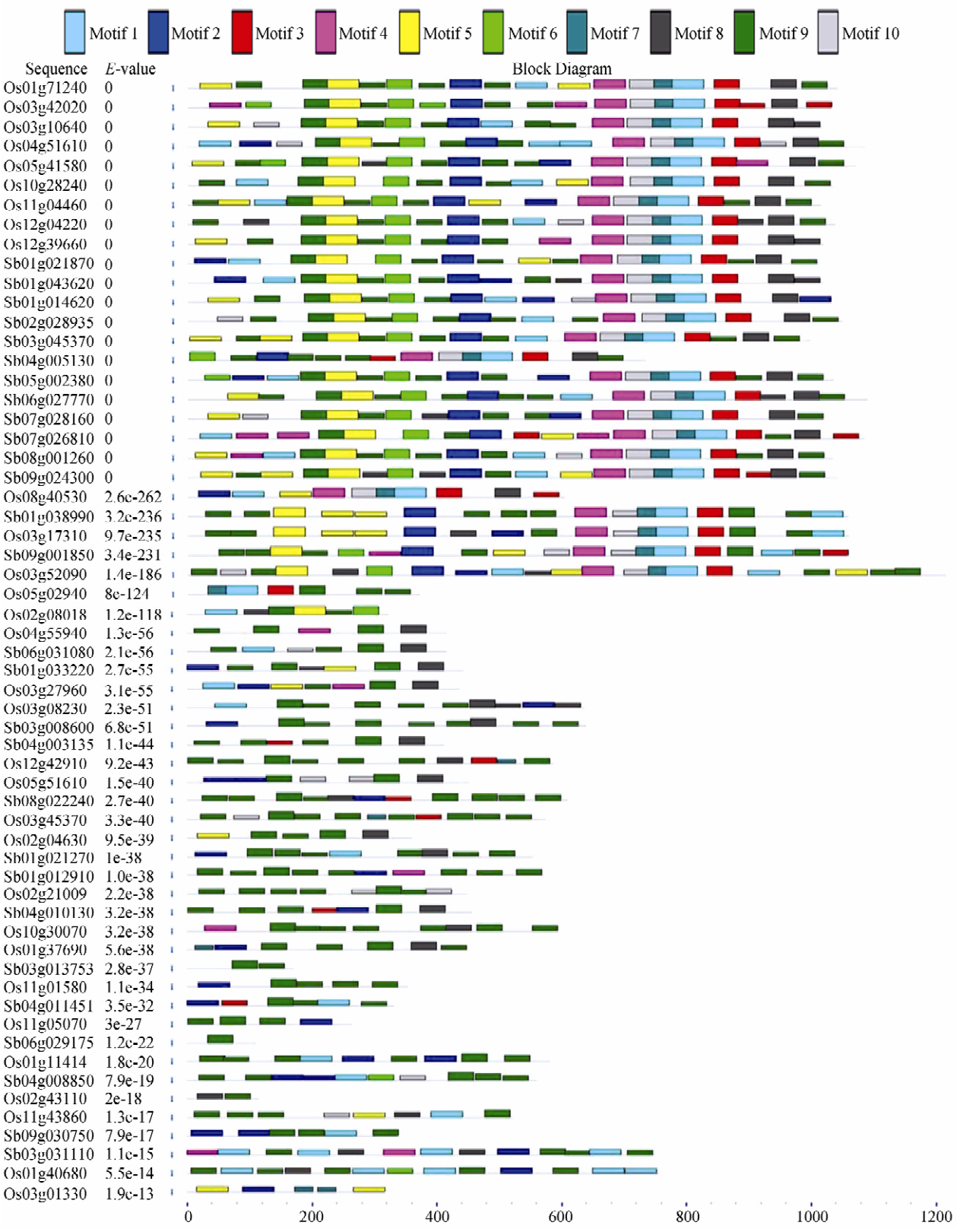

Figure 3 Block diagram of multilevel consensus sequences for the MEME defined motifs of transporter proteins. Ten motifs were obtained using MEME software. Different motifs are indicated by filled boxes with different colors and numbered 1 to 10 in order. Names of all different proteins and E values are shown to the right of the image. Scale bar below the image indicates the relative sizes of the motifs.

5 containing casein kinase II phosphorylation site is present in channels from rice but absent in those from sorghum. In addition, Motif 6 containing CHAP domain (cysteine, histidine-dependent amidohydrolase/peptidase) is missing in all channels and exchangers except two (Os03g45370 and Os03g01330) from rice. Furthermore, Motifs 7 and 10 are commonly observed in all transporters, both have casein kinase II phosphorylation site. Motif 9, which contains 2 functional sites, N-myristoylation site and C-terminus cation transporting ATPase, was observed in all calcium transporters. Multilevel consensus sequences for the motifs defined by MEME software is shown in Table 3.

Cis-regulatory elements analysis

The *cis*-regulatory element analysis was carried out by retrieving nucleotide sequences 1,000 bp upstream of the initiation codon of 31 and 28 transporter genes from rice and sorghum, respectively. A large number of *cis*-regulatory elements were detected when all transporter genes were subjected to search against PlantCARE database (*27*). Distribution of the *cis*-regulatory elements among these transporter genes was shown in **Figure 4**. These elements were mainly associated with five important physiological processes, including light responsiveness, stress responsiveness, hormone responsiveness, meristem-specific expression and seed-specific gene expression. In addition, elements associated with other functions are placed in another category.

Many light-responsive elements were revealed in the promoter regions of all transporter genes from rice and sorghum, including LAMP, ACA motif, ATCT, AT1, TCCC, AAAC, GAG, GT1, TCT, GATA motif, Sp1, G Box, I Box and AE Box. These elements suggest that the calcium transporter proteins might be involved in regulation of photoperiodic control during flowering. The frequency of light-responsive elements in transporter gene promoter region ranges from 7 to 12 and the highest number was found in the promoter region of Os01g48680. In sorghum, calcium channel contains all four seed-specific elements, which were also detected in one ATPase and one exchanger (Sb07g028160 and Sb04g003135), respectively. In addition, 10 calcium ATPases and 8 calcium exchangers contain one or more seed-specific elements in their promoter region, while no seed-specific expression was detected in 5 ATPases (Sb08g001260, Sb01g038990, Sb06g027770, Sb09g001850 and Sb02g028935) and 3 exchangers (Sb06g031080, Sb03g008600 and Sb05g026100). Hormone-specific expression, *i.e.*, TGA-element, GARE, ERE, TGACG, ABRE, IIb and CGTCA-motif, has also been observed in all calcium transporters. Meristem-specific regulation, *i.e.*, presence of dOCT box, CCGTCC-box and CAT-box, were present only in some transporters while the presence of *cis*-regulatory elements like Skn-1 motif, GCN4 motif, O2-site and RY element confers endosperm-specific gene expression. Finally, the involvement of calcium transporter in regulation of abiotic stress has been identified according to the presence of different stress-responsive elements like HSE, MBS, although transporters Sb08g001260, Sb09g001850 and Sb06g031080 do not contain any stress-related elements.

Conclusion

In this study, a complete analysis of the calcium transporter gene family in rice and sorghum is presented, including multiple sequence alignment, domain analysis of Ca^{2+} transporters, phylogeny, chromosomal locations and analysis of their *cis*-regulatory elements. *In silico* analysis has revealed the existence of 28 and 31 members of transporter genes encoding channel, ATPase and exchanger in genomes of sorghum and rice, respectively. Multiple sequence alignment of these transporter proteins of sorghum and rice showed certain conservedness. Phylogenetic analysis further segregated these proteins into four clusters each representing calcium channel, IIA & IIB type calcium ATPases and calcium exchanger. Motif analysis revealed some conserved motifs between channel, ATPase and exchanger in rice and sorghum. The analysis of *cis*-regulatory elements for the predicted calcium transporters revealed that the major putative functions of these genes are associated with gene regulation of seed storage proteins, abiotic and biotic stress, photoperiod, growth hormone and meristem. Although the function of calcium transporters for calcium accumulation in seed has not been reported so far, the presence of seed-specific motif in upstream region

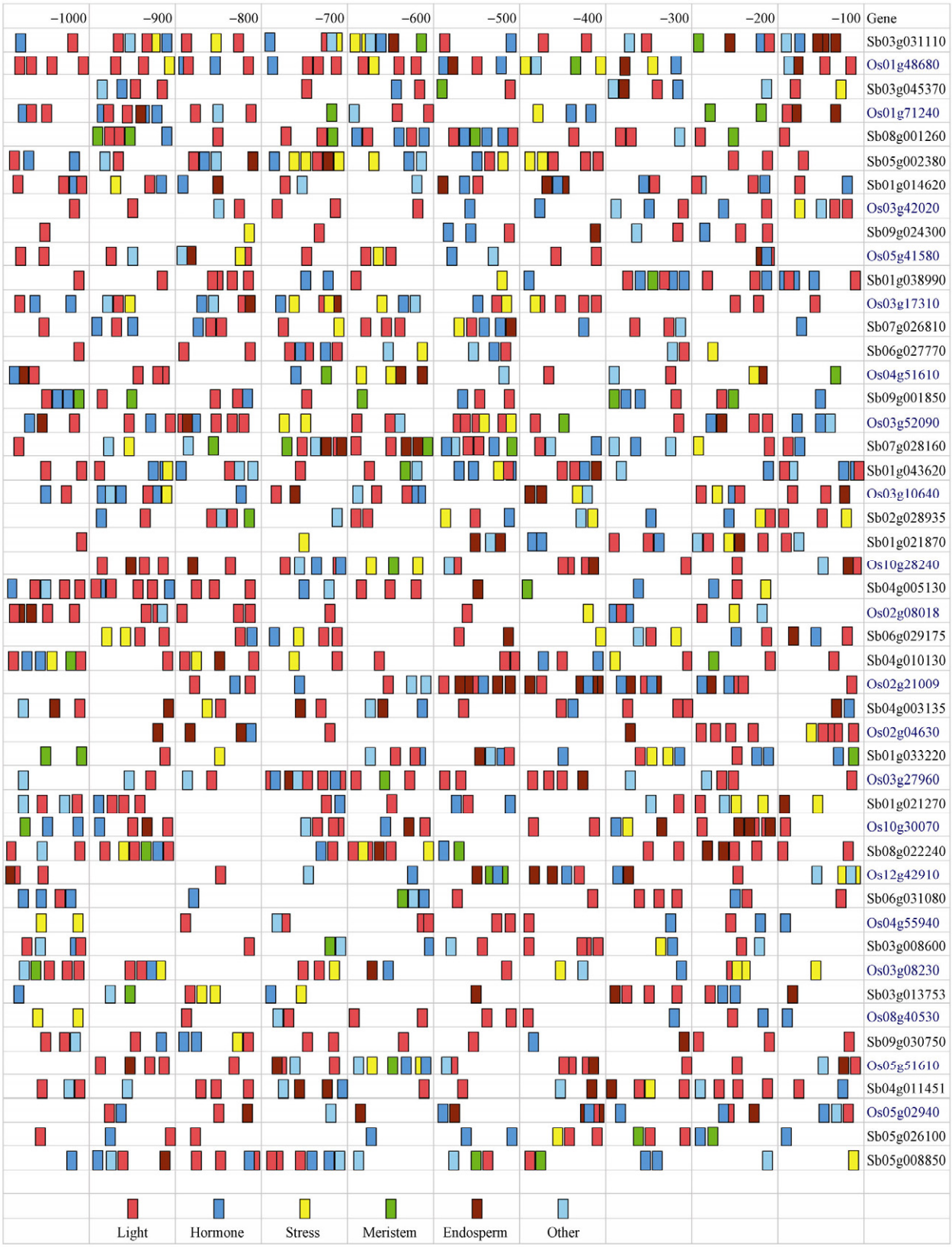

Figure 4 *Cis*-regulatory elements in the upstream region of calcium transporter genes from sorghum and rice. Different elements were indicated using different colors. Elements responsive to light, hormone and stress were indicated in red, blue and yellow. Elements for meristem and seed specific expression were indicated in green and brown while other elements were indicated in light blue.

suggests that these proteins might be involved in calcium accumulation.

Characterization of these proteins at physiological, molecular and structural level might shed some light on their functionality. The cause of gene movement and the erosion of gene colinearity between rice and sorghum chromosomes has been an unsolved task. However, advent of comparative genomics will provide insights for such changes in chromosomal organization. The identification of syntenic relationship between rice and sorghum in the present study showed that calcium transporters are not uniformly distributed but rather clustered on certain chromosomes. The present *in silico* analysis of the predicted transporter gene family has provided significant clues for exploring the expression and function of these calcium transporters under different environmental conditions. These studies would further help in understanding the molecular basis of many agriculturally important traits such as calcium accumulation in developing grains and their roles in development and defense against biotic and abiotic stresses. Furthermore, deciphering the diversity, organization and phylogeny of calcium transporter gene family in sorghum would also facilitate future annotation of transporter genes within genomes of other cereals.

Materials and Methods

Search of databases for the identification of calcium transporter family members

In order to perform comparative analysis, the sequences of Ca^{2+} transporters from rice (*Oryza sativa*) and sorghum (*Sorghum bicolor*) were downloaded from two different sources, TIGR (The Institute for Genomic Research; http://www.tigr.org/) pseudomolecule database (version 4) for rice and MIPS (Munich information center for protein sequence; http:// mips.helholtz-muenchen.de/plant/sorghum/) for sorghum, respectively. For the identification of Ca^{2+} transporters in rice and sorghum, the homology search of the Ca^{2+} transporter proteins was performed by BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) using blastp and tblastn algorithm. Detailed analysis of Ca^{2+} transporters in whole genomes of rice and sorghum, including their identification, classification, sequence analysis, domain analysis, chromosomal locations and phylogenetic relationships were carried out as described below.

Phylogenetic analysis

Amino acid sequences of all the identified Ca^{2+} transporters from sorghum and rice were aligned separately using ClustalW (*28*) and the phylogenetic tree was constructed using UPGMA method of MEGA version 4.0.02 (*29*). Each node was tested using the bootstrap approach by taking 100 replications to ascertain the reliability of nodes. The number indicated percentages against each node.

Chromosomal distribution of calcium transporters and synteny between rice and sorghum

The sorghum genome is approximately two-fold larger in size but has two fewer chromosomes relative to rice. The difference in chromosome number between sorghum and rice is due to two chromosome fusion events that occurred prior to the divergence of sorghum from Pennisetum (*30*). The chromosomal position of each calcium transporter gene in the rice and sorghum was determined by blastn search with NCBI genome (chromosome) database against genomic sequence of each chromosome for the rice (Accession No. NC008394-NC008405) and sorghum (Accession No. CM000760-CM000769) and then manually marked. The chromosomal distribution was visualized by NCBI Map Viewer (http://www.ncbi.nlm.nih.gov/mapview/).

Analysis of conserved motifs

To identify the conserved motifs within the protein sequences of calcium transporter family, the deduced protein sequences of all the calcium transporters in rice and sorghum were analyzed using online MEME (Multiple Expectation Maximization for Motif Elicitation) tool version 3.5.7 (*31*). For motif analysis, maximum number of motifs was set to 10 and optimal motif width was set as 20 to 50 amino acids while other factors were set as default.

Domain analysis of calcium transporters

The putative protein sequences of calcium transporters from rice and sorghum were subjected to protein functional analysis using Pfam version 23.0 (*32*) and SMART (Simple Modular Architecture Research Tool) version 5.1 (*33*).

Analysis of *cis*-regulatory elements

For promoter analysis, sequences 1,000 bp upstream of the initiation codon of the putative calcium transporter genes were retrieved and subjected to search using CARE program (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) of Plant CARE database to identify *cis*-regulatory elements.

Acknowledgements

This work was financially supported by Department of Biotechnology, Govt. of India as Programme Support for research and development in Agricultural Biotechnology at G.B. Pant University of Agriculture and Technology, Pantnagar (Grant No. BT/PR7849/AGR/02/2006). The facilities provided by DBT-funded SUB-DIC Bioinformatics Centre, Pantnagar for carrying out this *in silico* investigation is also acknowledged.

Authors' contributions

AK conceived the idea, designed the project and interpreted the results. AG performed most of the data analysis and drafted the manuscript. GT, DP and SG provided advice for upstream promoter analysis and revised the manuscript. All the authors read and approved the final manuscript.

Competing interests

The authors have declared that no competing interests exist.

References

1 Blaustein, M.P. and Lederer, W.J. 1999. Sodium/calcium exchange: its physiological implications. *Physiol. Rev.* 79: 763-854.
2 Carafoli, E. 2002. Calcium signaling: a tale for all seasons. *Proc. Natl. Acad. Sci. USA* 99: 1115-1122.
3 Smyth, J.T., *et al*. 2006. Emerging perspectives in store-operated Ca^{2+} entry: roles of Orai, Stim and TRP. *Biochim. Biophys. Acta* 1763: 1147-1160.
4 Berridge, M.J. 1997. The AM and FM of calcium signalling. *Nature* 386: 759-760.
5 Santella, L., *et al*. 2005. The cell cycle: a new entry in the field of Ca^{2+} signaling. *Cell. Mol. Life Sci.* 62: 2405-2413.
6 Lipskaia, L. and Lompre, A.M. 2004. Alteration in temporal kinetics of Ca^{2+} signaling and control of growth and proliferation. *Biol. Cell* 96: 55-68.
7 Strehler, E.E. and Zacharias, D.A. 2001. Role of alternative splicing in generating isoform diversity among plasma membrane calcium pumps. *Physiol. Rev.* 81: 21-50.
8 Monteith, G.R. and Roufogalis, B.D. 1995. The plasma membrane calcium pump—a physiological perspective on its regulation. *Cell Calcium* 18: 459-470.
9 Sze, H., *et al*. 2000. Diversity and regulation of plant Ca^{2+} pumps: insights from expression in yeast. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 51: 433-462.
10 Monteith, G.R., *et al*. 2007. Calcium and cancer: targeting Ca^{2+} transport. *Nature Rev. Cancer* 7: 519-530.
11 Nagata, T., *et al*. 2004. Comparative analysis of plant and animal calcium signal transduction element using plant full-length cDNA data. *Mol. Biol. Evol.* 21: 1855-1870.
12 Palmgren, M.G. and Harper, J.F. 1999. Pumping with plant P-type ATPases. *J. Exp. Bot*. 50: 883-893.
13 Axelsen, K.B. and Palmgren, M.G. 2001. Inventory of the superfamily of P-type ion pumps in Arabidopsis. *Plant Physiol.* 126: 696-706.
14 Tuteja, N. and Mahajan, S. 2007. Calcium signaling network in plants: an overview. *Plant Signal Behav.* 2: 79-85.
15 White, P.J. 2000. Calcium channels in higher plants. *Biochim. Biophys. Acta* 1465: 171-189.
16 White, P.J., *et al*. 2002. Genes for calcium-permeable channels in the plasma membrane of plant root cells. *Biochim. Biophys. Acta* 1564: 299-309.
17 Schantz, H.L. 1954. The place of grasslands in the earth's cover of vegetation. *Ecology* 35: 143-145.
18 Evans, L.T. 1998. *Feeding the Ten Billion: Plants and Population Growth.* Cambridge University Press, Cambridge, UK.
19 Paterson, A.H., *et al*. 2009. The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457: 551-556.
20 Bowers, J.E. 2003. A high-density genetic recombination map of sequence-tagged sites for sorghum, as a framework for comparative structural and evolutionary genomics of tropical grains and grasses. *Genetics* 165: 367-386.
21 Altschul, S.F., *et al.* 1990. Basic local alignment search

tool. *J. Mol. Biol.* 215: 403-410.
22 Altschul, S.F., *et al.* 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.
23 Peng, Y., *et al.* 1999. Comparative genome mapping of *Sorghum bicolor* (L.) Moench using an RFLP map constructed in a population of recombinant inbred lines. *Plant Breed*. 118: 225-235.
24 Klein, P.E., *et al.* 2000. A high-throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. *Genome Res*. 10: 789-807.
25 Paterson, A.H., *et al*. 2004. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl. Acad. Sci. USA* 101: 9903-9908.
26 Devos, K.M. 2005. Updating the "Crop circle." *Curr. Opin. Plant Biol.* 8: 155-162.
27 Lescot, M., *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.
28 Thompson, J.D., *et al*. 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.
29 Tamura, K., *et al*. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596-1599.
30 Kellogg, E.A. 2001. Evolutionary history of the grasses. *Plant Physiol*. 125: 1198-1205.
31 Bailey, T.L., *et al*. 2006. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res*. 34: W369-373.
32 Finn, R.D., *et al*. 2006. Pfam: clans, web tools and services. *Nucleic Acids Res.* 34: 247-251.
33 Schultz, J., *et al*. 1998. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl. Acad. Sci. USA* 95: 5857-5864.

Supplementary Material

Figure S1

DOI: 10.1016/S1672-0229(11)60017-X