HORIZON
e-Publishing Group
HePG

Check for updates   OPEN ACCESS

RESEARCH ARTICLE

Genome-wide identification, characterization and expression analysis of the expansin gene family under drought stress in tea (*Camellia sinensis* L.)

Kuntala Sarma Bordoloi, Pallabika Dihingia, Debasish B Krishnatreya & Niraj Agarwala*

Department of Botany, Gauhati University, Gopinath Bordoloi Nagar, Jalukbari, Guwahati 781 014, India
*Email: niraj_botany@gauhati.ac.in

ABSTRACT

During several developmental processes, expansins contribute to cell enlargement by promoting cell wall loosening. To explore the biological roles of expansins during drought stress response and to characterize different expansins in tea, we performed a detailed analysis of the expansin gene family covering phylogeny, gene structure, profiling of gene expression and co-expression network analysis. We identified a total of 40 expansin genes in the tea genome belonging to 3 subfamilies, out of which 29 tea expansins belong to EXPA, 9 to EXLA and 2 to EXPB subfamilies. A minimum of 3 and a maximum of 13 exons are present in the gene structure of expansins. Presence of drought stress responsive *cis-acting* elements in the upstream of promoter regions of 40% of the identified expansins shows that the putative expansins may have been involved in tea plant's response to drought stress. At least 15 out of the 40 expansin genes are found to be differentially expressed in response to drought in each of the drought stress related public datasets analysed *in-silico*. TEA022767 belonging to EXPA subfamily is seen to be upregulated during drought stress, as revealed from the analysis of all three publicly available bio-projects. Co-expression network analysis shows that TEA022767 and TEA032954 form a connecting link between two expression correlation groups that further signifies their role in drought stress response in tea. This study helps to interpret and to understand the biological roles of diverse expansin genes in tea plants under drought stress conditions.

Introduction

Expansins are cell wall-loosening proteins lacking hydrolytic activity. These proteins regulate and uphold the shape by loosening/disrupting H-bonds of cellulose microfibrils in the extracellular matrix in a non-enzymatic pH dependent manner. This leads to remodelling and loosening of cell wall and cell extension (1, 2). Expansins play diverse roles in various biological processes associated with modification of cell-wall, root hair growth and root development (3, 4), mycorrhizal interaction and root nodule formation (5, 6), abscission (7), leaf initiation, growth and development (8, 9), flower development and abscission (10), stem and pollen tube elongation (11, 12), fiber development (13), fruit maturing, development and dehiscence (14-16), fruit inflexibility and softening (17), seed yield and

germination (18) and several biotic and abiotic stress responses. The first expansins were found in cucumber (*Cucumis sativus* L.) (19). Since then, plant expansins have been studied in hundreds of different plants. A large superfamily of expansin in plants encompassing four subfamilies has been discovered by genome-wide investigation (20). They are defined as expansin A (EXPA), expansin B (EXPB), expansin-like A (EXLA) and expansin-like B (EXLB). The two major subfamilies EXPA and EXPB have been experimentally proved to exhibit cell-wall loosening properties. Very less is known about the exact functions of EXLA and EXLB subfamilies (20, 21). Expansins in plants are typically 250–275 amino acids long and contain two domains (1 and 2) followed by a signal peptide of about 20–30 amino acid residues (22). Double psi beta-barrel or DPBB is the domain homologous to glycoside hydrolase

family 45 (GH45). It comprises a conserved HFD motif and several conserved cysteine residues. The 2nd domain is distantly related to G2As or grass group-2 pollen allergens (22, 23).

A diverse set of functions are dependent on plant cell wall. It acts as the first line of defense during pathogen attack and provides a structural framework to maintain plant growth and development. The cell wall must retain the capacity to respond to both biotic and abiotic stimuli. Genes encoding enzymes that produce the plant cell wall show distinct expression patterns when subjected to various stresses. This signifies that they may be responsible for facilitating stress tolerance through modifications in cell wall composition (22).

Tea (*Camellia sinensis* L.) is an important evergreen woody plant belonging to the family Theaceae. Tea made from the leaves of the plant is one of the popular non-alcoholic beverages of the world with its important economic, health and cultural values. It comprises of several important secondary metabolites such as caffeine, polyphenols, vitamins and minerals beneficial to human health. Tea plant gets affected by several abiotic stresses like temperature (both low and high) and light (low light and UV), that affect the crop productivity.

Drought is a major limiting factor that significantly influences crop growth and yield. A 14-33% loss of tea productivity and a 6-19% plant mortality as a result of drought stress was reported in earlier study (24). A positive correlation exists between expression of expansin genes and drought tolerance in crops. A recent study showed that overexpression of TaEXPA2 in wheat helped the plants to cope up with drought stress by induced lateral root formation as well as by increased antioxidant activity (25). In another study, transgenic tobacco overexpressing TaEXPA2 improved drought tolerance in the plants (18). In a similar way, ectopical expression of RhEXPA4 in transgenic *Arabidopsis* increased survival rate of the plants under drought stress (26). Although expansins have been characterized in many plant species, such as *Arabidopsis* (27), rice (27), maize (28), wheat (29), soybean (30), tobacco (31), tomato (32), apple (33) and grape (34), the features and functions of expansin gene family in tea are unexplored. Considering the importance of expansins in mediating drought tolerance in crops, a genome wide survey of tea expansins was conducted and expression profiling of expansin genes during drought stress in tea using publicly available data was carried out. Our findings will help in elucidating important information about the role of expansins in drought tolerance in tea plants.

Materials and Methods

Genome-wide identification, classification and characterization of expansin genes in tea

In order to construct a local protein database, the peptide sequences of tea were downloaded from the publicly available database TPIA (Tea Plant Information Archive). Expansin protein sequences of *Arabidopsis thaliana* (L.) Heynh. and *Oryza sativa* L. were downloaded from their respective databases viz. TAIR and Oryzabase. We used the expansin protein sequences of both as query sequences to perform BLASTX with an e-value cutoff of $1e^{-10}$ against the locally constructed tea protein database. The Hidden Markov Model (HMM) profiles of DPBB domain (PF03330) and pollen allergen domain (PF01357) from the Pfam database (http://pfam.xfam.org/) was also used to identify the expansin genes in tea genome through the HMMER 3.0 program. All the putative tea EXP protein sequences were further confirmed for conserved domains in the NCBI Conserved Domain Database (CDD) (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). To categorise the putative expansins into the four subfamilies, the annotated *Arabidopsis* expansin proteins were used to construct a phylogenetic tree with tea expansins. We used the Protparam tool (http://web.expasy.org/protparam/) to analyse parameters like molecular weight, number of amino acids, isoelectric points (pI) etc. (35). FUEL-mLoc (http://bioinfo.eie.polyu.edu.hk/FUEL-mLoc/) tool was used to determine the sub-cellular locations of the predicted expansins (36) keeping all the parameters to default. We used the online tool SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP-4.1/) for signal peptide prediction (37) and the C, S and Y scores were used to predict the presence or absence of signal peptides.

Gene structure and motif analysis

The genomic and coding sequences of expansins were downloaded from the TPIA database. The online tool, GSDS (Gene Structure Display Server) (http://gsds.cbi.pku.edu.cn) (38) was used to predict the exon-intron structures of the predicted expansins by setting the parameters to default. Setting the maximum motif width to 300 and motif number to 10, the online tool MEME (http:/meme.nbcr.net/ meme/intro.html) (39) was used to find out the conserved motifs in the expansin protein sequences.

Analysis of cis-acting elements in tea expansin gene promoter region

The promoter regions (2000 bp sequence upstream of the genomic DNA sequence) of the tea expansin genes were extracted from the TPIA database (http://tpia.teaplant.org/), and they were analysed using PlantCARE Database (http://bioinformatics.psb. ugent.be/webtools/plantcare/html/) to predict the *cis-*elements (40).

Sequence alignment and phylogenetic tree construction

Twenty one *Arabidopsis* expansin protein sequences were downloaded from the TAIR (The Arabidopsis Information Resource) database (https://www. arabidopsis.org/). A phylogenetic tree of tea expansins with the *Arabidopsis* expansin proteins was constructed using the MEGA-X software (41) through the Maximum Likelihood method with 1000 bootstrap replicates setting the other parameters to default.

Estimation of synonymous and non-synonymous substitution rates and calculation of Ka/Ks ratios of tea expansins with their orthologs in Arabidopsis

Extraction of the orthologs of tea expansins in the *Arabidopsis* genome from the Phytozome database (42) was followed by calculation of Ka and Ks values for each pair of orthologous genes through the PAL2NAL tool (http://www.bork.embl.de/pal2nal/) (43). By using the Ka/Ks ratios, selection pressure on the orthologous gene pairs was analyzed. Using the formula $Ks/(2 \times 6.5 \times 10^{-9}) \times 10^6$ (44), we calculated the divergence time of each expansin gene.

Gene expression profiling of the predicted expansin genes and construction of co-expression network

Three public RNA-seq datasets with NCBI IDs PRJNA596070 (drought with FA treatment), PRJNA539837 (drought stress) and PRJNA297732 (drought stress) were downloaded from Sequence Read Archive database of NCBI (http://www.ncbi.nlm.nih.gov/sra) and converted to fastq format by SRA toolkit (https://www.ncbi. nlm.nih.gov/books/NBK158900/). FastQC program (https://www.bioinformatics.babraham.ac.uk/project s/fastqc/) was used for quality check of the raw sample data. We performed the differential expression analysis using HISAT2, StringTie (45) and DESeq tools (46) pipeline. The online tool ClustVis was used to generate the expression heat maps (https://biit.cs.ut.ee/clustvis/) (47). For building the gene co-expression network, we used the FPKM data of the genes in all samples and constructed the network with the help of Cytoscape 3.8 version (48).

Results

Genome-wide identification and characterization of expansin genes in tea

In order to identify expansin genes in the tea genome, a combination of BLAST searches using the corresponding sequences from *Arabidopsis thaliana* and *Oryza sativa* as queries and the HMMER 3.0 program was used to predict expansins from tea genome. A total of 40 protein sequences were identified as putative expansin genes in *C. sinensis* after scanning for the conserved domains. All the 40 EXP genes of tea have both the conserved domains DPBB and Pollen allergen domains. There are 29 tea EXPs under the EXPA subfamily (Supplementary Fig. 1), 9 under EXLA and 2 under EXPB subfamilies. None of the tea EXPs belonged to the EXLB subfamily. The identified expansin genes encoded proteins ranged from 195 (TEA008282) to 800 (TEA005109) amino acids (AA) in length and the predicted molecular weights (MWs) ranged from 20824.82 to 89950.99 Da. Isoelectric point (pI) values of most of the members were above 7.0 while ten of them have pI values below 7.0 (Table 1).

The exon-intron distribution of the expansin genes in tea was analysed by the online GSDS (Gene Structure Display Server) tool (Fig. 1). The lengths of boxes and lines are scaled based on gene length. It was found that all the members had at least 3 exons. The number of exons ranged from 3 to 13 in the 40 expansin genes in tea.

To estimate the divergence among the expansin genes, we could predict different conserved motifs and their distribution pattern among the genes. Among these, motif 5 is the longest of 97 residues and motifs 6 and 7 are shortest of 15 residues each. Motif 7 is present in all the 40 genes except three *i.e.* TEA020839, TEA020841 and TEA032954. The longest motif i.e. motif 5 is present only in four genes viz. TEA033642, TEA004336, TEA022147 and TEA010004 (Fig. 2). To investigate the sub-cellular location of the putative expansins, we used FUEL-mLoc Server. The analysis suggested that among the 40 EXP genes, 32 were secreted to cell wall while 5 were secreted extra-cellularly, 2 to cell membrane and 1 to chloroplast (Table 1). Also, the presence and absence of signal peptide was noted. The signal peptide prediction showed that not all but half of the expansins contain signal peptides in their sequence (as shown in Table 1).

Phylogenetic analysis of the predicted expansins

The deduced protein sequences were aligned using CLUSTALW and a phylogenetic tree was constructed by the MEGA-X software using the maximum likelihood method with 1000 bootstrap replicates (Fig. 3). It produced a rooted tree that does not assume a constant rate of evolution but offers superior accuracy. In the Fig. 3, length of the branches represents genetic distance. The three classes of tea expansins (EXPA, EXPB, EXLA) formed three separate groups that have made them easily distinguishable among the classes. Two EXPB expansins TEA008282 and TEA002540 formed the smallest clade, nine EXLA expansins formed a larger clade and rest 29 EXPA expansins formed the largest group together (Fig. 3).

Calculation of divergence time and ratio of non-synonymous to synonymous substitution rates of tea expansins with their orthologs in Arabidopsis

The ka/ks ratio is calculated as the ratio of the number of non-synonymous substitutions per non-synonymous site (ka) to the number of synonymous substitutions per synonymous site (ks). In our study, ka/ks ratio of all the tea expansin genes were found to be less than unity i.e., <1 which shows that natural selection is suppressing protein changes. In other way, we can say that all the 40 expansin gene pairs are under purifying selection. The average divergence time of the expansin genes in tea was found to be 0.7×10^3 million years ago (Supplementary Table 2).

Analysis of cis-acting elements in upstream regions of tea expansin genes

The *cis-acting* elements of the promoter region among the 40 expansin members were analysed and the results are tabulated (Supplementary Table 1). The results revealed that the *cis-acting* elements present in the promoter region of the expansin genes can be divided into 26 categories. Among the 26 categories, light responsive *cis-acting* elements were present in all the genes. Among all the categories,

Table 1. Physicochemical characterization of expansins in tea

Tea Ids	Class	Genome location	Strand	Signal Peptide	No. of introns	No. of amino acids	Mol. Wt.	pI value	Sub-cellular localization
TEA020839	EXPA	Scaffold1997:284458-308191	negative	No	5	321	35057.5	6.39	Cell-wall
TEA020841	EXPA	Scaffold1997:354255-359961	negative	1-26	4	358	38752.51	6.47	Cell-wall
TEA003333	EXPA	Scaffold871:109810-119399	positive	No	5	402	44418.69	9.61	Cell-wall
TEA016947	EXPA	Scaffold1302:716830-721896	negative	No	4	311	33052.23	8.75	Cell-wall
TEA008211	EXPA	Scaffold712:1645440-1649200	negative	1-24	2	251	26673.93	6.91	Cell-wall
TEA002962	EXPA	Scaffold912:236913-239839	positive	No	2	247	26964.38	9.77	Cell-wall
TEA010311	EXPA	Scaffold2408:1084516-1087638	positive	1-31	2	268	29523.17	9.08	Cell-wall
TEA001505	EXPA	Scaffold558:4125342-4126434	positive	1-27	2	254	27515.97	8.42	Cell-wall
TEA002242	EXPA	Scaffold1983:724808-727178	positive	1-17	2	240	26071.49	9.13	Cell-wall
TEA026647	EXPA	Scaffold216:348230-351793	negative	No	2	256	27663.39	8.56	Cell-wall
TEA013500	EXPA	Scaffold980:117008-119742	positive	No	2	279	30007.45	7.01	Cell-wall
TEA016522	EXPA	Scaffold1761:2553417-2556221	positive	1-26	2	254	27152.28	6.78	Cell-wall
TEA027164	EXPA	Scaffold3500:3224145-3226210	negative	1-26	2	253	27052.11	6.01	Cell-wall
TEA015776	EXPA	Scaffold3826:900539-909810	positive	No	11	564	62198.2	7.02	Cell-wall
TEA013504	EXPA	Scaffold980:123067-124868	positive	1-25	2	252	27101.22	7.5	Cell-wall
TEA022767	EXPA	Scaffold9600:98191-100202	positive	1-27	2	254	27194.45	8.12	Cell-wall
TEA031014	EXPA	Scaffold605:1478140-1486823	positive	No	2	215	23223.76	7.56	Cell-wall
TEA007366	EXPA	Scaffold1345:557454-559926	positive	No	3	259	27896.1	8.31	Cell-wall
TEA004163	EXPA	Scaffold1522:831087-835799	positive	1-23	2	260	27836.63	9.26	Cell-wall
TEA013846	EXPA	Scaffold2753:1598056-1601496	positive	1-21	2	258	27828.65	9.39	Cell-wall
TEA008349	EXPA	Scaffold142:368598-379730	negative	1-14	9	785	88849.91	6.96	Cell-wall
TEA009532	EXPA	Scaffold258:214264-217597	negative	1-23	2	260	28214.14	9.3	Cell-wall
TEA024881	EXPA	Scaffold160:865524-871878	positive	No	3	333	36302.48	8.75	Cell-wall
TEA015188	EXPA	Scaffold1167:1176466-1179493	negative	1-29	2	286	32117.81	9.44	Cell-wall
TEA000787	EXPA	Scaffold3620:559700-590230	negative	No	7	472	53691.71	8.89	Cell-wall
TEA022584	EXPA	Scaffold502:559143-585716	positive	No	10	493	55171.06	7.56	Cell-wall
TEA026533	EXPA	Scaffold5206:335908-341652	positive	1-25	4	415	45128.59	7.91	Cell-wall
TEA008282	EXPB	Scaffold2348:1486653-1490744	positive	No	2	195	20824.82	6.8	Cell-wall
TEA005109	EXPA	Scaffold1857:3534737-3551250	negative	1-25	9	800	89950.99	9.2	Cell-wall
TEA014210	EXLA	Scaffold2985:491364-497140	positive	1-24	4	251	27461.45	4.81	Extra-cellular
TEA002540	EXPB	Scaffold393:1371897-1384376	negative	1-27	7	529	57503.27	8.48	Cell-wall
TEA021837	EXLA	Scaffold1713:1646430-1649208	positive	No	4	328	35948.26	8.78	Cell-membrane
TEA020837	EXPA	Scaffold1997:272173-284059	negative	No	2	231	25343.09	8.68	Cell-wall
TEA012391	EXLA	Scaffold3880:131477-133774	positive	1-24	4	254	27542.69	4.52	Extra-cellular
TEA033642	EXLA	Scaffold973:627240-628912	negative	No	3	227	25102.61	8.02	Extra-cellular
TEA004336	EXLA	Scaffold10551:46387-56896	negative	No	12	601	65901.01	9.06	Chloroplast
TEA022147	EXLA	Scaffold1552:1473827-1475891	negative	1-24	3	250	27545.21	7.42	Extra-cellular
TEA032954	EXLA	Scaffold2938:990566-992945	negative	1-18	4	259	28074.14	8.43	Cell-membrane
TEA010004	EXLA	Scaffold21350:198655-200274	positive	1-24	3	250	27440	5.2	Extra-cellular
TEA000961	EXLA	Scaffold78:1180204-1187100	negative	No	3	311	35734.05	9.01	Cell-wall

anaerobic induction and abscisic acid response are some of the most abundantly occurring functions. Some other important functions include drought inducibility, low temperature response, gibberellin-responsive, zein metabolism regulation, auxin response. Presence of drought responsive *cis-acting* elements in 40% of the genes is in line with the hypothesis that expansins might have roles in drought stress response in tea (33).

Differential in silico expression of expansin genes in response to drought and co-expression network analysis

Heat maps constructed to analyse expression profiles of tea expansin genes in response to drought stress, showed differential expression patterns of the genes. Fig. 4a shows that the expression level of genes TEA020841, TEA033642, TEA010004, TEA022147, TEA021837, TEA012391, TEA031014, TEA032954, TEA000961, TEA007366, TEA008211, TEA009532 and TEA022767 were highly increased in response to severe drought. However, with fulvic acid (FA)

treatment, the expression levels of these transcripts got substantially reduced. This may suggest that these genes have definite roles in severe drought response. Gene TEA022584 showed a decrease in its expression level in drought treatments coupled with FA treatment, with relatively high expression in the mild and severe drought treatments without FA treatment. Expression of transcripts TEA008349 and TEA015188 were low in mild drought, mild drought with FA and severe drought treatments, but their expression levels got elevated during the severe drought with FA treatment. Downregulation in expression of TEA003333 and TEA015776 genes were noted in severe drought as compared to mild drought. Genes TEA008282, TEA013500, TEA000787, TEA024811 and TEA013504 were highly expressed during mild drought treatment as compared to the other treatment groups, which might implicate their probable role in mild drought response rather than severe drought. TEA020841, TEA021837, TEA012839, TEA024881 genes showed a similar kind of
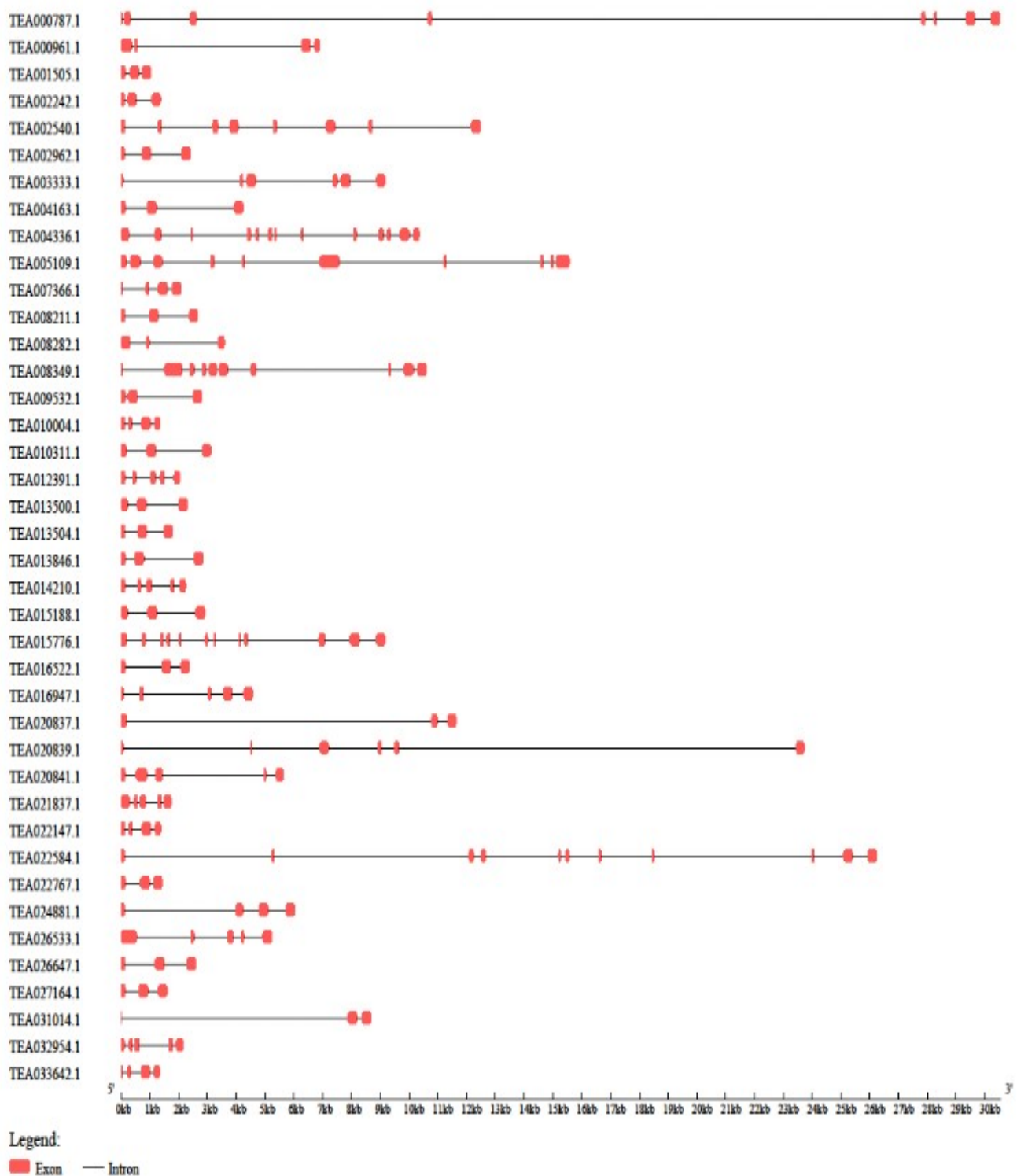
Fig. 1. Exon-intron structures of the tea expansin genes. Boxes: exons; lines: introns. The lengths of boxes and lines are scaled based on gene length.

expression in both mild and severe drought with FA treatment.

In a drought stress and re-watering experiment (Fig. 4b), the expression levels of genes TEA022584, TEA033642, TEA016947, TEA026533, TEA005109, TEA021837, TEA008282, TEA015776, TEA004336, TEA013504, TEA020837, TEA008349, TEA022147, TEA022767, TEA014210 and TEA032954 were noticeably higher in response to drought than the control and re-watering treatment conditions. Drought and re-watering treatments considerably

reduced the gene expression levels of TEA008211, TEA000787 and TEA015188 in the treated samples than the untreated control. The expression of genes TEA009532, TEA002962, TEA002242 and TEA003333 got highly downregulated during the drought stress condition, while re-watering treatment restored their expression similar to untreated control. Re-watering conditions elevated the expression levels of TEA013500, TEA031014, TEA007366, TEA013846, TEA027164, TEA016522, TEA000961, TEA010004 than the control and drought treated plants
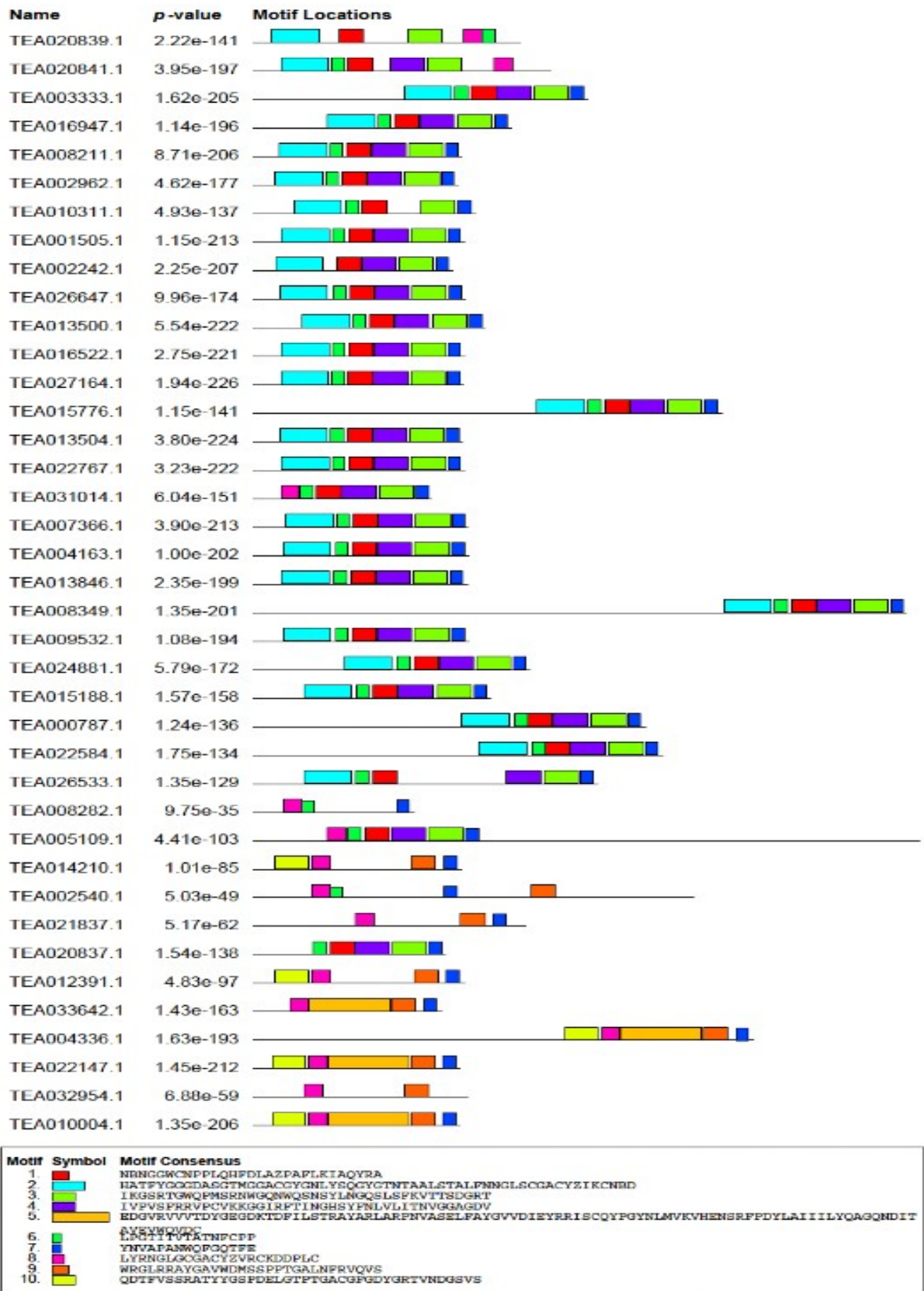
Fig. 2. MEME motif search results. Conserved motifs are indicated in coloured boxes.
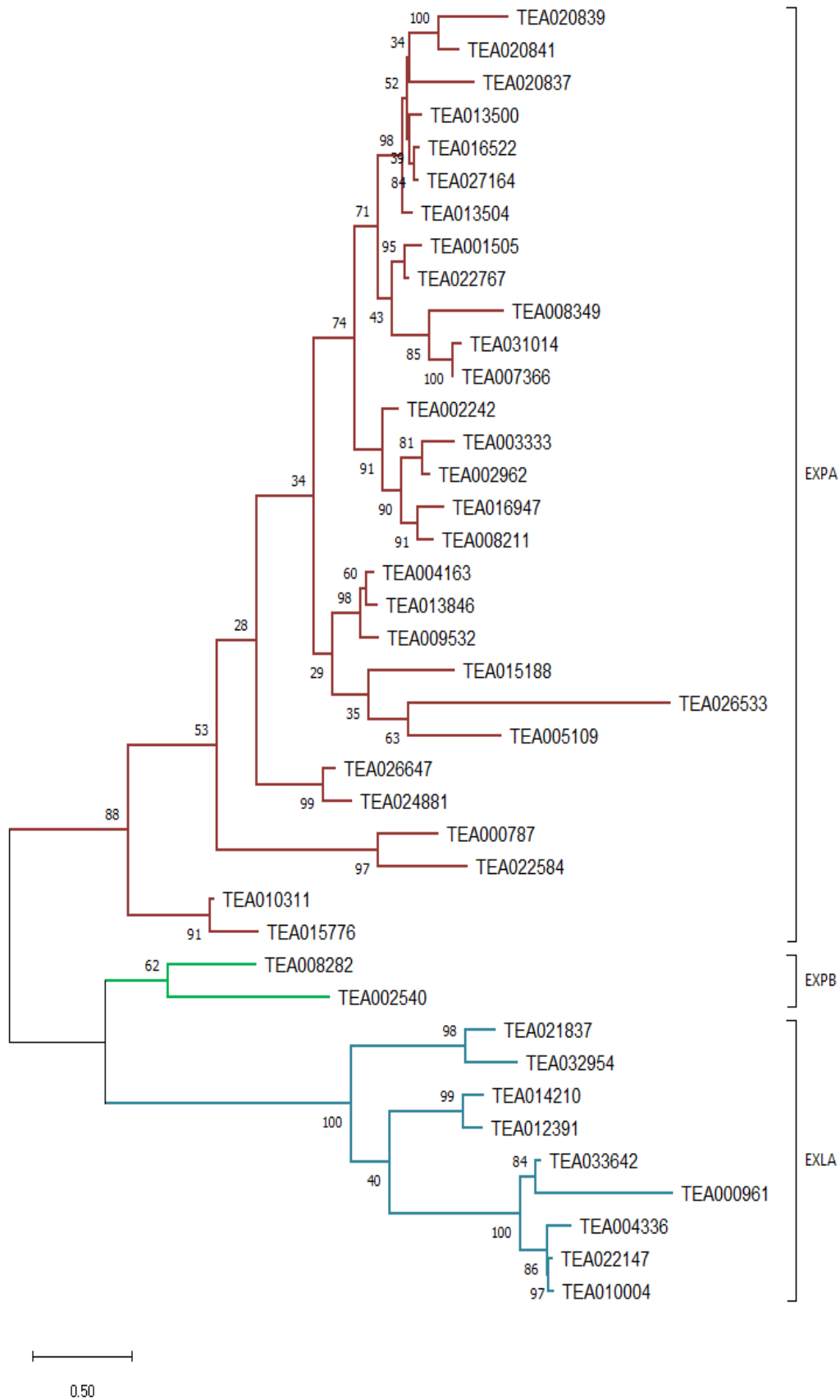
Fig. 3. Phylogenetic analysis of the expansin gene family in tea. The phylogenetic tree was constructed by MEGA-X software using the maximum likelihood method with 1000 bootstrap replicates.
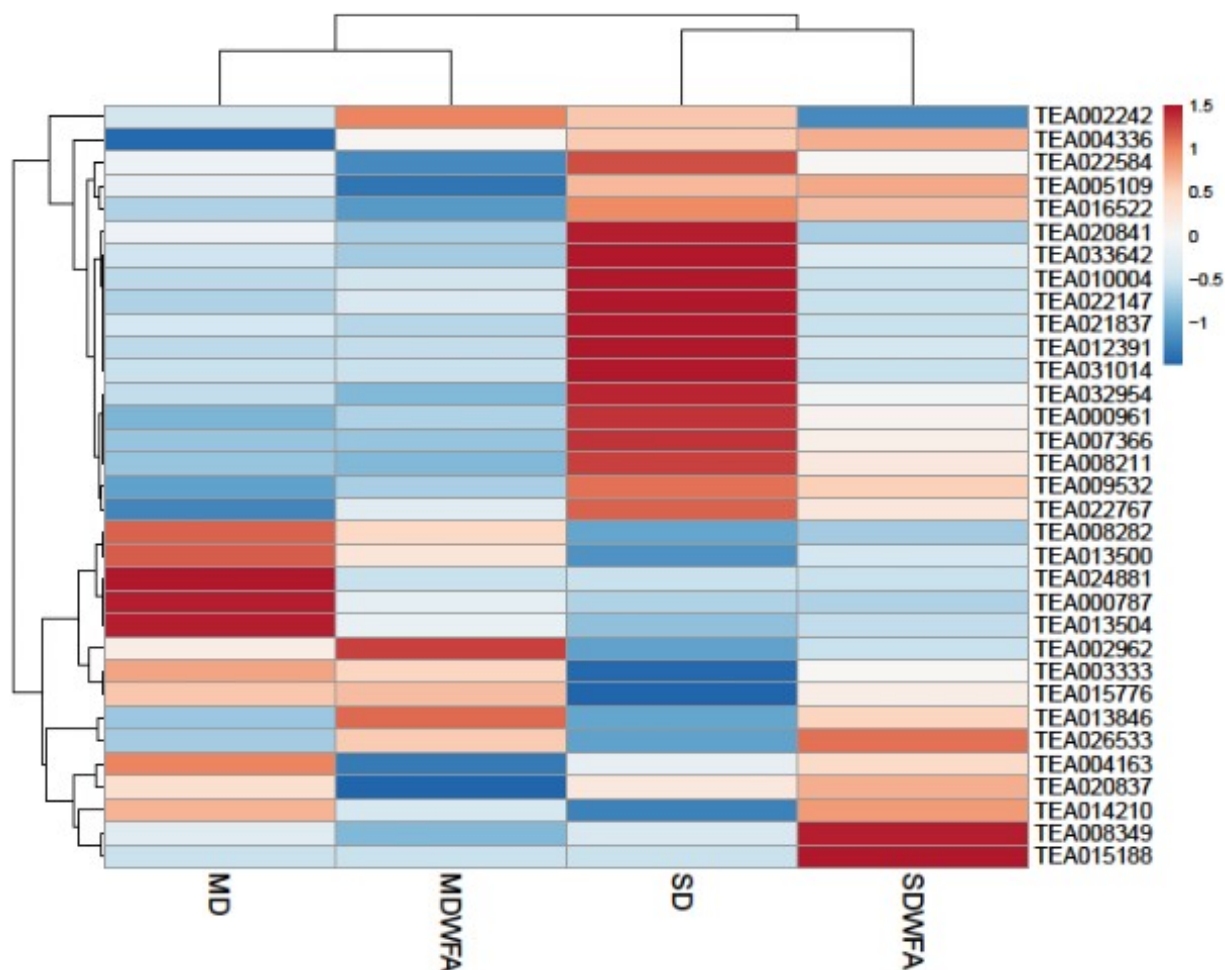
Fig. 4a. Expression of expansin genes in mild drought (MD), severe drought (SD); mild drought with FA (MDWFA) and severe drought with FA (SDWFA). The legend represents the logarithmic normalized FPKM. Dark red and dark blue boxes indicate highest and lowest expression levels.
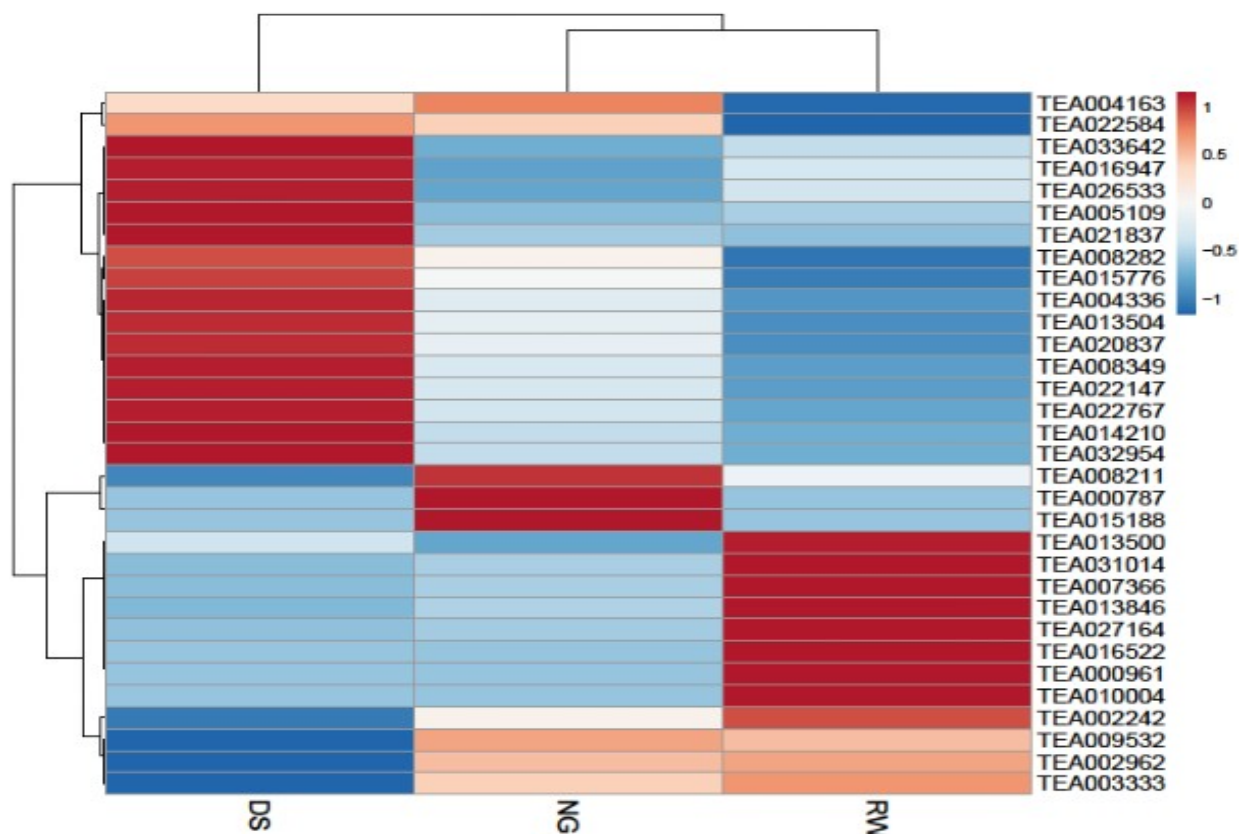

Fig. 4b. Expression of expansin genes in drought stress (DS), no stress (NG) and re-watering conditions (RW). The legend represents the logarithmic normalized FPKM. Dark red and dark blue boxes indicate highest and lowest expression levels.
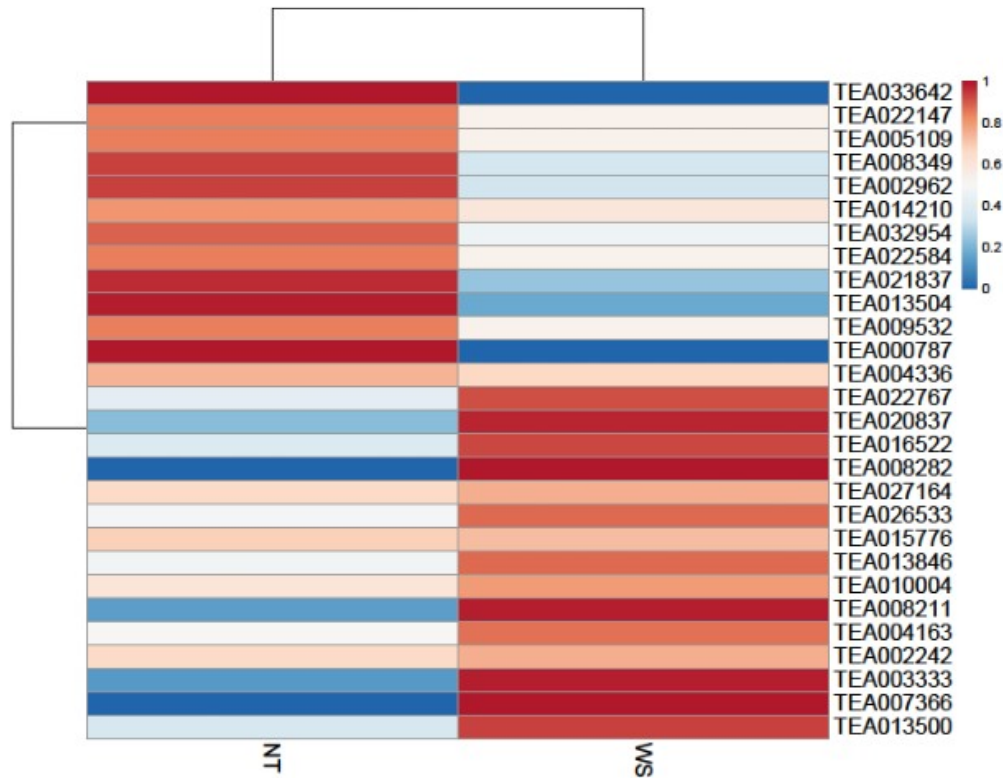
Fig. 4c. Expression of expansin genes in water stress (WS) and no treatment (NT) conditions. The legend represents the logarithmic normalized FPKM. Dark red and dark blue boxes indicate highest and lowest expression levels.

Another drought stress experiment (Fig. 4c) shows that the genes TEA022767, TEA020837, TEA016522, TEA008282 TEA008211, TEA003333, TEA007366 and TEA013500 exhibited high expression levels during water stress condition as compared to the control. The expression of transcripts TEA033642, TEA021837, TEA013504, TEA032954, TEA008349, TEA002962 and TEA000787 was strongly downregulated in water-stressed plants than the control plants.

Co-expression network analysis showed the clustering of genes in three distinct groups. TEA024881 interacts with TEA003333 and TEA000787 and forms a separate group with no edge connection with the other two groups (Fig. 5).

Discussion

Several studies help us to understand the involvement of expansins in acclimatizing the plants to drought stress by modifying cell expansion (49, 50). Transgenic plants overexpressing expansin genes showed drought tolerant phenotypes and relatively lower water potential than the wild type plants (26). An increase in reactive oxygen species (ROS) concentration in cells of plants is a common phenomenon in plants experiencing abiotic stress like drought. Accumulation of ROS scavenging enzymes like superoxide dismutase (SOD), peroxidase (POD), catalase (CAT) etc. has been reported in plants overexpressing expansin genes (18, 25). The flowchart (Fig. 6) demonstrates the possible ways as to how expansins mitigate drought stress in plants. From the studies conducted on overexpression of expansin genes in plants, it is evident that expansins play considerable role in developing drought tolerance in plants.

Just like any other plant, tea is also subjected to several abiotic and biotic stresses. So, tea plant uses different means to adapt to such stresses. Despite expansin playing an impending role in enhancing plant performance under unfavourable conditions, molecular and functional characterization of the expansin genes remains insufficient or unexplored in many species. To date, the expansin gene family has been investigated and characterized in plant species such as *Arabidopsis thaliana* L., *Glycine max* (L.) Merr., *Nicotiana tabacum* L., *Cucumis sativus* L., *Saccharum officinarum* L., *Triticum aestivum* L., *Solanum lycopersicum* L., *Ziziphus jujuba* Mill., *Vitis vinifera* L., *Oryza sativa* L., *Zea mays* L., *Citrullus lanatus* (Thunb.) Matsum. & Nakai etc. (28-34, 51-58). However, no systematic and thorough study of the expansin gene family was carried out earlier in tea. In this study, identification of expansins, its superfamily organization, characterization and expression analysis in response to drought stress was investigated. Out of the three bioprojects taken, research works from two bioprojects related to drought stress in tea have already been published (59, 60). Using the transcriptomic data from all three bioprojects, we have explored the biological role of expansin genes during drought stress in tea.

Our study identified 40 expansin genes in tea genome containing two conserved domains double psi-beta barrel (DPBB) and pollen allergen domains, characteristic of other expansins identified to date. Half of the numbers of the putative expansin genes comprise the signal peptide. In addition, as observed by FUEL-mLoc analysis, the majority of the tea expansins were secreted to the cell wall. Analysis of gene structure showed that 45% of the tea expansins had no introns in their gene structure, whereas all the other tea expansins had at least one intron.
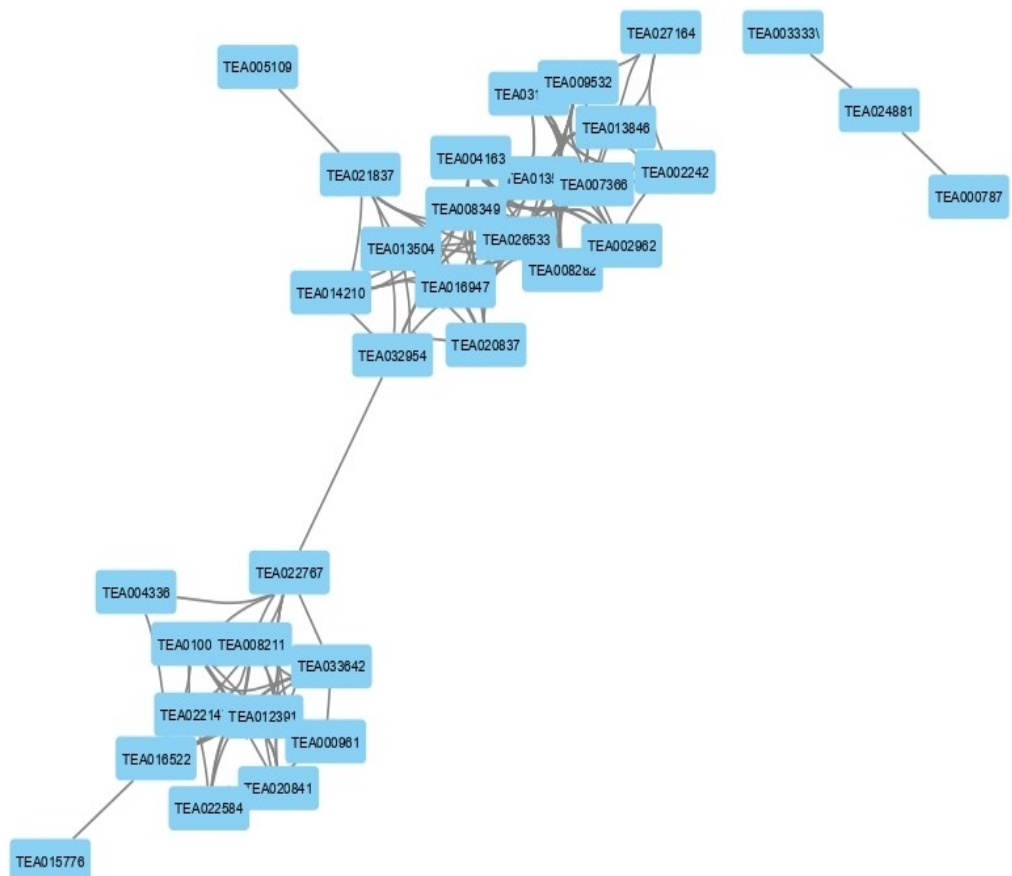
Fig. 5. Co-expression network of expansin genes with themselves involved drought stress responses using absolute PCC ≤ 0.9. Nodes represent genes, edges represent correlations. Straight edges represent positive correlation while edge bends represent negative correlation.
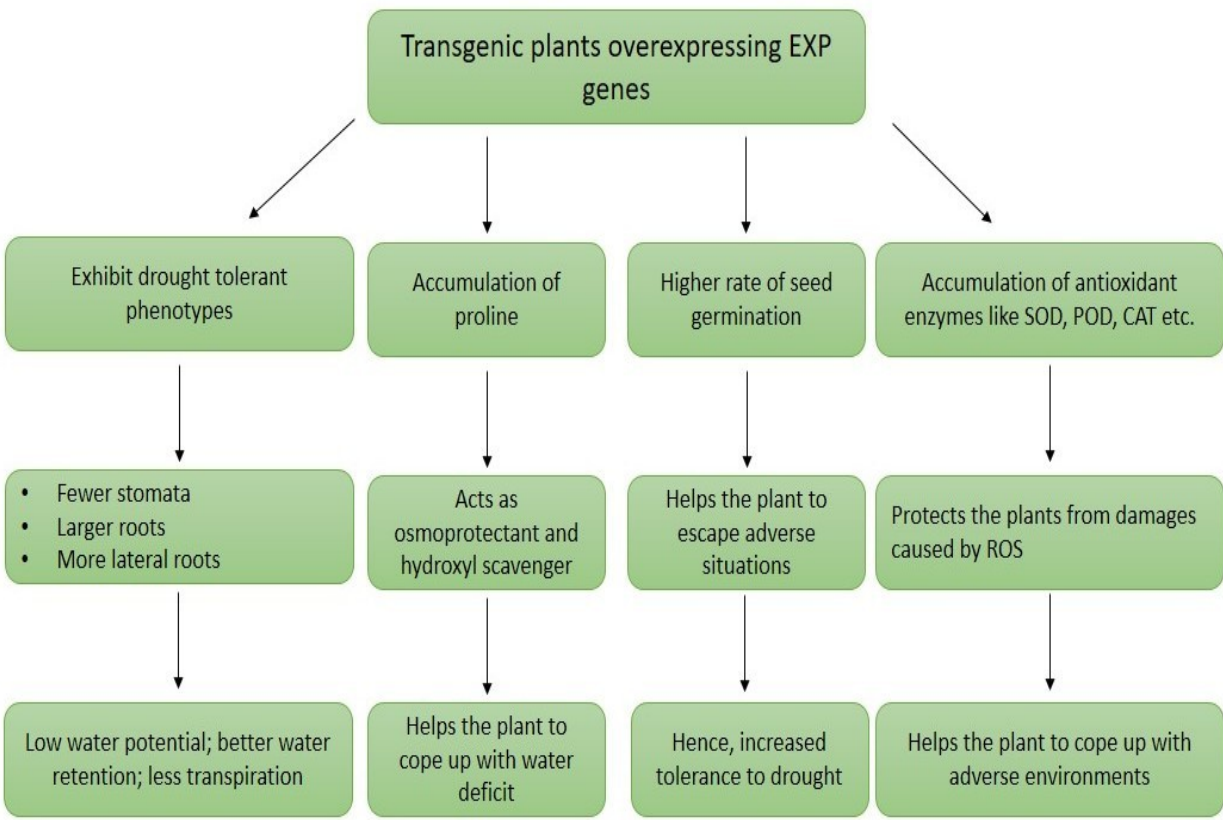

Fig. 6. A flowchart demonstrating the possible role of expansins in mitigating drought tolerance in earlier studied plants.

In tea, out of the 40 expansins, alpha-expansins (EXPA) were dominant or in abundance than the other expansin categories in line with the earlier studies already mentioned above (20, 28, 30-32, 34, 51-58). Tea plant lacks EXLB expansins but at least one EXLB expansin is present in other plant species studied till date except for that of sugarcane (53) and wheat (56).

To understand the regulation of tea expansin expression, we used the PlantCARE program to analyze *cis-acting* elements in all tea expansin promoter regions. Three classes of *cis-acting* elements viz., light responsive elements, core promoter element around -30 transcription start site and a promoter and enhancer region are present in all the 40 tea expansins. Plants are affected by both abiotic and some biotic factors which can impact the expression of expansin genes. This ultimately leads the genes to involve in some defensive or adaptive responses. Several *cis-acting* elements responsive to different functional categories have been found to occur in the gene upstream regions of expansins. Presence of drought responsive *cis-acting* elements in 40% of the identified expansins in tea demonstrates the potential involvement of expansins in drought stress response. As seen in Fig. 4b, 16 expansins showed high levels of expression in drought conditions and subsequent re-watering of the plants restored the expression of the genes. This finding suggests the involvement of expansins in drought stress response in tea. Analysis of the presence of these *cis-acting* elements shows that they are vital components of genetic regulatory elements in the tea genome.

Differential gene expression analysis showed that although 8 out of 9 EXLA expansins got upregulated during severe drought stress in one of the public datasets (Fig. 4a), the overall differential expression pattern of the genes is not class specific. TEA022767 was observed to be highly upregulated during drought stress conditions in all the datasets analysed, while re-watering treatment restored its expression to normal (Fig. 4b). Here, it may be concluded that TEA022767 belonging to EXPA subfamily is an important gene for drought stress response in tea.

Conclusion

In this study, we performed a genome-wide analysis of the expansin family in tea and also carried out expression analysis using publicly available drought stress data. The finding of this study will provide a piece of first-hand information on tea expansin genes in relation to drought stress. Further investigation is required to identify the exact factors responsible for regulating the expression of drought induced expansins in tea. Future studies involving functional genomics analysis will further aid in understanding the biological roles of expansins in the tea plant.

Acknowledgements

Authors' contributions

KSB carried out data collection, analysis and interpretation; PD carried out data collection, analysis and interpretation; DBK took part in data analysis and critical revision of the article; NA conceived the study and involved in its design and coordination. All authors read and approved the final manuscript.

Supplementary files

Supplementary Fig. 1: Classification of tea expansins into its types through phylogenetic relationship with *Arabidopsis* expansins.

Supplementary Table 1: Presence of cis-acting elements in the upstream of promoter regions of tea expansin genes.

Supplementary Table 2: Data on synonymous and non-synonymous substitution rates of tea expansins with their orthologs in *Arabidopsis*.

Conflict of interests

Authors do not have any conflict of interests to declare.

References

1. McQueen-Mason S, Cosgrove DJ. Disruption of hydrogen bonding between plant cell wall polymers by proteins that induce wall extension. Proc Natl Acad Sci. 1994;91(14):6574-78. https://doi.org/10.1073/pnas.91.14.6574

2. Choi D, Cho HT, Lee Y. Expansins: expanding importance in plant growth and development. Physiol Plant. 2006;126:511–18. https://doi.org/10.1111/j.1399-3054.2006.00612.x

3. Cho HT, Cosgrove DJ. Regulation of root hair initiation and expansin gene expression in *Arabidopsis*. Plant Cell. 2002;14:3237–53. https://doi.org/10.1105/tpc.006437

4. Che J, Yamaji N, Shen RF, Ma JF. An Al-inducible expansin gene, OsEXPA10 is involved in root cell elongation of rice. Plant J. 2016;88:132–42. https://doi.org/10.1111/tpj.13237

5. Balestrini R, Cosgrove DJ, Bonfante P. Differential location of α-expansin proteins during the accommodation of root cells to an arbuscular mycorrhizal fungus. Planta. 2005;220:889–99. https://doi.org/10.1007/s00425-004-1431-2

6. Flemetakis E, Efrose RC, Desbrosses G, Dimou M, Delis C, Aivalakis G, Udvardi MK, Katinakis P. Induction and spatial organization of polyamine biosynthesis during nodule development in *Lotus japonicus*. Mol Plant Microbe Interact. 2004;17:1283–93. https://doi.org/10.1094/MPMI.2004.17.12.1283

7. Belfield EJ, Ruperti B, Roberts JA, McQueen-Mason S. Changes in expansin activity and gene expression during ethylene-promoted leaflet abscission in *Sambucus nigra*. J Exp Bot. 2005;56(413):817–23. https://doi.org/10.1093/jxb/eri076

8. Goh HH, Sloan J, Dorca-Fornell C, Fleming A. Inducible repression of multiple expansin genes leads to growth suppression during leaf development. Plant Physiol. 2012;159(4):1759–70. https://doi.org/10.1104/pp.112.200881

9. Kuluev BR, Knyazev AV, Mikhaylova EV et al. The role of expansin genes PtrEXPA3 and PnEXPA3 in the regulation of leaf growth in poplar. Russ J Genet. 2017;53: 651–60. https://doi.org/10.1134/S1022795417060084

10. Saito T, Tuan PA, Katsumi-Horigane A et al. Development of flower buds in the Japanese pear (*Pyrus pyrifolia*) from late autumn to early spring, Tree Physiol. 2015;35(6):653–62. https://doi.org/10.1093/treephys/tpv043

11. Cho HT, Kende H. Expression of expansin genes is correlated with growth in deepwater rice. Plant Cell. 1997;9:1661–71. https://doi.org/10.1105/tpc.9.9.1661

12. Tabuchi A, Li LC, Cosgrove DJ. Matrix solubilization and cell wall weakening by b-expansin (group-1 allergen) from maize pollen. Plant J. 2011;68:546–59. https://doi.org/10.1111/j.1365-313X.2011.04705.x

13. Harmer S, Orford S, Timmis J. Characterisation of six α-expansin genes in *Gossypium hirsutum* (upland cotton). Mol Gen Genomics. 2002;268:1–9. https://doi.org/10.1007/s00438-002-0721-2

14. Palapol Y, Kunyamee S, Thongkhum M, Ketsa S, Ferguson IB, Van Doorn WG. Expression of expansin genes in the pulp and the dehiscence zone of ripening durian (*Durio zibethinus*) fruit. J Plant Physiol. 2015;182:33–39. https://doi.org/10.1016/j.jplph.2015.04.005

15. Nardi CF, Villarreal NM, Rossi FR, Martínez S, Martínez GA, Civello PM. Over expression of the carbohydrate binding module of strawberry expansin2 in *Arabidopsis thaliana* modifies plant growth and cell wall metabolism. Plant Mol Biol. 2015;88:101–17. https://doi.org/10.1007/s11103-015-0311-4

16. Brummell DA, Harpster MH, Dunsmuir P. Differential expression of expansin gene family members during growth and ripening of tomato fruit. Plant Mol Biol. 1999;39:161–69. https://doi.org/10.1023/A:1006130018931

17. Perini MA, Sin IN, Villarreal NM, Marina M, Powell AL, Martínez GA, Civello PM. Over expression of the carbohydrate binding module from *Solanum lycopersicum* expansin 1 (Sl-EXP1) modifies tomato fruit firmness and *Botrytis cinerea* susceptibility. Plant Physiol Biochem. 2017;113:122–32. https://doi.org/10.1016/j.plaphy.2017.01.029

18. Chen Y, Han Y, Zhang M, Zhou S, Kong X, Wang W. Over expression of the wheat expansin gene TaEXPA2 improved seed production and drought tolerance in transgenic tobacco plants. PLoS One. 2016;11:e0153494. https://doi.org/10.1371/journal.pone.0153494

19. Shcherban TY, Shi J, Durachko DM, Guiltinan MJ, McQueen-Mason SJ et al. Molecular cloning and sequence analysis of expansins-a highly conserved, multigene family of proteins that mediate cell wall extension in plants. Proc Natl Acad Sci U S A. 1995;92:9245–49. https://doi.org/10.1073/pnas.92.20.9245

20. Sampedro J, Cosgrove DJ. The expansin superfamily. Genome Biol. 2005;6(12):242. https://doi.org/10.1186/gb-2005-6-12-242

21. Cosgrove DJ. Assembly and enlargement of the primary cell wall in plants. Annu Rev Cell Dev Biol. 1997;13:171-201. 10.1146/annurev.cellbio.13.1.171

22. Yennawar NH, Li LC, Dudzinski DM, Tabuchi A, Cosgrove DJ. Crystal structure and activities of EXPB1 (*Zea* m 1), a β-expansin and group-1 pollen allergen from maize. Proc Natl Acad Sci. 2006;103(40):14664-71. https://doi.org/10.1073/pnas.0605979103

23. Cosgrove DJ. Loosening of plant cell walls by expansins. Nature. 2000;407:321–26. https://doi.org/10.1038/35030000

24. Cheruiyot EK, Mumera LM, Ngetich WK, Hassanali A, Wachira FN. High fertilizer rates increase susceptibility of tea to water stress. J Plant Nutr. 2009;33:115–29. https://doi.org/10.1080/01904160903392659

25. Yang J, Zhang G, An J, Li Q, Chen Y, Zhao X, Wu J, Wang Y, Hao Q, Wang W, Wang W. Expansin gene TaEXPA2 positively regulates drought tolerance in transgenic wheat (*Triticum aestivum* L.). Plant Sci. 2020;298:110596. https://doi.org/10.1016/j.plantsci.2020.110596

26. Lü P, Kang M, Jiang X, Dai F, Gao J, Zhang C. RhEXPA4, a rose expansin gene, modulates leaf growth and confers drought and salt tolerance to *Arabidopsis*. Planta. 2013;237(6):1547-59. https://doi.org/10.1007/s00425-013-1867-3

27. Lee Y, Choi D, Kende H. Expansins: ever-expanding numbers and functions. Curr Opin Plant Biol. 2001;4:527-32. https://doi.org/10.1016/s1369-5266(00)00211-9

28. Zhang W, Yan H, Chen W, Liu J, Jiang C, Jiang H, Zhu S, Cheng B. Genome-wide identification and characterization of maize expansin genes expressed in endosperm. Mol Genet Genomics. 2014;289(6):1061-74. https://doi.org/10.1007/s00438-014-0867-8

29. Li N, Pu Y, Gong Y et al. Genomic location and expression analysis of expansin gene family reveals the evolutionary and functional significance in *Triticum aestivum*. Genes Genom. 2016;38:1021–30. https://doi.org/10.1007/s13258-016-0446-y

30. Zhu Y, Wu N, Song W, Yin G, Qin Y, Yan Y, Hu Y. Soybean (*Glycine max*) expansin gene superfamily origins: segmental and tendem duplication events followed by divergent selection among subfamilies. BMC Plant Biol. 2014;14:93. https://doi.org/10.1186/1471-2229-14-93

31. Ding A, Marowa P, Kong Y. Genome-wide identification of the expansin gene family in tobacco (*Nicotiana tabacum*). Mol Genet Genomics. 2016;291:1891–907. https://doi.org/10.1007/s00438-016-1226-8

32. Lu Y, Liu Lifeng, Wang X, Han Z,Ouyang B, Zhang J, Li H. Genome-wide identification and expression analysis of the expansin gene family in tomato. Mol Genet Genomics. 2016;291:597-608. https://doi.org/10.1007/s00438-015-1133-4

33. Zhang S, Xu R, Gao Z et al. A genome-wide analysis of the expansin genes in *Malus × Domestica*. Mol Genet Genomics. 2014;225–36 (2014). https://doi.org/10.1007/s00438-013-0796-y

34. Dal Santo S, Vannozzi A, Tornielli GB, Fasoli M, Venturini L, Pezzotti M et al. Genome-wide analysis of the expansin gene superfamily reveals grapevine-specific structural and functional characteristics. PLoS One. 2013;8(4):e62206. https://doi.org/10.1371/journal.pone.0062206

35. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. The Proteomics Protocols Handbook. 2005; (edited) Walker, J.M., Totowa, N.J.: Humana Press. Available from https://link.springer.com/protocol/10.1385/1-59259-890-0:571

36. Wan S, Mak MW, Kung SY. FUEL-mLoc: feature-unified prediction and explanation of multi-localization of cellular proteins in multiple organisms. Bioinformatics. 2017;33:749–50. https://doi.org/10.1093/bioinformatics/btw717

37. Petersen TN, Brunak S, Von Heijne G, Nielsen H. SignalP 4.0: Discriminating signal peptides from transmembrane regions. Nat Methods. 2011;8:785. https://doi.org/10.1038/nmeth.1701

38. Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. GSDS 2.0: An upgraded gene feature visualization server. Bioinformatics. 2015;31:1296–97. https://doi.org/10.1093/bioinformatics/btu817

39. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Noble WS. MEME SUITE: Tools for motif discovery and searching. Nucleic Acids Res. 2009;37:W202–08. https://doi.org/10.1093/nar/gkp335

40. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S. PlantCARE, a database of plant *cis-acting* regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 2002;30:325-27. https://doi.org/10.1093/nar/30.1.325

41. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35(6):1547–49. https://doi.org/10.1093/molbev/msy096

42. Goodstein DM, Shu SQ, Howson R, Neupane R, Hayes RD, Fazo J et al. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 2012;40:D1178–86. https://doi.org/10.1093/nar/gkr944

43. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 2006;34:W609–12. https://doi.org/10.1093/nar/gkl315

44. Wei C, Yang H, Wang S et al. Draft genome sequence of *Camellia sinensis* var. *sinensis* provides insights into the evolution of the tea genome and tea quality. Proc Natl Acad Sci. 2018;115(18):E4151-58. https://doi.org/10.1073/pnas.1719622115

45. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT,

StringTie and Ballgown. Nat Protoc. 2016;11:1650–67. https://doi.org/10.1038/nprot.2016.095

46. Anders S, Huber W. Differential expression analysis for sequence countdata. Genome Biol. 2010;11:R106. https://doi.org/10.1186/gb-2010-11-10-r106

47. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Nucleic Acids Res. 2015;43:W566-70. https://doi.org/10.1093/nar/gkv468

48. Shannon P, Markiel A, Ozier O, Baliga NS, Wan JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504. https://dx.doi.org/10.1101%2Fgr.1239303

49. Marowa P, Ding A, Kong Y. Expansins: roles in plant growth and potential applications in crop improvement. Plant Cell Rep. 2016;35:949-65. https://doi.org/10.1007/s00299-016-1948-4

50. Koech RK, Malebe PM, Nyarukowa C et al. Functional annotation of putative QTL associated with black tea quality and drought tolerance traits. Sci Rep. 2019;9:1465. https://doi.org/10.1038/s41598-018-37688-z

51. Krishnamurthy P, Hong JK, Kim JA, Jeong MJ, Lee YH, Lee SI. Genome-wide analysis of the expansin gene superfamily reveals *Brassica rapa*-specific evolutionary dynamics upon whole genome triplication. Mol Genet Genomics. 2015;290(2):521-30. https://doi.org/10.1007/s00438-014-0935-0

52. Hou L, Zhang Z, Dou S, Zhang Y, Pang X, Li Y Genome-wide identification, characterization, and expression analysis of the expansin gene family in Chinese jujube (*Ziziphus jujuba* Mill.). Planta. 2019;249(3):815-29. https://doi.org/10.1007/s00425-018-3020-9

53. Santiago TR, Pereira M, de Souza WR, Steindorff A, Cunha B, Gaspar M, Favaro LC, Formighieri EF, Kobayashi AK, Molinari HB. Genome-wide identification, characterization and expression profile analysis of expansins gene family in sugarcane (*Saccharum* spp.). PLoS One. 2001;13:e0191081. https://doi.org/10.1371/journal.pone.0191081

54. Gao WLD, Fan X, Sun Y, Han B, Wan X, Xu G. Genome-wide identification, characterization, and expression analysis of the expansin gene family in watermelon (*Citrullus lanatus*). 3 Biotech. 2020;10(7). https://doi.org/10.1007/s13205-020-02293-3

55. Guimaraes LA, Mota APZ, Araujo ACG, de Alencar Figueiredo LF, Pereira B, de Passos Saraiva MA, Brasileiro ACM. Genome-wide analysis of expansin super family in wild *Arachis* discloses a stress-responsive expansin-like B gene. Plant Mol Biol. 2017;94(1-2):79–96. https://doi.org/10.1007/s11103-017-0594-8

56. Zhang JF, Xu YQ, Dong JM, Peng LN, Feng X, Wang X, Li FL. Genome-wide identification of wheat (*Triticum aestivum*) expansins and expansin expression analysis in cold-tolerant and cold-sensitive wheat cultivars. PLoS One. 2018;13(3):e0195138. https://doi.org/10.1371/journal.pone.0195138

57. Lv LM, Zuo DY, Wang XF, Cheng HL, Zhang YP, Wang QL, Ma ZY. Genome-wide identification of the expansin gene family reveals that expansin genes are involved in fibre cell growth in cotton. BMC Plant Biol. 2020;20:1. https://doi.org/10.1186/s12870-020-02362-y

58. Han Z, Liu Y, Deng X, Liu D, Liu Y, Hu1Y and Yan Y. Genome-wide identification and expression analysis of expansin gene family in common wheat (*Triticum aestivum* L.). BMC Genom. 2019;20:101. https://doi.org/10.1186/s12864-019-5455-1

59. Liu S-C, Jin J-Q, Ma J-Q, Yao M-Z, Ma C-L, Li C-F, et al. Transcriptomic analysis of tea plant responding to drought stress and recovery. PLos One. 2016;11:1–21. https://doi.org/10.1371/journal.pone.0147306

60. Sun J, Qiu C, Ding Y, Wang Y, Sun L, Fan K, et al. Fulvic acid ameliorates drought stress-induced damage in tea plants by regulating the ascorbate metabolism and flavonoids biosynthesis. BMC Genom. 2020;21:411. https://doi.org/10.1186/s12864-020-06815-4