Characterization of a Calmodulin-binding Transcription Factor from Strawberry (Fragaria × ananassa)

Xiangpeng Leng, Jian Han, Xiaomin Wang, Mizhen Zhao, Xin Sun, Chen Wang,* and Jinggui Fang

Abstract
Calmodulin-binding transcription activator (CAMTA) is a calmodulin-binding transcription factor that has a broad range of functions from sensory mechanisms to regulating many growth and developmental processes. In this study, we isolated four strawberry CAMTA (FaCAMTA) genes using HMMER and BLAST analysis. The chromosome scaffold locations of these CAMTA genes in the strawberry genome were determined and the protein domain and motif organization (CG-1, transcription factor immunoglobulin, ankyrin [ANK] repeats, calmodulin-binding IQ motif) of FaCAMTAs were also assessed. All FaCAMTAs were predicted to be Ca- and calmodulin-binding proteins. The expression profiles of FaCAMTA genes were measured in different tissues and revealed distinct FaCAMTA gene expression patterns under heat, cold, and salt stress. These data not only contribute to a better understanding of the complex regulation of the FaCAMTA gene family but also provide evidence supporting the role of CAMTAs in multiple signaling pathways involved in stress responses. This investigation can provide useful information for further study.

Plants have to endure stresses such as temperature extremes, ultraviolet light, salt, and pathogen attacks by modulating the expression of specific genes. Regulated gene expression is one of the most complex activities in cells and involves many transcription factors (TFs) (Yang and Poovaiah, 2002). Transcription factors play a crucial role in regulating every aspect of the organism’s lifecycle and are able to respond to signals originating from within and outside the organism. Calcium (Ca2+) plays a key role in regulating gene transcription (Ikura et al., 2002). Accumulating evidence indicates that Ca2+ regulates many growth and developmental processes and plays a crucial role in stress signaling and adaptation (Poovaiah et al., 1993; Snedden and Fromm, 2001; Reddy and Reddy, 2004) in response to various biotic and abiotic stresses (Knight, 1999; Rentel and Knight, 2004). Within cells, Ca signatures are perceived by the elongation factor-hand families of Ca-modulated proteins (calmodulin (CaM), Ca-dependent protein kinase and calcineurin B-like protein), which are well characterized in plants (Zielinski, 1998; Luan et al., 2002; Harmon, 2007).
Calmodulin is one of the most well characterized Ca sensors and functions as a modulator through the regulation of numerous target proteins that play roles in metabolism, ion transport, transcriptional regulation, protein phosphorylation, and other critical functions (Zielinski, 1998; Yang and Poovaiah, 2002, 2003; Ranty et al., 2006). Previous studies have shown that CaM is a highly conserved and multifunctional regulatory protein in eukaryotes. Its regulatory activities are triggered by its ability to modulate the activity of a certain set of CaM-binding proteins after binding to Ca\(^{2+}\), thereby generating physiological responses to various stimuli (Zielinski, 1998; Snedden and Fromm, 2001; Yang and Poovaiah, 2002).

In recent years, an increasing number of CaM-binding proteins in plants have been reported. Calmodulin-binding transcription activators, also referred to as signal-responsive proteins or ethylene-induced CaM-binding proteins, were first discovered in plants in a screen for the CaM-binding proteins with nonspecific DNA-binding activity (Bouché et al., 2002; Yang and Poovaiah, 2002; Reddy et al., 2011). This family of TFs is highly conserved and possesses multiple domains: (i) a highly conserved DNA-binding domain containing a predicted bipartite nuclear localization signals (NLS) in the N-terminal portion, designated CG-1 (the binding site includes the CGCG and CGTG motif); (ii) a transcription factor immunoglobulin (TIG) domain reported to be involved in nonspecific DNA contacts in various transcription factors; (c) ANK repeats, which are known to be involved in protein–protein interactions and present in a large number of functionally diverse proteins; and (ii) calmodulin-binding IQ motifs, known as CaM-binding sites, localized in the C-terminal part of CAMTAs (Bouché et al., 2002; Du et al., 2009; Reddy et al., 2011).

Calmodulin-binding transcription activator genes have also been identified in the genomes of multicellular organisms, including mammals, flies, and worms (Han et al., 2006; Song et al., 2006). In *Arabidopsis thaliana* (L.) Heynh., there are six CAMTA (CAMTA1–6), the transcript of which accumulates (upregulation) or diminishes (downregulation) rapidly and transiently in response to various abiotic and biotic stress (Galon et al., 2008; Reddy et al., 2011; Pandey et al., 2013). Each member has distinct or overlapping spatial and temporal expression patterns in different plant developmental stages under various biotic and abiotic stresses. CAMTA transcript levels are induced under cold and heat treatment (CAMTA1 and CAMTA3–6) as well as under salinity (CAMTA1–4 and CAMTA6) (Yang and Poovaiah, 2002). Furthermore, CAMTA expression responds to phytohormones and secondary messengers known to mediate plant responses to biotic and abiotic stress, such as abscisic acid (CAMTA2 and CAMTA4–6), methyl jasmonate (CAMTA1, 3, and 4), ethylene (CAMTA1, 3, and 4), H\(_2\)O\(_2\) (CAMTA2–6), salicylic acid (CAMTA2 and CAMTA4–6), and auxin (CAMTA1) (Galon et al., 2010a; Yang and Poovaiah, 2002). All CAMTAs are induced on wounding (Yang and Poovaiah, 2002). The patterns of induction to multiple chemical and physical stimuli suggest the involvement of individual CAMTAs in multiple signal transduction pathways and stress responses. For example, CAMTA1 repressor lines and camt1 mutants showed enhanced responsiveness to auxin, suggesting that in wild-type plants, enhanced expression of CAMTA1 in response to stresses suppresses the plant’s responsiveness to auxin (Galon et al., 2010b). During developmental stages, CAMTA3 knockout plants accumulate high levels of reactive oxygen species and showed enhanced resistance toward fungal and bacterial pathogens by suppressing plant responses. It negatively regulates the defense response to pathogens and interacts with the WRKY33 TF in CAMTA3 mutants (Galon et al., 2008; Du et al., 2009). The promoters of the drought-responsive element binding protein 1C and ZAT12 binds with CAMTA3 in plants (Doherty et al., 2009), indicating Ca signal-driven gene expression. Besides various findings on function of CAMTA protein on stress physiology. Mitsu‐suda et al. (2003) were the first to report the gene downstream of the CAMTA protein and showed pollen-specific expression of AtCAMTA1 and AtCAMTA5, which possibly increased *A. thaliana* vacuolar H\(^+\)-pyrophosphatase (AVP1) gene expression in pollen by binding to the pollen-specific cis-acting region of AVP1. Another study by Yang et al. (2012) in tomato (*Solanum lycopersicum* L.) indicates that SISR expression is influenced by both the ripening inhibitor mutant-mediated developmental network and ethylene signaling. The results suggest that Ca signaling is involved in the regulation of fruit development and ripening through Ca–CaM–SISR interactions. In brief, CAMTA family genes are very important for plant growth and development and reveal differential responses to developmental signals and a variety of environmental signals.

Strawberry (*Fragaria × ananassa*) is one of the most economically important fruit crops worldwide and has a long history of cultivation as well as high value economically and nutritionally (Dong et al., 2012). With the release of the strawberry genome sequence and development of bioinformatics tools and resources (http://bioinformatics.ca/links_directory/), accessed 27 Nov. 2014, mining of strawberry family genes is becoming possible, which is important to research on the functional genomics of strawberry. In this study, two sequence search tools and a strawberry genome dataset were used for FaCAMTA family gene identification, accompanied with sequence alignment and protein structural analysis. In addition, the reliability of predicted CAMTA members was also analyzed. The results of this study can be useful in further functional analysis of the CAMTA family genes in strawberry.

**Materials and Methods**

**Retrieval of CAMTA Gene Sequences**

The strawberry genome sequence has been completed and filtered protein and coding sequences have also become available. Several approaches were used to identify the CAMTA genes from strawberry. A whole-genome peptide sequence dataset of strawberry was downloaded from the
Genome Database for Rosaceae (http://www.rosaceae.org/, accessed 28 Nov. 2014). HMMPER version 3.0 software was obtained from the HMMPER website (http://hmmer.janelia.org/, accessed 28 Nov. 2014), and BLAST (stand-alone) software was retrieved from NCBI (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/, accessed 28 Nov. 2014). HMMPER and BLAST were used to screen the putative CAMTA genes from strawberry genome sequences using default search parameters. The CAMTA family gene domain model file (CG-1, pfam03859; IQ, pfam00612; ANK, pfam00023) was downloaded from the Sanger database (http://pfam.sanger.ac.uk/family/ PF03859#curationBlock). Mining all the strawberry peptide sequences was carried out by HMMPER software. All the putative FaCAMTA peptide sequences were submitted to the NCBI Conserved Domains Search website tool (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi, accessed 28 Nov. 2014) to check whether they contained the CG-1, ANK, and IQ domains. Finally, BLAST was used to remove any repetitive sequences of FaCAMTA family genes.

Multiple Sequence Alignment and Phylogenetic Tree Construction

All putative FaCAMTA peptide sequences alignment were performed in MacVector version 11.04 (available at http://www.macvector.com/downloads.html, accessed 28 Nov. 2014) using the global alignment program ClustalW (Thompson et al., 1994) with the default parameters. The consequential alignment was used to create a phylogenetic tree in MEGA version 4.0 (available at http://www.megasoftware.net, accessed 28 Nov. 2014) using the Neighbor-Joining method (Saitou and Nei, 1987), with bootstrapping set at 1000 replications.

Chromosome Location and Conserved Domain Search

Genomic locations of putative strawberry CAMTAs were identified through the NCBI-BLAST function in the Genome Database for Rosaceae website (http://www.rosaceae.org/tools/ncbi_blast, accessed 4 Feb. 2015). Conserved CG-1, TIG, and IQ domains were identified in putative multiple domains of CAMTAs using the Conserved Domain Database (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi, accessed 28 Nov. 2014) and verified by manual alignment with those in the corresponding A. thaliana CAMTAs (Bouché et al., 2002).

Protein Structure Analysis

The number of amino acids, the molecular weight, and the theoretical isoelectric point of putative strawberry CAMTAs were obtained from the Prot-Param analyses (http://cn.expasy.org/tools/protparam.html, accessed 28 Nov. 2014) (Kyte and Doolittle, 1982; Guruprasad et al., 1990) on the basis of their sequences. Protein secondary structure element prediction was conducted using the SOPMA server (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html, accessed 4 Feb. 2015) (Geourjon and Deléage, 1995). Putative amino acid sequences of various CAMTAs were submitted to the SWISS-MODEL workspace server (http://swissmodel.expasy.org, accessed 28 Nov. 2014) (Arnold et al., 2006). The proteins models were viewed and edited using Deepview/Swiss PDB Viewer version 4.0 software (http://spdbv.vital-it.ch, accessed 28 Nov. 2014).

Subcellular Localization Prediction

The computer programs used to predict the subcellular locations were WoLF PSORT (http://wolfpsort.org, accessed 4 Feb. 2015) and BaCelLo (http://gpcr.biocomp.unibo.it/bacello/, accessed 28 Nov. 2014).

Expressed Sequence Tag Retrieval and Function Prediction

Expressed sequence tags (ESTs) corresponding to the CAMTA genes in strawberry were isolated from the NCBI EST database (http://www.ncbi.nlm.nih.gov/nucest, accessed 28 Nov. 2014) using the BLAST program and specifying the target organism from the ‘Organism’ menu, using the FaCAMTA nucleotide sequences as queries. Each putative FaCAMTA expression profile and biological function was accomplished via an in silico strawberry expression analysis platform developed in our lab (Shangguan et al., 2013).

Plant Material and Stress Treatment

The strawberry cultivar used in our study was ‘Ningyu’, which was a new, early-maturing strawberry cultivar with disease resistance derived from ‘Sachinoka’×’Akihime’. Strawberry plants were grown in a greenhouse at 24 ± 2°C with a photoperiod of 14 h light and 10 h dark. Young leaves (20 d old), mature leaves (50 d old), flowers (fully open), small berries, mature berries, and young roots were collected in 2013 from Ningyu grown in a greenhouse. Strawberry plantlets grown for 6 wk were used for different stress treatments. For the cold and heat treatments, strawberry plants were incubated at 4 or 42°C (Yu et al., 2012); for the salt stress treatment, 200 mM NaCl was applied to the soil; for the ethylene treatment, 100 µM ethylene was sprayed onto the strawberry plants. All plant material was harvested at time courses of 2 and 12 h after treatment. In all cases, parallel and untreated plants at the same stage were used as controls. All the treatments were performed at room temperature, except the temperature stress treatment. After each treatment, whole plants were collected and immediately frozen in liquid nitrogen and stored at −80°C until RNA extraction.

RNA Isolation and Expression Analysis

Strawberry RNA extraction followed a simple and efficient modified cetyl trimethyl ammonium bromide method. Briefly, 0.05 g of samples was ground in a mortar with liquid N. The powdered tissue was transferred to a microfuge tube containing 500 µL 2% CTAB buffer [2% CTAB, 100 mM tris(hydroxymethyl)-aminomethane (pH 8.0), 20 mM ethylenediaminetetraacetic acid, and 1.4
M NaCl) with 2% β-mercaptoethanol added just before use. Samples were incubated for 20 min at 65°C. Suspensions were extracted twice with chloroform and isoamyl alcohol (24:1). To the water phase, 1:4 vol. 10 M LiCl was added and the RNA was precipitated for 60 min at −20°C. The precipitate was dissolved in tris(hydroxymethyl)aminomethane–ethylenediaminetetraacetic acid buffer and extracted with phenol, chloroform, and isoamyl alcohol and chloroform/isoamyl alcohol. The RNA was precipitated, washed twice, dried on a clean bench, and resuspended in 30 µL of diethylpyrocarbonate water. Genomic DNA was removed after 15 min incubation at 37°C with RNase-free DNase (Takara, Otsu, Japan). Next, an RNA clean purification kit was used to purify the sample RNA (Biotake Company, Beijing, China). RNA (2.0 µg) was converted to cDNA by reverse transcription, using the M-MLV RTase cDNA Synthesis Kit (Takara). The reverse transcription primers were P1 (GCAGGACTGAGCT-GACTGACTT30V5, 5′-3′) and P2 (GACCACTGG-TATCAACGCAGAGTACGCGGG, 5′-3′).

Semiquantitative reverse transcription polymerase chain reaction (PCR) was performed as described by Yu et al. (2009). The semiquantitative PCR program was 94°C for 5 min, followed by 38 cycles of denaturation for 30 s at 94°C, annealing for 30 s at 58.6°C, extension for 45 s at 72°C, and a final extension for 10 min at 72°C. Each reaction was repeated three times. A control PCR with primers that amplified a 180-bp fragment of Actin was performed in parallel to verify that there were similar amounts of cDNA in each sample. The PCR products were analyzed by electrophoresis on a 2.0% agarose gel containing ethidium bromide and photographed under ultraviolet light. PCR products with clear bands were collected from agarose gel and extracted with the AxyPrepTM DNA Gel Extraction Kit (Axygen Biosciences, Hangzhou, China). Details of the primers are shown in Supplemental Table 1.

Quantitative real-time PCR (qRT-PCR) was used to detect the expression pattern of FaCAMTA family genes under different stress conditions. The primers used are listed in Supplementary Table 1. Quantitative real-time PCR was performed using an ABI PRISM 7500 real-time PCR system (Applied Biosystems, Carlsbad, CA). Each reaction contained 10 µL 2x SYBR Green Master Mix Reagent (Applied Biosystem), 2.0 µL cDNA sample, and 400 nM of the gene-specific primer in a final volume of 20 µL. The thermal cycle was as follows: amplification reactions were initiated with a pre-denaturing step at 95°C for 10 s and followed by denaturing (95°C for 5 s), annealing (60°C for 10 s) and extension (72°C for 15 s) for 40 cycles during the second stage; a final stage of 55–95°C was used to determine the dissociation curves of the amplified products. The relative mRNA level for each gene was calculated as ΔΔCT values (Livak and Schmittgen, 2001) in comparison to unstressed seedlings (Applied Biosystems). The strawberry Actin gene was used as an internal control for normalization. At least three replicates of each cDNA sample were performed for qRT-PCR analysis.

<table>
<thead>
<tr>
<th>Table 1. Strawberry calmodulin-binding transcription activator (FaCAMTA) protein information.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FaCAMTA1</td>
</tr>
<tr>
<td>Gene ID</td>
</tr>
<tr>
<td>Protein length</td>
</tr>
<tr>
<td>Scf number†</td>
</tr>
<tr>
<td>Scf length‡</td>
</tr>
<tr>
<td>Scf position start§</td>
</tr>
<tr>
<td>Scf position end¶</td>
</tr>
<tr>
<td>Exon number</td>
</tr>
<tr>
<td>Molecular weight¶¶</td>
</tr>
<tr>
<td>Theoretical isoelectric point§§</td>
</tr>
<tr>
<td>Aliphatic index¶¶¶</td>
</tr>
<tr>
<td>GRAVY¶¶¶¶</td>
</tr>
<tr>
<td>Predicted localization††</td>
</tr>
</tbody>
</table>
| α-helix††† | 40.69% | 41.02% | 42.96% | 40.90%
| Extended strand¶¶ | 9.60% | 9.56% | 7.34% | 9.67%
| β-turn¶¶¶ | 5.28% | 6.12% | 4.46% | 5.97%
| Random coil¶¶¶¶ | 44.43% | 43.29% | 45.24% | 45.27%


‡ The BLAST result of strawberry CAMTA genes and strawberry scaffold, containing scaffold (Scf) number, Scf length, start and end in the scaffold.

§ The molecular weight, isoelectric point, aliphatic index, and GRAVY of the full-length protein were predicted by the ProtParam analyses (http://cn.expasy.org/tools/protparam.html, accessed 1 Dec. 2014).


† Protein secondary structure element prediction was conducted using the SOPMA server (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html, accessed 4 Feb. 2015).

†† Nuc, Nucleus; Chlo, Chloroplast; GRAVY, grand average of hydropathicity.

Results

Identification and Physical Locations of CAMTA Proteins in Strawberry

There are six CAMTAs in the A. thaliana genome (AtCAMTA1–6) and all of them are expressed in tissues (Yang and Poovaiah, 2002). To identify FaCAMTA family genes from the strawberry genome sequence dataset, HMMER and BLAST searches were performed using the CAMTA family domain file (CG_1, pfam03859; IQ, pfam00612; ANK, pfam00023) as a query. After conserved domain and BLAST analysis, four FaCAMTA protein sequences were identified and described in strawberry (Table 1). For the nomenclature of strawberry CAMTAs, we followed the previously reported rule for A. thaliana: the proteins were named CAMTA for the DNA-binding domain and labeled according to their orthology with reported isoforms in A. thaliana and on their estimated molecular weights (Table 1). In some cases, similarities in domain structures and isoelectric point have also been considered. Compared with A.thaliana CAMTA family genes, the homologous genes to AtCAMTA2 and AtCAMTA6 were not identified in strawberry.
The CAMTA proteins consist of multiple predicted functional domains, evolutionally conserved in amino acid sequences and organized in a conserved order. Alignment of the amino acid sequences of CAMTAs using the NCBI Conserved Domains Search website tool and comparisons with protein domain databases (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi, accessed 4 Feb. 2015) revealed different types of conserved regions in all FaCAMTAs. The conserved domains include the following: a conserved DNA-binding domain in the N-terminal portion of all CAMTA proteins, a CaM-binding domain in the C-terminal portion, and Ank repeats in the middle, suggesting that FaCAMTAs are true orthologs of the CAMTA transcription factor family (Fig. 1). In addition, FaCAMTA1 has an immunoglobulin-like fold–plexins–transcription factors (IPT) domain and FaCAMTA3, FaCAMTA4 and FaCAMTA5 contain a TIG domain. Both the IPT and TIG domains are located between the N-terminal DNA-binding domain and the Ank repeats (Fig. 1). Although the spacing is highly variable, the overall domain organization is conserved in all proteins.

According to the result of multiple sequence alignments, the deduced amino acid sequences of the four FaCAMTAs have 36.9, 45.9, 50.0, and 47.4% overall similarity compared with the corresponding A. thaliana sequence, respectively (Supplemental Fig. 1). Although CG-1 is a highly conserved uncharacterized domain of about 110 amino acid residues in A. thaliana, the length of the conserved domain diverges among different members of the FaCAMTA family. The CG-1 domain of FaCAMTA1 and FaCAMTA4 have about 110 amino acid residues; however, the CG-1 domain of FaCAMTA3 and FaCAMTA5 have about 160 and 40 amino acid residues, respectively (Fig. 2).

Phylogenies are useful for organizing knowledge of biological diversity, for structuring classifications, and for providing insight into events that occurred during evolution. The phylogenetic relationship of FaCAMTA-encoded polypeptides and A. thaliana, tomato, and rice (Oryza sativa L.) CAMTAs was further analyzed. FaCAMTA3 is closely related to SISR1 and AtCAMTA3, FaCAMTA5 is in the same clade as AtCAMTA5 and AtCAMTA6, and FaCAMTA4 is closely related to SISR2, SISR2L, and AtCAMTA4. In contrast, FaCAMTA1 is in the same clade as SISR4 and has a weak homology to AtCAMTA1 and AtCAMTA2 (Fig. 3).

All FaCAMTAs Encode Ca- and CaM-binding Proteins
Calmodulin-binding motifs were detected in all CAMTAs except in that from Caenorhabditis elegans (Bouché et al., 2002). These motifs often appear adjacent to other CaM-binding domains and mediate complex regulatory properties in the presence and/or absence of Ca2+ (Bähler and Rhoads, 2002). This raises the possibility that most CAMTAs are regulated through direct binding of CaM. All CAMTA proteins reported thus far are known to be Ca- and CaM-binding proteins. To ascertain whether FaCAMTAs encode Ca- and CaM-binding proteins, their putative CaM-binding regions were aligned with the corresponding regions in AtCAMTAs. The corresponding sites in AtCAMTAs are well characterized Ca- and CaM-binding regions and share very high homology with their counterparts in the FaCAMTAs. Within this region, FaCAMTA4 has almost the same amino acid sequence as AtCAMTA4; FaCAMTA5 shows high similarity to AtCAMTA5 and AtCAMTA6 (Fig. 4A).

Calmodulin is known to bind to peptides 12 to 30 amino acids long that tend to form amphipathic
Figure 2. The length of the CG-1 domain is distributed among different members of the strawberry calmodulin-binding transcription activator (CAMTA) family. Domains were identified by alignment of CAMTA proteins with the ClustalX program.

Figure 3. Phylogenetic analysis of amino acid sequences encoded by four strawberry calmodulin-binding transcription activators (CAMTAs) and CAMTA orthologs from various species of plants. A neighbor-joining phylogenetic tree was constructed with MEGA5 software. CAMTA accession numbers: Arabidopsis thaliana: AtCAMTA1, NP_196503.3; AtCAMTA2, NP_201227.3; AtCAMTA3, NP_850023.1; AtCAMTA4, NP_176899.2; AtCAMTA5, NP_193350.5; AtCAMTA6, NP_188319.2. Solanum lycopersicum: SISR1, GU170838; SISR1L, JN558810; SISR2, JN566047; SISR2L, JN566048; SISR3, JN566049; SISR3L, JN566051; SISR4, JN566050. Oryza sativa: Os01g69910, Os03g09100, Os03g27080, Os04g31900, Os07g43030, and Os10g22950.
α-helices with one face of the helix being positively charged (Vogel, 1994; Arazì et al., 1995). All can form a basic amphipathic α-helix structure (Arazì et al., 1995), which can be recognized by CaM. We selected FaCAMTA5 as an example to show a basic amphipathic α-helix structure projection (Fig. 4B). This analysis revealed an 18-amino acid region with typical CaM-binding characteristics between the amino acids valine 873 and leucine 890. When drawn in the form of an α-helix structure, it exhibits an amphipathic structure with a positively charged binding face and an opposing hydrophobic face (Fig. 4B). The results indicate that all FaCAMTAs are Ca- and CaM-binding proteins.

Subcellular Location Analysis and Scaffold Location
Predicting protein subcellular localization involves the computational prediction of where a protein resides in a cell. Predicting protein subcellular localization is an important component of bioinformatics-based prediction of protein function and genome annotation and it can aid the identification of drug targets. WoLF PROST and BaCelLo are the most widely used computational methods for protein subcellular localization. According to the WoLF PROST results, all FaCAMTAs were predicted to be located in the nucleus. According to the BaCelLo results, FaCAMTA1 was predicted to be located in the chloroplast and FaCAMTA3, FaCAMTA4, and FaCAMTA5 were nuclear proteins (Table 1). Subcellular localization prediction of strawberry tissues revealed the presence of FaCAMTAs predominantly in the nucleus.

Four FaCAMTA family genes were mapped in different strawberry scaffolds. FaCAMTA1, FaCAMTA3, FaCAMTA4, and FaCAMTA5 were mapped on scf0513094, scf0513168, scf0513160, and scf0513154, respectively.

Expressed Sequence Tag Sequences and Function Prediction
Expressed sequence tag sequences can be used to identify gene transcripts and as an instrument in gene sequence determination and gene discovery. We investigated the EST numbers of each FaCAMTA through the strawberry EST database using NCBI’s BLASTn program and the cDNA sequences of each FaCAMTA as a query. The BLAST results (Table 2) showed that FaCAMTA1 and FaCAMTA3 had two homologous EST sequences. FaCAMTA4 contained five EST sequences and FaCAMTA5 had only one EST sequence. All ESTs

Table 2. Strawberry calmodulin-binding transcription activator (FaCAMTA) family genes’ expressed sequence tag (EST) retrieval and function prediction.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Homologous ESTs</th>
<th>Scaffold number</th>
<th>Clone library</th>
<th>Biological function</th>
<th>Total score</th>
<th>Identity</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FaCAMTA1</td>
<td>EX659634</td>
<td>scf0513094</td>
<td>Strain</td>
<td>Salt stressed</td>
<td>1673</td>
<td>99%</td>
<td>0.0</td>
</tr>
<tr>
<td>FaCAMTA3</td>
<td>DY672460</td>
<td>scf0513094</td>
<td>Seedlings</td>
<td>Cold stressed</td>
<td>1595</td>
<td>100%</td>
<td>0.0</td>
</tr>
<tr>
<td>FaCAMTA4</td>
<td>DY673027</td>
<td>scf0513168</td>
<td>Seedlings</td>
<td>Cold stressed</td>
<td>1615</td>
<td>99%</td>
<td>0.0</td>
</tr>
<tr>
<td>FaCAMTA5</td>
<td>DY669949</td>
<td>scf0513160</td>
<td>Seedlings</td>
<td>Cold stressed</td>
<td>1216</td>
<td>99%</td>
<td>0.0</td>
</tr>
</tbody>
</table>

*Figure 4.* Strawberry calmodulin-binding transcription activators (FaCAMTAs) are Ca- and calmodulin-binding proteins. (A) Alignment of the conserved putative calmodulin-binding regions of FaCAMTAs with AtCAMTAs. (B) Helical wheel projection showing the basic amphipathic α-helix structure in the predicted Ca- and calmodulin-binding domain of FaCAMTA5 amino acid residues (Val873–Leu890). By default the output presents the hydrophilic residues as circles, hydrophobic residues as diamonds, potentially negatively charged residues as triangles, and potentially positively charged residues as pentagons.
in strawberry could be mapped to scaffold locations. For example, the homologous ESTs (EX658634 and DY672460) of FaCAMTA1 were mapped on scf0513094, the same as FaCAMTA1. FaCAMTA3 (Table 2), FaCAMTA4 and FaCAMTA5 also had the same phenomenon.

Expressed sequence tag sequences have become a tool to refine the predicted genes transcripts, which leads to the prediction of their protein products and ultimately their function, after a process of removing vectors and redundant sequences. Moreover, the situation in which these ESTs are obtained (tissue, organ, disease state) gives information on the conditions in which the corresponding gene is acting. ESTs contain enough information to permit the design of precise probes for DNA microarrays that then can be used to determine the gene expression. The Strawberry in silico expression platform version 2.0 is a quick and reliable digital analysis platform that contains the information of 60,012 strawberry EST sequences. Expressed sequence tag sequences whose total score and similarity were larger than 100 and 90%, respectively, for each FaCAMTA family gene were selected. The result showed that FaCAMTA1 and FaCAMTA4 might play important roles in response to salt and cold stress. FaCAMTA3 may be linked to cold stress. FaCAMTA5 was expressed in the shoot apex. It should be noted that the function prediction and expression patterns were only a predicted result and although the genes may have related functions and be expressed in the predicted tissues, it does not necessarily mean that the genes had no other functions or expression patterns. Therefore, experimental confirmation is needed to supplement the prediction results.

The Expression Pattern of FaCAMTA Genes in Different Strawberry Tissues
Gene expression patterns not only verify the existence of the predicted genes but also provide important clues to gene function. To investigate the temporal and spatial expression profiles of these FaCAMTAs, we employed a semiquantitative reverse transcription PCR approach to detect the mRNA accumulation of each FaCAMTA gene in young leaves, mature leaves, root, flowers, young berries, and mature berries. By using semiquantitative PCR, a signal was detected at around 38 cycles under normal growth conditions, whereas the exponential stage for the internal control FaActin was detected at around 28 cycles. By combining these results together, we found that the FaCAMTA genes were constitutively expressed in different tissues at similar levels and the expression level of four FaCAMTAs was relatively low in strawberry. For instance, FaCAMTA1 and FaCAMTA4 were detected in all tissues but relatively higher transcript levels were observed in young berries. The FaCAMTA5 gene was expressed at lower levels in roots and flowers compared to other tissues, whereas FaCAMTA3 mainly resided in roots and flowers (Fig. 5). Collectively, our results indicate that the four FaCAMTAs genes are ubiquitously expressed, suggesting general roles for the remainder in regulating organ and tissue development.

Expression Analysis of the FaCAMTA Gene Family under Stress
Expression patterns of some CAMTAs were reported to change under stress conditions, including heat, cold, high salinity, drought, and ultraviolet light and to signal intermediates and phytohormones, such as ethylene, abscisic acid, salicylic acid, and methyl jasmonate (Yang and Poovaiah, 2002). The rapid response of CAMTA genes to these external chemical and physical stimuli suggests that they play a role in the cross-talk between multiple signal transduction pathways involved in stress tolerance. To examine changes in gene expression patterns under different forms of stress, quantitative real-time PCR was performed for four FaCAMTAs genes in 6-wk-old strawberry plantlets exposed to heat, cold, salt, and ethylene stresses. In these experiments, plants were exposed to stress conditions for short periods of time (4 h) to avoid possible secondary effects. These genes were variably expressed under different stress conditions. FaCAMTAs showed separate expression patterns in response to various stress. As shown in Fig. 6, FaCAMTA1 appeared to be rapidly and constantly suppressed at low (4°C) and high (42°C) temperatures; however, the expression level of FaCAMTA1 was increased at 2 h and then decreased at 12 h under salt and ethylene stress. The expression level of FaCAMTA3 was constantly increased both cold and salt stress, whereas it displayed contradictory expression patterns and decreased expression at 42°C and under ethylene stress. FaCAMTA4 was induced at low temperatures, high temperatures and under ethylene stress. However, the expression level of FaCAMTA4 increased at 2 h and then decreased at 12 h under salt stress. FaCAMTA5 was slowly induced by high temperature and almost had no change in response to cold and salt stress. However, the expression level of FaCAMTA5 was induced at 2 h
and then decreased at 12 h under ethylene stress. The qRT-PCR results were consistent with EST function prediction, suggesting that FaCAMTA1, -3, and -4 are more likely to be involved in low temperature signaling. Salt stress also had a similar effect but did not induce the expression of FaCAMTA5 (Fig. 6). Thus, although CAMTA genes are constitutively expressed under different conditions, several members exhibited distinct expression patterns under different forms of environment stress, implicating many of the strawberry CAMTA proteins in environmental signaling pathways.

Discussion

Transcription factors play a crucial role in regulating gene expression and are able to regulate diverse cellular processes by interacting with other proteins. A TF named CAMTA was identified in various eukaryotes including A. thaliana (Bouché et al., 2002), rapeseed (Brassica napus L.) (Bouché et al., 2002), rice (Choi et al., 2005), and tomato (Yang et al., 2012), as well as in human and Drosophila (Katoh and Katoh, 2003). Previous studies have shown that CAMTAs have a broad range of functions in phytohormones’ response to growth control and sensory mechanisms, highlighted by their apparent involvement in mediating plant responses to biotic and abiotic stress. Despite the completion of the genome sequence of strawberry, there have been no detailed reports on FaCAMTA family genes. Currently, there are more than 1800 bioinformatic resources available for the researchers (http://bioinformatics.ca/links_directory/, accessed 28 Nov. 2014). These tools have greatly accelerated the discovery of new genes in living organisms. In this study, we identified four CAMTA orthologs from strawberry, designated FaCAMTAs. With the help of bioinformatics, we have performed sequence- and bioinformatics-based characterizations of these genes and their proteins. These genes were found to be distributed in four scaffolds, unlike in A. thaliana, where CAMTAs were localized in the chromosomes. As the strawberry genome project develops, more and more scaffold sequences are likely to be mapped in the chromosomes of strawberry.

The CAMTA proteins consist of multiple functional domains. The protein’s primary structure contains four conserved regions arranged in the same colinear order, including a newly characterized domain designated CG-1, a TIG domain, ANK repeats, and a variable number of IQ motifs. In this study, all four FaCAMTAs contained four conserved regions and were arranged in the same colinear order. CG-1 domains are most likely to be sequence-specific DNA-binding domains and represent a new category of DNA-binding domain associated with CAMTA proteins. This should be regarded as an excellent incentive to continue exploring these proteins in different organisms while addressing structural and functional questions (Finkler et al., 2007). The TIG domain is a second type of DNA-binding
domain in all CAMTAs. TIG domains are involved in DNA contact but also in dimerization, whereas a different domain is typically involved in providing DNA-binding sequence specificity. CAMTAs might interact with DNA in a similar way, with the TIG domain being a nonspecific DNA-binding domain while the CG-1 domain provides sequence specificity. In strawberry, FaCAMTA3, FaCAMTA4, and FaCAMTA5 contain a TIG domain and FaCAMTA1 has an IPT domain, which is known as a TIG domain. They are present in intracellular transcription factors, cell surface receptors (such as plexins and scatter factor receptors), and cyclodextrin glycosyltransferase and similar enzymes. In these TFs, IPTs form homo- or heterodimers, with the exception of the nuclear factor of activated T-cell transcription factors, which are mainly monomers.

The Ank repeat is one of the most common protein-protein interaction motifs in nature. Ankyrin repeats are tandemly repeated modules of about 33 amino acids. They occur in a large number of functionally diverse proteins, mainly from eukaryotes. Ankyrin repeats, like other conserved domains with a specific secondary structure (e.g., Src homology 2 and 3 domains), evolved as a universal module mediating protein-protein interactions (Sedgwick and Smerdon, 1999; Bouché et al., 2002). The primary structure of the ANK motifs found in four FaCAMTAs is similar to a consensus sequence (Sedgwick and Smerdon, 1999) corresponding to a large number of ANK-containing proteins (data not shown). FaCAMTAs might therefore interact with other proteins or form heteromeric (or homomeric) complexes by means of their ANK domains. CAMTAs contain a variable number of IQ motifs. The IQ motifs consist of low-complexity regions with the repetitive motif IQXXRGXXX and are known to be associated with binding CaM and CaM-like proteins (Rhoads and Friedberg, 1997; Bähler and Rhoads, 2002). With the help of bioinformatics, we found that putative CaM-binding regions of FaCAMTAs share very high homology with their counterparts in the AtCAMTAs (Fig. 4A). All four FaCAMTAs show a basic amphipathic α-helix structure projection with typical CaM-binding characteristics (Fig. 4B). The results indicate that all FaCAMTAs are Ca- and CaM-binding proteins. The subcellular location of a protein is closely correlated to its function. Thus computational prediction of subcellular locations from the amino acid sequence information would help annotation and functional prediction of protein-coding genes in complete genomes (Park and Kanehisa, 2003). Most eukaryotic proteins are encoded in the nuclear genome and synthesized in the cytosol but many need to be sorted further before they reach their final destination. Proteins must be localized at their appropriate subcellular compartment to perform their desired function. Experimentally determining the subcellular localization of a protein is a laborious and time-consuming task. Through the development of new approaches in computer science, coupled with an increased dataset of proteins of known localization, computational tools can now provide fast and accurate localization predictions for many organisms. This has resulted in subcellular localization prediction becoming one of the challenges being successfully aided by bioinformatics. In this study, two publicly available computational tools, WoLF PSORT and BaCelLo, were used for predicting the subcellular localization of the FaCAMTAs. Predicting the subcellular localization of strawberry tissues revealed the presence of FaCAMTAs to be predominantly in the nucleus. These results were consistent with our previous study in A. thaliana (Bouché et al., 2002). Proteins targeted to the nucleus usually contain an NLS composed of the basic amino acids arginine, and lysine, organized in groups (Liu et al., 1999). A putative NLS was detected in the CG-1 domain of all CAMTAs identified so far. Therefore, this region might constitute a signal that directs FaCAMTAs to the nucleoplasm.

Abiotic and biotic stress is one of the major environmental factors limiting crop productivity worldwide. It is hoped that understanding how plants respond and adapt to a changing environment at the molecular level will help in developing plants that can better cope with stresses. In recent years, the CaM-binding proteins of CAMTAs have been identified; these play key roles in stress signaling pathways in plants, including responses to cold, heat, salt and ethylene stress. Temperature profoundly influences the metabolism of organisms and thus is a key factor determining the growing season and geographical distribution of plants. Cold stress adversely affects the growth and development of plants and significantly constraints the spatial distribution of plants and agricultural productivity (Chinnusamy et al., 2007). Most temperate plants exhibit cold acclimation, which involves changes in gene expression (Fowler and Thomashow, 2002; Kreps et al., 2002). The C-repeat binding factors (CBF1, 2, and 3) (also called DREB1B, 1C, and 1A, respectively) are TFs that induce the expression of a large number of genes (CBF regulon) involved in cold acclimation (Chinnusamy et al., 2007; Sakamoto et al., 2004; Vogel et al., 2005). Heat stress can disturb cellular homeostasis, impair plant growth, and even cause plant death. Plant thermotolerance can be achieved through accumulation of heat shock proteins, the transcription of which is tightly regulated by TFs (Reddy et al., 2011). Several lines of evidence have implicated Ca²⁺ in cold and heat acclimation. Recently, members of the CAMTA family proteins have been identified as transcriptional regulators of CBF2 expression. Cold-induced expression of CBF2 was considerably lower in the camta3 mutant than in wild-type plants. The CAMTA3 protein binds to conserved DNA motifs present in the CBF2 promoter and regulates CBF2 expression. The camta1/camta3 double mutant exhibited hypersensitivity to freezing stress compared with wild-type plants. Since CAMTA proteins can interact with CaMs, cold-induced Ca signals may regulate CBF expression through the CAMTA proteins (Doherty et al., 2009). CAMTA1, which is induced by auxin, has also been implicated in the heat shock response. CAMTA1: β-glucuronidase GUS plants
that were exposed to heat shock expressed CAMTA1 in leaf trichomes at the leaf base and in the root cortex (Galon et al., 2010b), whereas the promoter of an auxin-induced gene showed no expression. This suggests that CAMTA1 responds to heat shock through promoter elements that are different from those responding to auxin. In this study, both cold and heat stress differentially induced the expression of FaCAMTA genes. FaCAMTA1 was rapidly and constantly suppressed, FaCAMTA4 was slowly upregulated over 2 to 12 h by cold and heat stress, and the expression level of FaCAMTA3 was increased at 4°C and decreased at 42°C, suggesting a likely function in response to cold and heat signaling in these three genes.

Salt stress is a serious factor that limits the efficiency of crops, especially the quality and quantity of their metabolic products (secondary plant products). Salt stress is responsible for osmotic, ionic, and oxidative stresses, which lead to reduced plant growth and development (Ahmad et al., 2013). In response to salt stress, plants activate various signaling pathways, including those involving Ca2+ that promote adequate cellular responses (Zhu, 2002). The salt-overly-sensitive (SOS) pathway is important in the decoding of salt stress-mediated Ca2+ signatures (Chinnusamy et al., 2004; Mahajan et al., 2008). The SOS3 (calcineurin B-like 4)/SOS2 (CIPK24) complex regulates various signaling pathways, including those involving osmotic, ionic, and oxidative stresses, which may serve as one of the early hubs in multiple signal transduction cascades by differentially responding to the multiple upstream signals. Furthermore, CAMTAs recognize CGCG cis-elements and may regulate downstream gene expression, which ultimately leads to the physiological responses of a variety of stresses. Future studies on each individual member’s expression and the accurate mechanisms related to the FaCAMTAs involved in signaling pathways and stress responses need detailed investigation in the future. We expect that future research will provide the answer as to the biological role of the CAMTAs in plants and in mammals.

Acknowledgments

This work was supported by grants from the Important National Science & Technology Specific Projects (No. 2012FY110100-3), the Natural Science Foundation of China (No.31501759, No.31372027) and Special Program of China Postdoctoral Science Foundation (2014T70533).

References


