ISSN: 2690- 912X

Research Article

Journal of Genetic Engineering and Biotechnology Research

# Evolutionary and Expression Analysis of *CAMTA* Gene Family in Three Species (*Arabidopsis*, Maize and Tomato), and Gene Expression in Response to Developmental Stages

**Abbas Saidi[1*], Zohreh Hajibarat[1]**

[1]*Department of Plant Sciences and Biotechnology, Faculty of Life Sciences and Biotechnology, ShahidBeheshti University, Tehran, Iran.*

**[*]Corresponding author**
Abbas Saidi,Department of Plant Sciences and Biotechnology, Faculty of Life Sciences and Biotechnology, ShahidBeheshti University, Tehran, Iran

**Submitted:** 09 March 2022; **Accepted:** 15 March 2022; **Published:** 25 March 2022

***Citation:*** *Abbas Saidi, ZohrehHajibarat. (2022).Evolutionary and Expression Analysis ofCAMTA Gene Family in Three Species (Arabidopsis, Maize and Tomato), and Gene Expression in Response to Developmental Stages.J Gene Engg Bio Res, 4(1), 127-139.*

**Abstract**

*The calmodulin-binding transcriptional activator (CAMTA) family has been known to be one of the fast responsive stress proteins. In this study, 17 CAMTA genes were selected in Arabidopsis, tomato and maize. The chromosomal distributions, gene structures, duplication patterns, phylogenetic tree, and developmental stage of the 17 CAMTAgenes in the three species were analyzed to further investigate their functions. According to the synteny analysis, CAMTA genes of maize and tomatorevealed higher similarity with each other as compared with Arabidopsis. A higher than 90 percent identity was observed between maize CAMTA genes (ZmCAMTA2 and ZmCAMTA3) and tomato CAMTA genes (SlCAMTA4, SlCAMTA4.1). To detect expression levels in different plant tissues, mRNA analysis of CAMTA genes were performed using publicly available expression data in the genvestigator. The aim of study was to identify and characterize CAMTA genes in three species, for the first time, via insilico genome-wide analysis approach. AtCAMTA1 and AtCAMTA2 and SlCAMTA2 and ZmCAMTA1 and ZmCAMTA2 genes were up-regulated during all developmental stages. The conserved motifs and gene structure in most proteins in each group were similar, validating the CAMTA phylogenetic classification. This study could be considered as a useful source for future CAMTA comparative studies in different plant species.*

**Keywords:** *CAMTA*Gene Family, Gene Expression, Conserved Domain, Gene Structure

## Introduction

Plants are exposed to a variety of stress conditions such as abiotic and biotic stresses during their growth life cycle. These environmental conditions may activate pathways or mechanisms of signaling [1]. To identify responsive genes and transcription factors involved in environmental stress adaptation, interaction of TF with cis-elements was surveyed [2]. Due to the importance of TFs in regulation of stress related genes, increasing number of studies are being performed focusing on their functions. Several studies have shown that TFs affecting life cycles of organisms play a crucial role in plant and cell signaling [3].

In general, over 90 TFs have been known as CaM-binding proteins (CBPs), including CAMTAs, MYBs, WRKY, NACs, and MADS box proteins [4]. One of the most important TFs is calmodulin-binding transcriptional activators (CAMTA), containing four functional domains known as IPT/TIG (transcription factor immunoglobulin), CaMBD, and a varying number of IQ motifs (calmodulin-binding), CG-1 (a DNA-binding domain specific to sequence), and ankyrin (ANK) repeats. Calmodulin-binding transcription activators (CAMTAs), the well-studied CaM-binding transcription factors, are found in animals, fungi, and plants [5]. The binding site for CAMTAs are present in downstream promoter of genes and are designated as A/C/G) CGCG (T/C/G) or (A/C) CGTGT, which helps to regulate their expression [6].

The first CAMTA gene member, NtER1, was identified in tobacco while surveyingCaM-binding proteins [7]. In *Arabidopsis*, six CAMTA transporters, termed as AtCAMTA1 to AtCAMTA6, have been detected. These AtCAMTAs are involved in biotic, abiotic, and hormonal regulations and developmental stages [8]. For example, *AtCAMTA1* and *AtCAMTA3* genes play an important role in response to cold and drought stress and in regulation of auxin and salicylic acid in plants. AtCAMTA6 genes are expressed in plants in response to salt stress, while AtCAMTA4 performs pivotal role in defense responses of plants against pathogens, and in re-

sponse to ethylene, jasmonate acid (JA), and abscisic acid (ABA) [9]. However, comprehensive analyses of CAMTA proteins from a variety of plant species at diverse phylogenetic locations are still lacking. To understand the origin, phylogeny, and structural evolution of plant *CAMTA* genes, we systemically identified the *CAMTA* gene family in three genome plant species.

Loss-of-function *CAMTA3/AtSR1* mutants revealed chlorosis and autonomous lesions, and elevated resistance to pathogens. Similarly, the mutant of a rice CAMTA member *OsCBT* gene showed significant resistance to pathogens, indicating that OsCBT might also act as a negative regulator on plant defense. CAMTA3 also played important roles in plant defense against insect herbivore, the regulation of glucose metabolism, and ethylene-induced senescence in *Arabidopsis*[10]. Recently, *CAMTA1, CAMTA2*, and *CAMTA3* were reported to function together in suppressing SA biosynthesis and were involved in cold/freezing tolerance by CBF transcription induction [11].

*CAMTAs* from tomato were identified to be differentially expressed genes (DEGs) during fruit development and ripening stages, indicating that calcium signaling is involved in the regulation of fruit development and ripening through calcium/calmodulin/CAMTA interactions [12]. Recently, 15 *CAMTA* genes were identified in soybean, and expression profile revealed that they were responsive to various stresses and hormone signals [13]. Based on a study, *TaCAMTA4* may function as a negative regulator in response to *Pucciniatriticina,* since the gene silencing-based knockdown of *TaCAMTA4* resulted in the enhanced resistance to P. triticina race 165 [14].

There is little analysis of CAMTA transcription factors for comparison of *Arabidopsis,* tomato and maize. Therefore, this study focuses on the systematic bioinformatics analysis and expression pattern analysis of CAMTA transcriptional factors in three species. In this study, we report the identification and a comprehensive analysis of the *CAMTA* gene family in *A.thaliana, Z.mays,* and *S.tuberosome*. Specifically, comprehensive information is provided on the gene structures, chromosomal locations, and phylogenetic relationships, gene duplications, genes expression during developmental stages, and promoter cis-elements identification of 17 *CAMTA* genes in three species.

## Material and Methods

The Arabidopsis information resource (TAIR) database was retrieved to download the sequences of six CAMTA family members from Arabidopsis thaliana. BLASTP (E-value was ≤ 1e-7) search was made against the genome of flax in NCBI database using Arabidopsis CAMTA proteins as queries. Pfam(https://pfam.xfam.org/), SMART (http://smart.embl-heidelberg.de/), and conserved domain database were used to collect, screen out, and filter non-redundant CAMTA sequences for conserved CAMTA domains using Gendoc software. The sequences were evaluated against the potential features of the CAMTA transporters such as the presence of IQ (PF00612), ANK (PF12796), TIG (PF01833), and CG-1 (PF03859) domains. The undesired gene sequences were eliminated manually.

ExPASy server (https://www.expasy.org/) was employed to predict the theoretical isoelectric point (pI) and the molecular weight (Mw) of each CAMTA protein.

### Phylogenetic Analysis and Gene Structure of CAMTA Proteins

CAMTA protein sequences were aligned in Arabidopsis, maize, and tomato by using the Clustal W function of MEGA 7.0 andphylogenetic tree was constructed using MEGA 7, applying the Neiojoubor joining algorithm with 1000 bootstraps replicates. The Gene Structure Display Server v2.0 (GSDS, http://gsds.cbi.pku.edu.cn/) was used to obtain information on the exon – intron of CAMTA proteins [15]. Both genome sequences and the coding sequences were utilized for, predictingthe positional information of the *CAMTA* genes using Ensemble plant, chromosomal locations. The sizes (bp) and intron numbers of *CAMTA* genes were identified.

### Analysis of Transcription Factor Binding Sites in the CAMTA Promoters

1500bp upstream of the promotor region of CAMTA was retrieved from NCBI database. The sequences obtained were analyzed using PlantPAN database (http://plantpan.itps.ncku.edu.tw/) with default limitations to identify the key *TFBS* with respect to stress response.

### Chromosomal Distribution and Conserved Motif of the CAMTA Gene

For locating the CAMTA members on Arabidopsis, maize and tomato chromosomes, *CAMTA* genes were placed on each chromosome according to the physical location of the gene. For exon/intron structural analysis, the genomic DNA and CDS sequences corresponding to each predicted *Arabidopsis*, tomato and maize *CAMTA* factor gene were downloaded from NCBI database. The *CAMTA* genes were drawn on all chromosomes and presented with MapChart [16]. The conserved motifs of 17 CAMTA protein sequences were analyzed by MEME program (http://meme.nbcr.net/meme/cgi-bin/meme.cgi). Conserved motifs were identified by online MEME (http:// meme.sdsc.edu/meme/meme.html) program using the full length protein sequences of each CAMTA protein sequences, with the following parameters: maximum number of motifs (6 motifs) and motif width set as 6–100 amino acids.

### Expression Patterns of CAMTAGenes in Developmental Stages

In order to find DEGs under developmental stages, *CAMTA* gene expression data was extracted by Genevestigator from *A.thaliana, Z.mays,* and *S.tuberosome* database using Affymatrix Arabidopsis/maize /tomato Genome Array platform and 'Perturbations' tool. DEGs with p-values <0.05 and log fold-change values ≥2 and≤ −2 were selected for genes. The fold changes in the expression of CAMTA genes under developmental stages were used to generate gene expression heatmap using genevestigator (https://genevesti-

gator.com/gv/) with purple/white color schemes as markers where "purple" and "white" colors represent up and down-regulation of the respective genes.

Identification of Orthologous and Paralogous CAMTAGenes between Three Species The orthologous genes in all three species were detected from EnsemblPlants (https://plants.ensembl.org/index.html). Orthologous genes in CAMTA proteins were selected when the identity exceeded 70% whereas, the paralogous genes were selected when the identity was more than 85% from EnsemblPlants. Orthologous and paralogous CAMTA genes were obtained using Circos program (http://mkweb.bcgsc.ca/tableviewer/).

## Results and Discussion

In the present study, we identified 17 members of CAMTA gene family in Arabidopsis, maize and tomato, and named them according to their position on the chromosome (Table 1). Bioinformatics analyses such as phylogenetic relationships, domains, gene structures, protein motifs, chromosomal locations, and detection of CAMTA homologous and orthologous genes were performed. The expression profiles in different stages were analyzed. All *SlCAMTA, ZmCAMTA,* and *AtCAMTA* genes are listed in table 1 along with their gene nomenclature, genes details, and amino acid length, molecular weight, and point isoelectric. The SlCAMTA TFs vary in amino acid length from 916 to 1041 with a predicted isoelectric point (pI) ranging from 5.50 to 8.89 and molecular weight ranging from 103.94 to 117.42 kDa. The AtCAMTA TFs vary in amino acid length from 845 to 1050 with a predicted isoelectric point (pI) ranging from 5.17 to 8.05 and molecular weight ranging from 96.18 to 117.26 kDa.The ZmCAMTA TFs vary in amino acid length from 842 to 1025 with a predicted isoelectric point (pI) ranging from 6.38 to 8.43 and molecular weight ranging from 94.65 to 114.42 kDa. Three genome sequences and the coding sequences were utilized for predicting the positional information of the *CAMTA* genes using Plazza database v12.1.

**Table 1:** Information for CAMTA Transcription Factor Family Members in Arabidopsis, Maize, and Tomato

| Ensemble plant | Other gene name | Duplication | Length | MW | PI: |
|---|---|---|---|---|---|
| Solyc01g057270.2.1 | SlCAMTA3 | - | 958 | 107.79 | 8.89 |
| Solyc01g105230.2.1 | SlCAMTA2 | - | 1041 | 117.42 | 5.50 |
| Solyc04g056270.2.1 | SlCAMTA3.1 | - | 1022 | 114.15 | 6.08 |
| Solyc05g015650.2.1 | SlCAMTA4 | - | 936 | 105.30 | 6.25 |
| Solyc12g035520.1.1 | SlCAMTA4.1 | - | 974 | 108.98 | 6.68 |
| Solyc12g099340.1.1 | SlCAMTA6 | - | 916 | 103.94 | 7.18 |
| AT1G67310.1 | AtCMTA4 | - | 1016 | 113.06 | 5.17 |
| AT2G22300.1 | AtCMTA3 | - | 1032 | 116.11 | 5.25 |
| AT3G16940.1 | AtCMTA6 | - | 845 | 96.18 | 8.05 |
| AT4G16150.1 | AtCMTA5 | - | 923 | 104.85 | 7.39 |
| AT5G09410.1 | AtCMTA1 | Block | 989 | 111.99 | 6.23 |
| AT5G64220.1 | AtCMTA2 | Block | 1050 | 117.26 | 6.39 |
| Zm00001d003958_T007 | ZmCAMTA1 | Block | 996 | 112.21 | 6.52 |
| Zm00001d021516_T001 | ZmCAMTA2 | - | 913 | 101.84 | 8.43 |
| Zm00001d025235_T001 | ZmCAMTA5 | Block | 865 | 97.52 | 7.70 |
| Zm00001d028007_T002 | ZmCAMTA4 | Block | 1025 | 114.42 | 6.38 |
| Zm00001d042313_T001 | ZmCAMTA3 | - | 842 | 94.65 | 6.93 |

## Phylogenetic Relationships, Conserved Motifs and Gene Structures of the CAMTA Factor Family Genes in Arabidopsis, Tomato and Maize

To determine the phylogenetic relationships among the different members of the CAMTA proteins in *Arabidopsis*, tomato, and maize, a phylogenetic analysis based on alignments of the 17 CAMTA protein sequences was performed. As shown in Figure 1

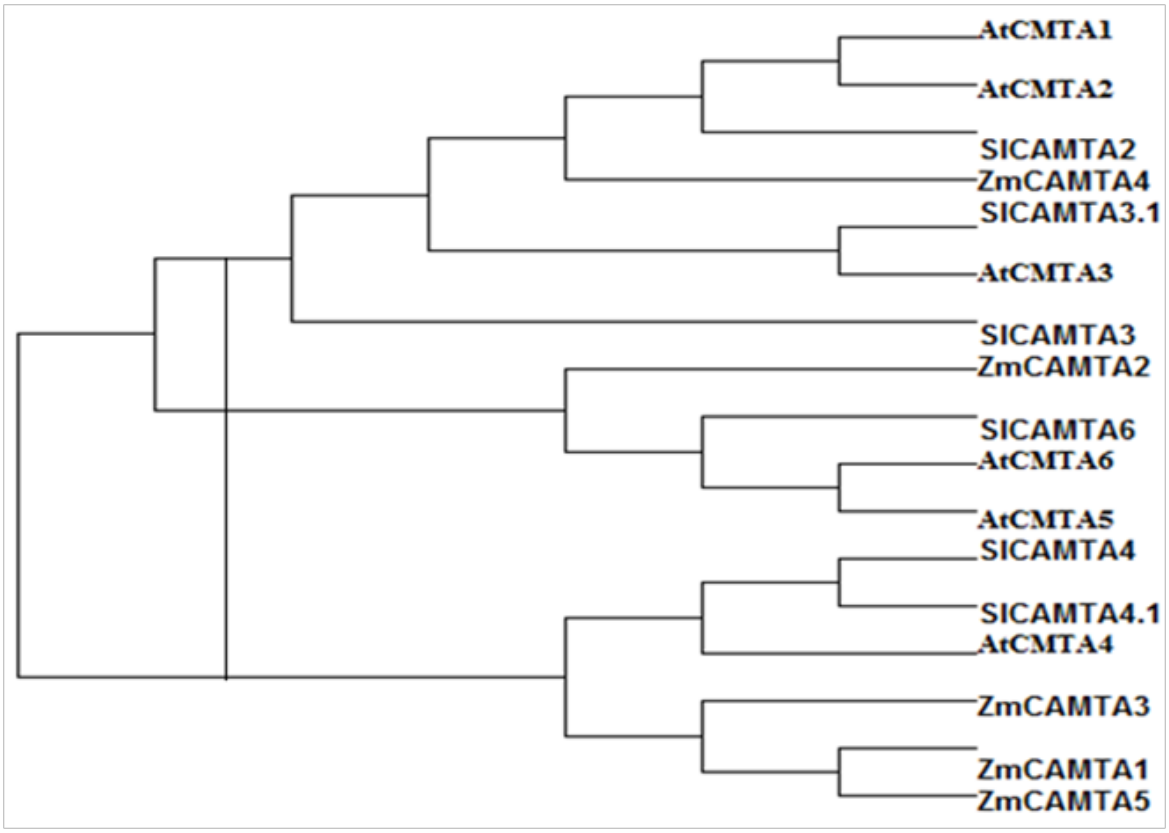



**Figure 1:** Phylogenetic Tree of CAMTA Genes Created By the Neighbor-Joining (NJ) Method in MEGA7.0 Software

The neighbor-joining phylogenetic tree divided 17 CAMTA genes into three clusters. In cluster I, seven CAMTA factors AtCAMTA1, AtCAMTA2, AtCAMTA3, SlCAMTA2, SlCAMTA3, SlCAMTA3.1and *ZmCAMTA4* gene were grouped. In cluster II, tomato CAMTA factor (SlCAMTA6) clustered with AtCAMTA5, AtCAMTA6, ZmCAMTA2 genes. In cluster III, two tomato CAMTA factors (SlCAMTA4, SlCAMTA4.1) clustered with four other *AtCAMTA4, ZmCAMTA1, ZmCAMTA3,* and ZmCAMTA5 genes. In contrast, *CAMTA* genes have been divided into seven subfamilies in *Arabidopsis* and rice [17]. It has been reported that the subfamilies I, II and III of the *CAMTA* genes simultaneously occur in both dicotyledons and monocots, but the subfamily IV genes did not exist in monocots [18]. Another study showed that CAMTA genes were divided into four (I, II, III, and IV) subfamilies in soybean [18]. The genes within each cluster showed similar exon/intron structures and conserved motif.To reveal the evolutionary relationship between Arabidopsis,tomato and maize we comprehensively analyzed the phylogeny between the orthologs of CAMTAs of three species.In this study, results of our clustering analysis of CAMTA factors in dicotyledons and monocots showed three subfamilies in the phylogenetic tree, similar results as obtained in soybean [18]. Most of the *AtCAMTA, SlCAMTA* and *ZmCAMTA* genes were classified in the same subfamily. The number of IQ motifs in all CAMTAs varied from one to three. All CAMTAs contain two IQ motifs in C-terminal (Figure 2a).

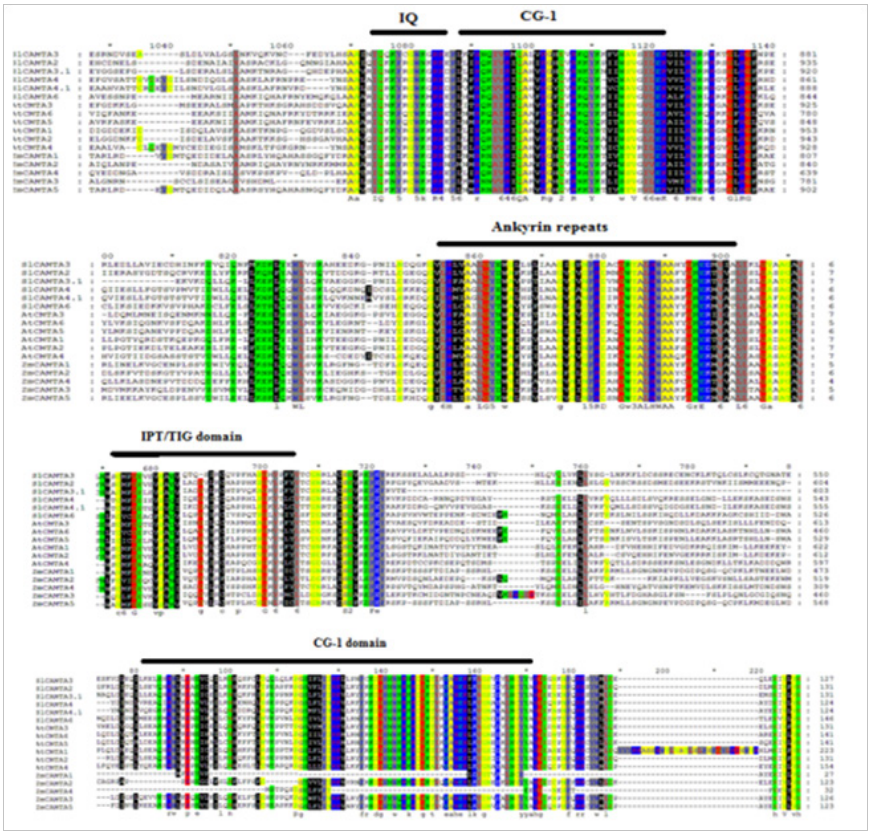



**Figure 2a:** Multiple Alignment of *Arabidopsis,* maize, and tomato CAMTA Family Proteins. Multiple Alignment Was Performed Via Clustal W and the Residues Were Colored Using Genedoc Software

This study revealed that CAMTAs share the same domain organization as reported previously [18].Using the MEME tool, a total of eight conserved motifs were detected in the *CAMTA* genes, and the lengths of these conserved motifs varied from 6 to 50 amino acids. Among them, motifs 2, 3, 4, 5, 7 and 8 were widely identified in all CAMTAs (Figure 2b). In general, the CAMTA factors in the same clade may have similar functions. The composition of conserved motifs in most proteins of each group was similar, which validated the CAMTA transcription factors phylogenetic classification.

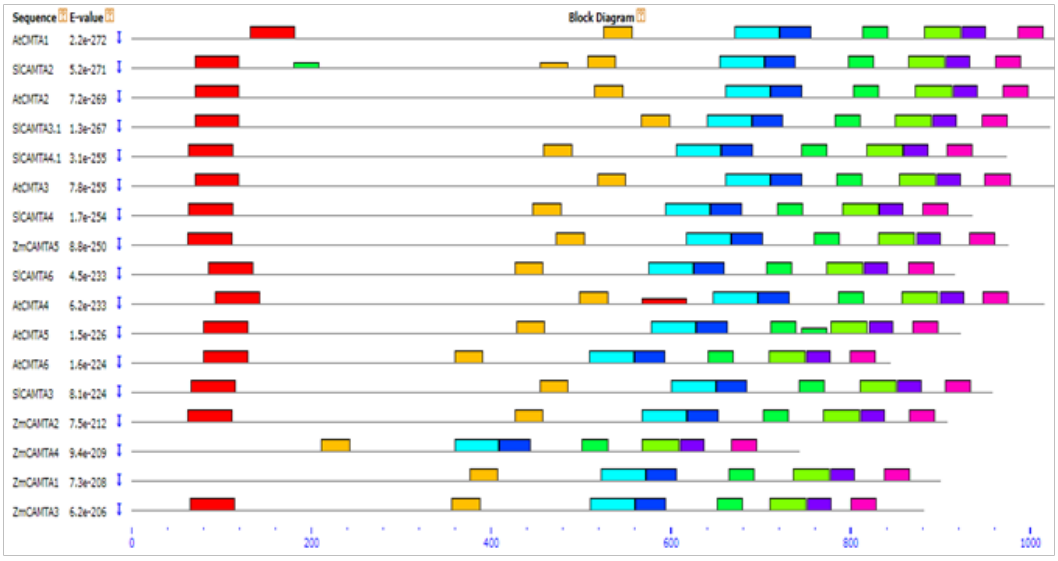



**Figure 2b:** Conserved CAMTA Protein Motifs In Maize, Recognized By MEME Database. Motifs 1-8 are Identified By Different Colors

## Gene Structure

*The highest numbers of introns i.e. 9–12, while group I and group III were disrupted by 1–12 introns disrupted CAMTA genes of group II.* The fixed number of introns and exons is a conserved character of CAMTAs among *Arabidopsis, maize and tomato [19].*

Study of protein structure is important to understand its mode of action. Using the MEME, the conserved motifs of all CAMTAs protein were analyzed. The plant CAMTA-encoded proteins were characterized with the presence of four functional domains, known as CG-1 (a sequence-specific DNA-binding domain), ANK (ankyrin repeats), IPT/TIG (transcription factor immunoglobulin), and IQ motifs (calmodulin-binding) which, are highly conserved among different plant species. All four CG-1, ANK, IPT/TIG, and IQ domains were detected within each of the *StCAMTA, ZmCAMTA*, *and AtCAMTA* CAMTA proteins,

## Chromosomal location and gene duplication

Analysis of the chromosomal distribution of *AtCAMTA* genes indicated that the 6 *CAMTA* genes of *Arabidopsis* were located on all chromosomes, while the 6SlCAMTA genes of tomato were located on chromosomes 1, 3, 4, and 5 but not on chromosomes 2 and 6, 7, 8, 9, 10, 11 (Figure 3a and 1b). In maize, *ZmCAMTA* genes were located on chromosomes 1,2,3,4, 7, and 10 but not on the chromosomes 5/6/8/9.

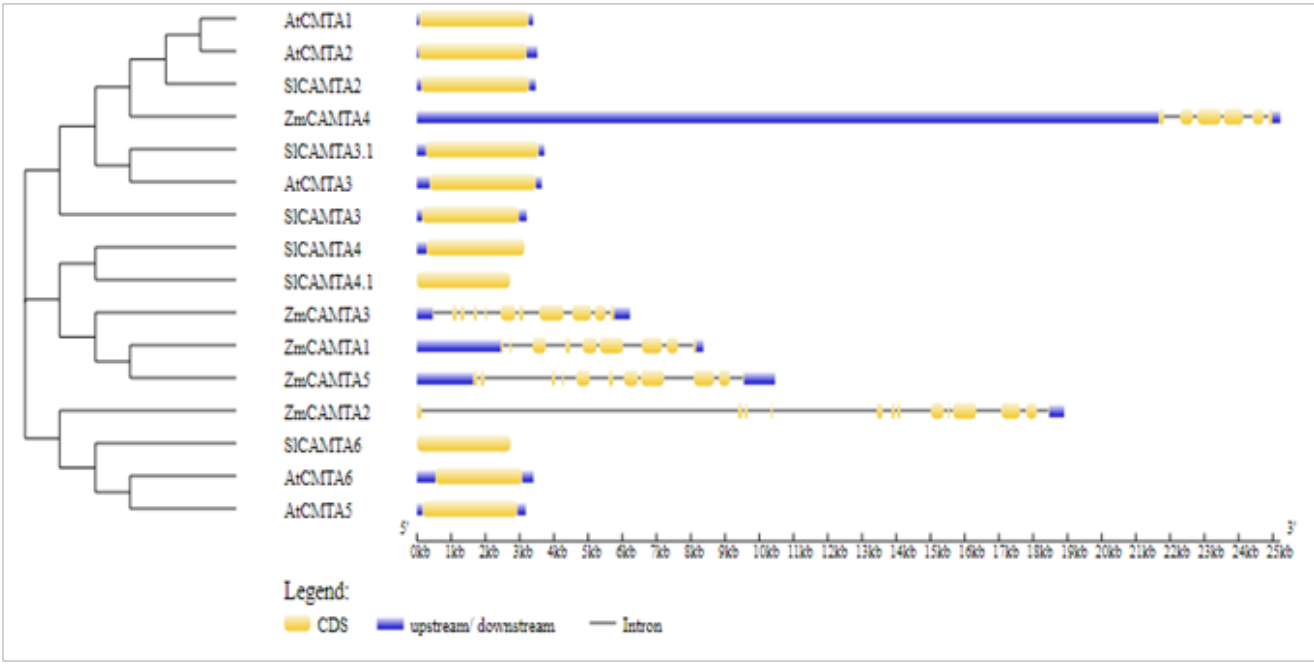
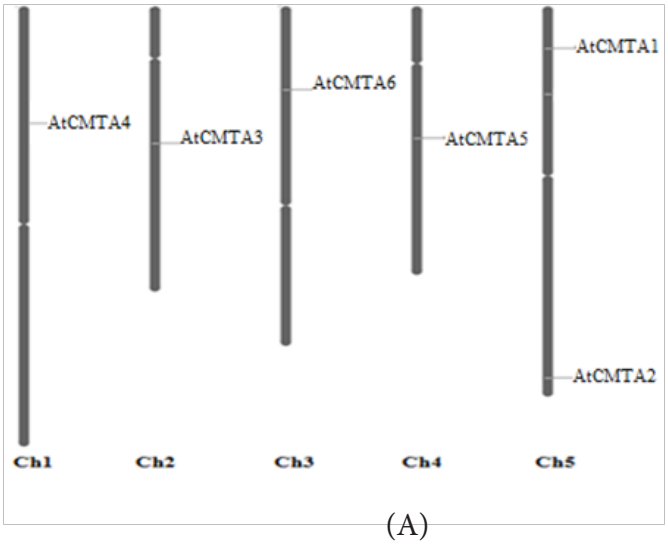
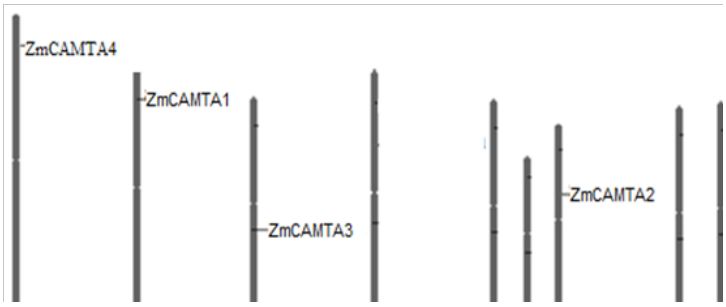



**Figure 3:** The Exon-Intronic Structure of *A. Thaliana, Z.Mays, and S.TuberosomeCamta* Genes According To Their Phylogenetic Relationships. Yellow And Blue Colors Represent Gene Exon and Intron, Respectively

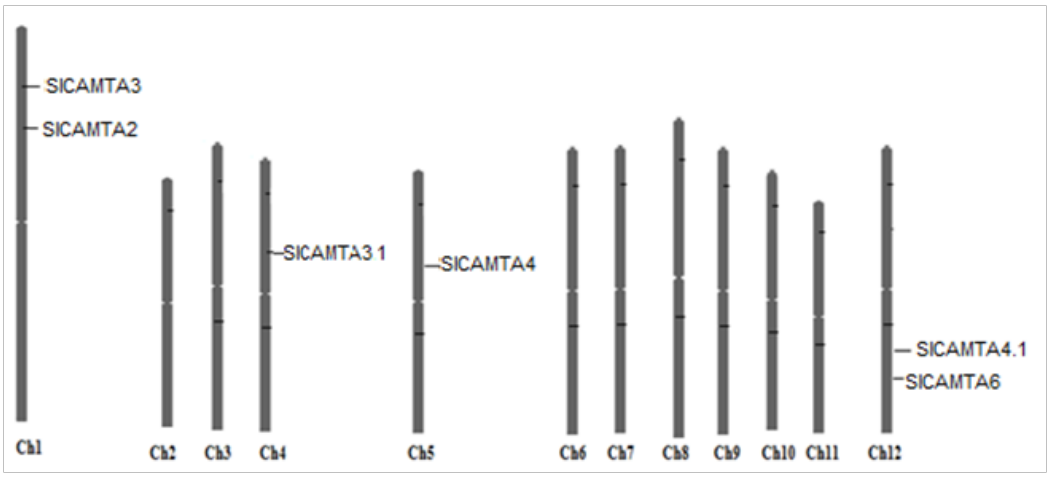
It can be concluded from Figure 4 that there is tandem duplication in Arabidopsis, tomato, and maize. Gene duplication can assist plants to adjust to different environments during their development and growth [20].




(A)




(B)




(C)

**Figure 4:**Chromosomal Distribution and Expansion Patterns of *CAMTA* Genes in *Arabidopsis(a)*, Maize (b), and Tomato (c).

## Prediction of TFBS in the CAMTA Genes

To understand the transcriptional, hormonal, and developmental regulation in response to stress, PlantPAN database was evaluated for the prediction of transcription factor binding sites in 1000 bp upstream promotor region of *CAMTA*. The results showed that there are different known stresses -related cis-elements that existed in the promoter regions of the 17 CAMTA genes.

The *CAMTAs* promoters have various cis-regulatory elements, believed to be involved in responses to environmental and hormonal regulations (Figure 5). In this study, the elements identified in response to various stresses included bZIP, CAMTA, MYB, bHLH, NF-YA NF-YB NF-YC and AP2/ERF.

Analysis of promoters showed that the bZIP and G-box cis-elements were expressed in response to ABA [20]. WRKY cis-elements are regulated in response to auxin hormone. Our results identified AT-hook, YABBY, and *EIL/EIN3* in the promoter regions of CAMTA. Among surveyed CAMTA, *ZmCAMTA* gene had the maximum cis elements in its promoter region. *ZmCMATA2, ZmCMATA4* and *ZmCMATA5* genes had the highest number of cis-elements in their promotor regions. The promoter sequences contained approximately 25 types of elements (Figure 5).

| | MYB;ARR-B | Myb/S | C2H2 | Homeodomain | bZIP | HD-ZIP | MYB- | Dof | GATA | MYB | EIN3;E | TCR | AT-H | bHLH | Trihelix | NAC | SBP | NF-YB;NF-YA | Dehydr | TCP | Sox;YABB | WRKY | LOB | MADS box | AP2;ERF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AtCMTA4 | 28 | 100 | 24 | 62 | 136 | 39 | 12 | 118 | 190 | 80 | 7 | 4 | 8 | 98 | 89 | 12 | 44 | 101 | 45 | 8 | 2 | 59 | 2 | 9 | 126 |
| AtCMTA3 | 25 | 91 | 17 | 86 | 168 | 46 | 22 | 135 | 197 | 95 | 4 | 11 | 24 | 125 | 110 | 11 | 73 | 129 | 42 | 8 | 0 | 101 | 0 | 4 | 151 |
| AtCMTA6 | 58 | 158 | 41 | 74 | 120 | 42 | 34 | 119 | 277 | 120 | 10 | 0 | 30 | 74 | 90 | 13 | 46 | 121 | 41 | 5 | 0 | 102 | 2 | 7 | 161 |
| AtCMTA5 | 44 | 105 | 27 | 41 | 144 | 13 | 18 | 88 | 208 | 94 | 1 | 1 | 24 | 99 | 88 | 22 | 43 | 90 | 57 | 6 | 0 | 69 | 1 | 16 | 120 |
| AtCMTA1 | 37 | 91 | 30 | 56 | 172 | 29 | 29 | 126 | 207 | 100 | 9 | 9 | 26 | 146 | 106 | 25 | 29 | 108 | 51 | 17 | 2 | 64 | 2 | 4 | 135 |
| AtCMTA2 | 32 | 80 | 33 | 81 | 178 | 39 | 20 | 102 | 218 | 105 | 7 | 13 | 21 | 103 | 86 | 12 | 55 | 133 | 62 | 15 | 2 | 77 | 1 | 10 | 169 |
| SlCMTA3 | 108 | 33 | 73 | 178 | 51 | 15 | 110 | 179 | 100 | 4 | 18 | 22 | 74 | 90 | 19 | 49 | 114 | 33 | 2 | 2 | 90 | 1 | 1 | 144 | 100 |
| SlCMTA2 | 41 | 120 | 23 | 93 | 149 | 52 | 18 | 105 | 213 | 134 | 7 | 17 | 23 | 87 | 91 | 19 | 21 | 147 | 38 | 8 | 2 | 48 | 2 | 5 | 156 |
| SlCMTA3.1 | 33 | 116 | 34 | 90 | 141 | 56 | 8 | 112 | 187 | 90 | 5 | 10 | 21 | 105 | 99 | 20 | 45 | 156 | 51 | 19 | 2 | 90 | 0 | 6 | 155 |
| SlCAMTA4 | 40 | 121 | 30 | 106 | 235 | 74 | 18 | 82 | 219 | 143 | 1 | 10 | 14 | 242 | 71 | 9 | 62 | 117 | 54 | 14 | 0 | 36 | 1 | 5 | 113 |
| SlCAMTA4.1 | 34 | 93 | 17 | 85 | 196 | 54 | 25 | 70 | 178 | 125 | 3 | 13 | 16 | 162 | 68 | 18 | 43 | 95 | 44 | 3 | 0 | 74 | 0 | 4 | 109 |
| SlCAMTA6 | 32 | 98 | 36 | 61 | 61 | 28 | 18 | 62 | 204 | 100 | 1 | 3 | 24 | 103 | 61 | 16 | 37 | 100 | 40 | 15 | 0 | 60 | 3 | 12 | 143 |
| ZmCAMTA1 | 70 | 208 | 70 | 160 | 121 | 73 | 29 | 235 | 406 | 264 | 27 | 17 | 118 | 31 | 6 | 26 | 115 | 50 | 316 | 42 | 3 | 151 | 3 | 16 | 356 |
| ZmCAMTA2 | 356 | 359 | 135 | 329 | 1096 | 192 | 72 | 447 | 899 | 495 | 48 | 33 | 180 | 716 | 465 | 77 | 330 | 621 | 464 | 147 | 6 | 287 | 4 | 39 | 1008 |
| ZmCAMTA5 | 98 | 308 | 106 | 222 | 455 | 112 | 55 | 313 | 588 | 388 | 29 | 42 | 175 | 246 | 315 | 46 | 151 | 382 | 136 | 52 | 4 | 293 | 2 | 26 | 584 |
| ZmCAMTA4 | 192 | 598 | 164 | 445 | 1100 | 239 | 95 | 571 | 1079 | 732 | 37 | 40 | 204 | 955 | 621 | 124 | 423 | 872 | 704 | 215 | 14 | 330 | 12 | 53 | 1460 |
| ZmCAMTA3 | 49 | 123 | 45 | 165 | 368 | 116 | 22 | 172 | 352 | 130 | 23 | 18 | 61 | 456 | 185 | 30 | 116 | 234 | 83 | 17 | 4 | 65 | 5 | 13 | 322 |

**Figure 5:** Prediction of TFBS in the Promoter Regions of CAMTA Genes in *Arabdopsis*, Tomato, and Maize

Previous studies have shown that *CAMTAs* were responsive to diverse stresses and stimuli in Arabidopsis, tomato, and soybean [11-21]. Enrichment of cis-elements involved in stresses/stimuli response in *CAMTA* promoters suggest that they are likely to respond differently to various stresses and stimuli signals, like other *CAMTAs* in different species.

Within the promoter regions of each of the three species, the abiotic responsive cis elements, AP2/ERF, CAMTA, WRKY and MYBwere surveyed. Pervious study has shown that WRKY cis-elements is often synergistically linked to the occurrence of responsive bZIP TFs [22]. Several AP2/ERFs play key roles in response to the induction of specific stresses and diverse DNA binding preferences and enable these TFs to integrate responses of multiple stimuli in stress condition [23].

Researchers have proposed that AP2/ERF TFs work in tandem with bZIPs and MYBs to bring about synergistic regulation of cold stress tolerance by controlling ABA mediated gene expression in *Arabidopsis* [24,25]. Therefore, it can be suggested that a network of TFs is involved in co-regulating diverse stress-responsive genes, which potentially form the missing molecular link between primary and specialized metabolism genes under stress conditions [26].

Among the cis elements surveyed, dehydrin was the most cis element found in the CAMTA SlCAMTA/AtCAMTA2/ ZmCAMTA2 genes. Due to the presence of dehydrin, plants are found to respond better to drought stresses. The enrichment of the dehydrin in the promoter regions of most *ZmCAMTA/AtCAMTA/SlCAMTA* family genes suggested comprehensive transcriptional regulation by

the CAMTAs themselves, and indicated a complicated regulation network between them. Our finding showed that thesurveyed *Sl-CAMTAs* genes, such as SlCAMTA7, SlCAMTA2, SlCAMTA17, and *SlCAMTA11*, were up-regulated under drought stress. Physiologically, the WRKY TFs binding to W-boxes regulate various developmental activities (controlling senescence) and defense associated processes (regulating responses to pathogen infestation and other abiotic stresses) [27]. Similarly, the MYB-binding sites are present in the promoters of the genes *(StCAMTA26a/26b* and SlCAMTA19/5/18) and MYB (StCAMTA26a/26b and SlCAMTA4/5 /8/8A/16a/19/26a/26b). It has been recognized that MYB TFs binding to their respective cis-elements control changes in various processes like hormonal signaling, specialized metabolism (phenylpropanoid and anthocyanin biosynthesis), cellular morphogenesis, and formation of meristem [28, 29].

Next, we further analyzed the hormone responsive elements in Arabidopsis, maize, and tomato *CAMTA* promoters. The most dominant hormone responsive element in *CAMTAs* is ABRE for recognizing ABA signal, detected in the promoters of 17 CAMTAs. Another founding member is LOB (corresponding toASL3), which is intriguing in that it is expressed specifically at the base region of all lateral organs formed from shoot apical meristem (SAM) and floral meristem [2].

The maximum number of WRKY elements was related to AtCAMTA3/6, SlCAMTA4 and ZmCAMTA5/6. Therefore, we can conclude that WRKY, in the promoter regions of genes, has important roles in response to pathogens. The ethylene-insensitive3-like/ethylene-insensitive3 *(EIL/EIN3)* protein family can serve as a vital factor for plant growth and development under different environmental conditions. *EIL/EIN3* protein is a form of a contained nuclear protein with DNA-binding activity, contributing to the complex network of primary and secondary metabolic pathways of plants [30].AT-hook DNA binding proteins are known to contribute to a functional nuclear architecture by binding to the nuclear matrix. AT-hook motifs bind to the minor grooves in duplex DNA of matrix attachment regions (MAR) of target DNA sequences, containing characteristic AT-rich DNA sequences. Many plant AT-hook motif proteins have a plant and prokaryote conserved (PPC) domain with unknown functions.At-hook motifs are related with functional domains in chromatin proteins and in DNA-binding proteins like that homedomains and zinc fingers.

Our finding showed that YABBY, WRKY, and GATA are involved in response to developmental stages and abiotic stresses. The NAC and YABBY transcription factors are known to be involved in numerous biological processes [31]. Another group of transcription factors is YABBY, playing a critical role in determining organ polarity. It is involved in the establishment of abaxial-adaxial polarity in lateral organs [31]. YABBY family transcription factors contain a zinc-finger domain in the amino-terminal region and a YABBY domain in the carboxyl-terminal region.

## Orthologous and Paralogous Genes Study in CAMTA Transcription Factor

In the present study, a comparative analysis was performed to identify orthologs of CAMTA genes between *Arabidopsis*, maize and tomato genomes (Figure 6). Based on the results, AtCAMTA5 with AtCAMTA6; AtCAMTA2 with SlCAMTA3.1 and SlCAMTA4 with SlCAMTA4.1 genes showed high similarity (identity 70%) and in total resulted in 3 orthologous gene pairs. The analysis of syntenywas shown that ZmCAMTA1 with ZmCAMTA3 were paralogous. Syntenic analysis suggest the segmental and tandem duplication as a major force for the diversity in *CAMTA* gene family. The syntenic analysis reveals the structural and functional conservation of the genes which underlie the origins of evolutionary novelty. Based on the evolutionary history of genes, orthologs have similar functions which reflected to genes with conserved domains (Altenhoff and Dessimoz, 2009).

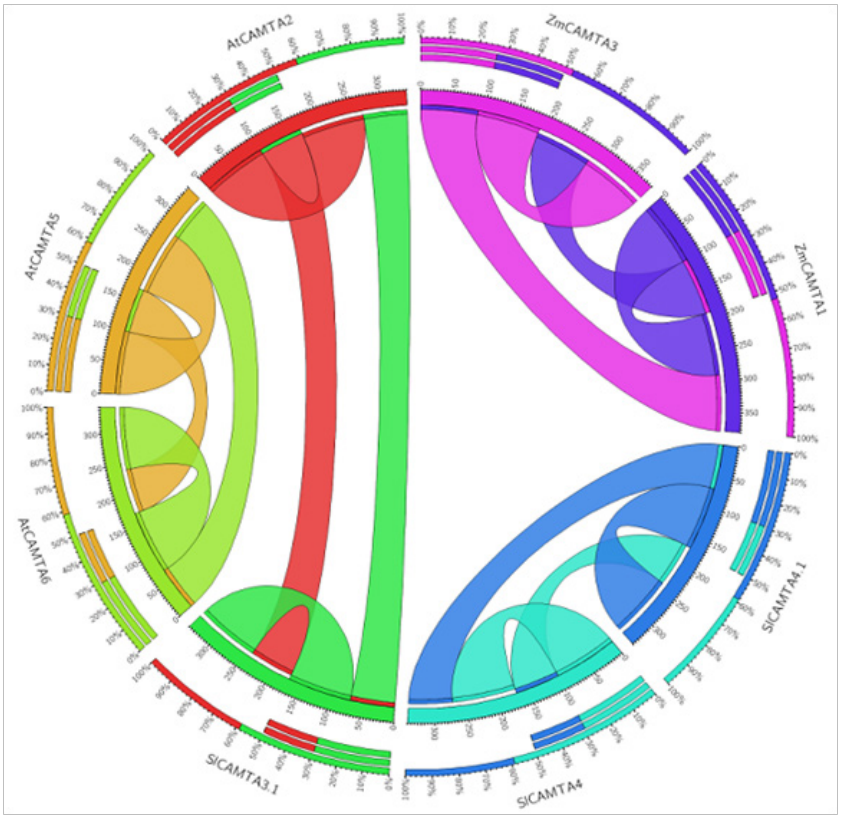



**Figure 6:**Orthologous and Paralogous Relationships of *CAMTA* Genes with *Arabidopsis*, Maize, and Tomato Genomes Visualized by Circos Database

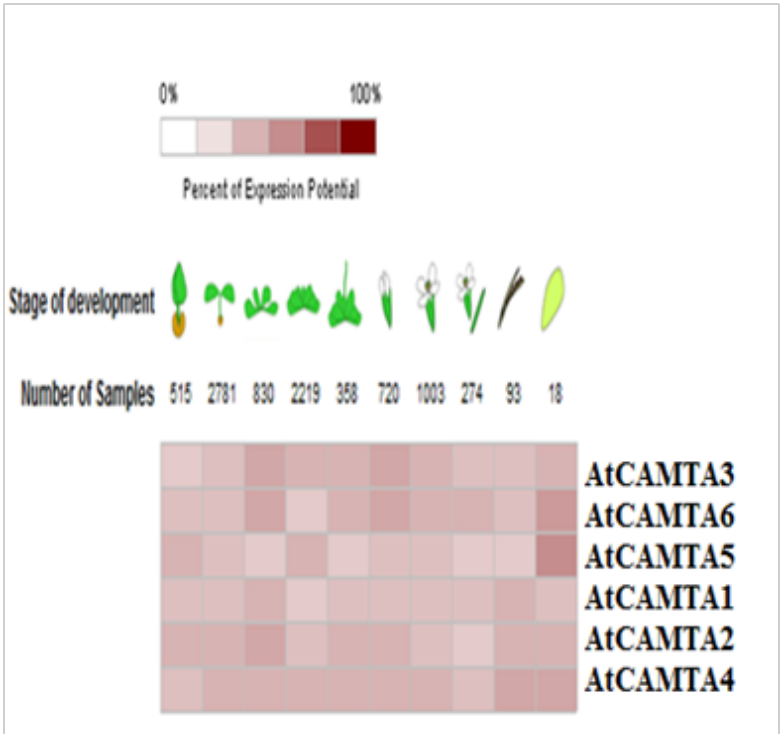
No duplicated *CAMTA* genes were identified from synteny analysis in monocot (maize) implying different expansion types of this gene family between maize and eudicots Arabidopsis and tomato, but *AtCAMTAwith SlCAMTA3.1* showed high similarity. Based on our findings, a comparative analysis was performed to identify orthologs of *CAMTA* genes among three genomes. Orthologous genes among Arabidopsis, maize, and tomato suggested that whole genome duplication (polyploidy) plays an important role in the expansion of *CAMTA* genes. In the pathway, gene duplication consisted of chromosomal duplications.

In addition, chromosomal/segmental duplication in *Arabidopsis,* maize and tomato genomes, was detected. Often orthologous genes have a similar function among different species. Therefore, the study of evolutionary genomics can shed light on the gene function. Analysis of *CAMTA* genes showed that whole genome duplication, tandem duplication, and chromosomal/segmental duplications play an important role in tomato genome expansion. However, the number of tandom/segmental duplications indicate that they are main factors in the evolution of *CAMTA* genes. Given the main role of these three species as model plants, their genomes provide a new resource for use in breeding. Also, three orthologous gene pairs were identified between *A.thaliana and Z.mays.*

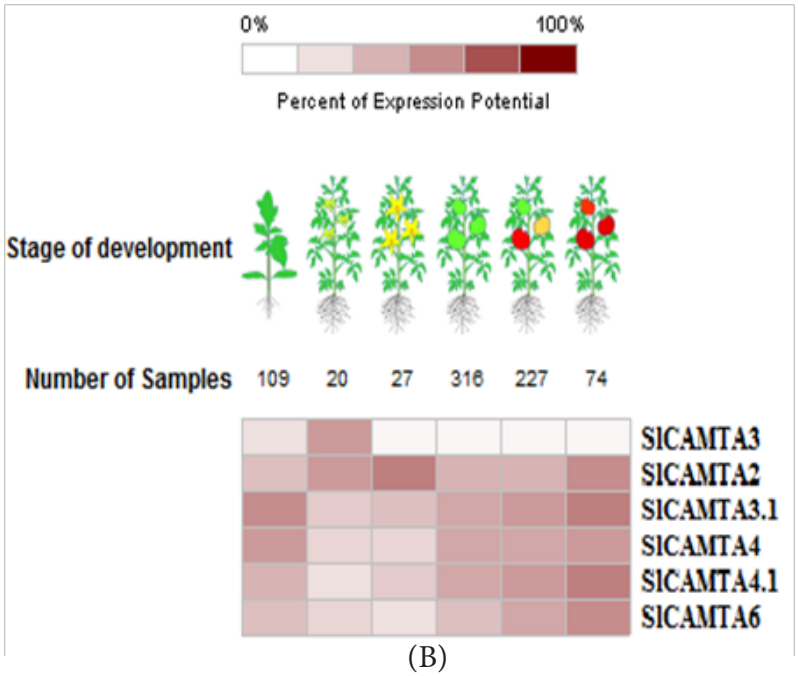
## Analysis of Expression Pattern of CAMTA Transcription Factors during Developmental Stages

Analysis of expression pattern of CAMTA in *Arabidopsis,* tomato and maize under developmental stage were performed. As shown in Figure 7a 7b and 7c in Arabidopsis, tomato and maize, respectively, we can clearly see the expression of *6AtCAMTA* transcription factors in developmental stages of *Arabidopsis* that AtCAMTA5 and AtCAMTA6 were up-regulated in almost of all developmental stages such as senescence and young rosette (Figure 7a). In tomato, As shown in Figure7b, most of SlCAMTA3 genes down-regulated gene, which up-regulated SlCAMTA2, SlCAMTA3, SlCAMTA3.1, SlCAMTA4, SlCAMTA4.1, and SlCAMTA6 in developmental stages. CAMTA TFs are regulated at many processes of plant growth and development, especially during light-mediated processes such as flowering, maturation, embryo development, and differentiation and expansion [32]. In maize, *ZmCAMTA1 and ZmCAMTA2* genes were up regulated in all developmental stages. It has been reported that some *ZmCAMTA* genes were expressed in different tissues. The CAMTA genes highly expressed in organs of plants are crucial for the functioning or development of a specific organ.*StCAMTA11* and *StCAMTA28* genes up regulated in root, shoot, and inflorescence. It has been reported that fewer differentially expressed CAMTA factor genes were found in soybean roots than in soybean leaves [34]. CAMTAs have been shown to be ex-

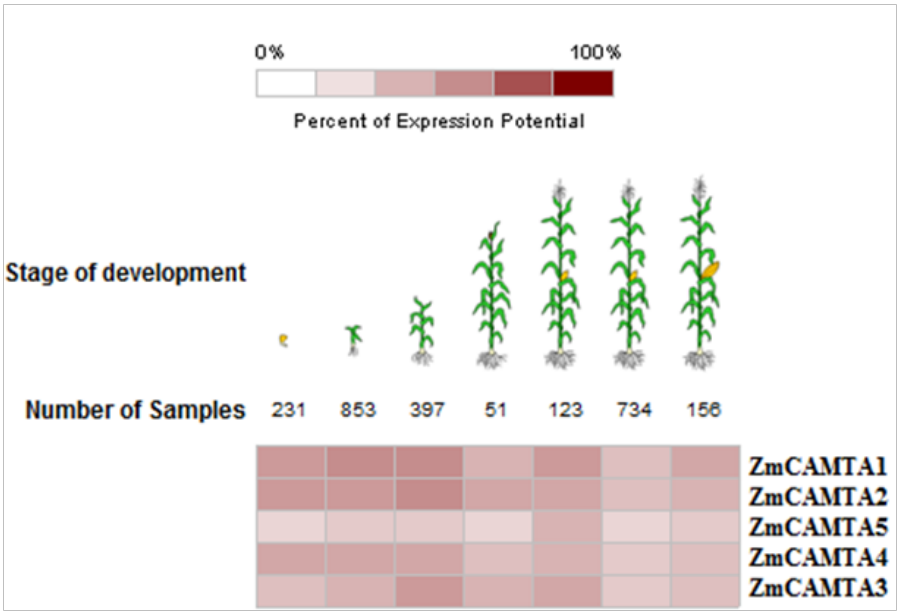
tensively involved in plant growth and developmental regulation, as well as in biotic and abiotic stress tolerance [33]. In *Arabidopsis, CAMTA1 and CAMTA2* genes use in concert with CAMTA3 to directly bind to the promoter of C-repeat binding factor2 *(CBF2)* to induce expression, leading to increased plant freezing tolerance. AtCAMTA3 can act as a negative regulator of plant immunity to modulate pathogen defense responses by activating the EDS1-mediated salicylic acid (SA) signaling. A recent study showed that *TaCAMTA4* may function as a negative regulator of the defense response against Pucciniatriticina, since the virus-induced gene silencing (VIGS)-based knockdown of *TaCAMTA4* resulted in the enhanced resistance to *P. triticina [18].* Our results suggest that one CAMTA member usually participates in multiple signaling pathways, while multiple CAMTA members often work together to participate in one signaling pathway.




(A)




(B)




(C)

**Figure 7:** Heat map Representation of Expression Analysis of **CAMTA** Genes at Different Developmental Stages of *Arabidopsis (A),* Tomato (B), Maize (C

## Conclusions

In conclusion, 17 *CAMTA* genes were identified in Arabidopsis, tomato, maize in the present study. Analysis of the gene structure- and protein domain, biochemical properties, and the phylogenetic tree indicated that the *CAMTA* gene family was highly conserved during plant evolution. Expression analysis showed that all 17 *CAMTA* genes were expressed in multiple tissues with different expression levels, suggesting that various *CAMTA* gene members maintain different functions in growth and development. Using prediction of cis elements, the present of AP2/ERF in all the *CAMTA* genes could respond to at least one abiotic stress or multiple stresses, implying different regulations and functions of *CAMTA* gene members for coping with various abiotic stresses. *CAMTA* genes in *Arabidopsis,* tomato, maize genome was predicted to be potential target genes by CAMTA, demonstrating that CAMTA can be widely involved in plant development and growth, as well as coping with stresses. Our findings provide new insight into the *CAMTA* gene family in three species as well as a foundation for further studies on the roles of *CAMTA* genes in wheat development, growth and stress response [35].This study will help to identify novel *CAMTA* genes for future breeding to improve plant production, quality and stress resistance, and open a new way for further elucidation for their roles underlying the signal transduction in among three species.

## References


1. Heil, M., &Bostock, R. M. (2002). Induced systemic resistance (ISR) against pathogens in the context of induced plant defences. Annals of botany, 89(5), 503-512.
2. Singh, B., Bohra, A., Mishra, S., Joshi, R., & Pandey, S. (2015). Embracing new-generation 'omics' tools to improve drought tolerance in cereal and food-legume crops. Biologiaplantarum, 59(3), 413-428.
3. Sahoo, K. K., Tripathi, A. K., Pareek, A., &Singla-Pareek, S. L. (2013). Taming drought stress in rice through genetic engineering of transcription factors and protein kinases. Plant Stress, 7(1), 60-72.
4. Reddy, A. S., Ali, G. S., Celesnik, H., & Day, I. S. (2011). Coping with stresses: roles of calcium-and calcium/calmodulin-regulated gene expression. The Plant Cell, 23(6), 2010-2032.
5. Rubtsov, A. M., &Lopina, O. D. (2000). Ankyrins. FEBS letters, 482(1-2), 1-5.
6. Ali, E., Raza, M. A., Cai, M., Hussain, N., Shahzad, A. N., Hussain, M., & Sun, P. (2020). Calmodulin-binding transcription activator (CAMTA) genes family: Genome-wide survey and phylogenetic analysis in flax (Linumusitatissimum). Plos one, 15(7), e0236454.
7. Du, L., Yang, T., Puthanveettil, S. V., &Poovaiah, B. W. (2011). Decoding of calcium signal through calmodulin: calmodulin-binding proteins in plants. In Coding and decoding of calcium signals in plants (pp. 177-233). Springer, Berlin, Heidelberg.
8. Bouché, N., Scharlat, A., Snedden, W., Bouchez, D., & Fromm, H. (2002). A novel family of calmodulin-binding transcription activators in multicellular organisms. Journal of Biological Chemistry, 277(24), 21851-21861.
9. Liu, J., Whalley, H. J., & Knight, M. R. (2015). Combining modelling and experimental approaches to explain how cal-

cium signatures are decoded by calmodulin binding transcription activators (CAMTA s) to produce specific gene expression responses. New Phytologist, 208(1), 174-187.
10. Yang, T., &Poovaiah, B. W. (2002). A calmodulin-binding/CGCG box DNA-binding protein family involved in multiple signaling pathways in plants. Journal of Biological Chemistry, 277(47), 45049-45058.
11. Galon, Y., Aloni, R., Nachmias, D., Snir, O., Feldmesser, E., Scrase-Field, S., & Fromm, H. (2010). Calmodulin-binding transcription activator 1 mediates auxin signaling and responds to stresses in Arabidopsis. Planta, 232(1), 165-178.
12. Kim, Y. S., an, C., Park, S., Gilmour, S. J., Wang, L., Renna, L., &Thomashow, M. F. (2017). CAMTA-mediated regulation of salicylic acid immunity pathway genes in Arabidopsis exposed to low temperature and pathogen infection. The Plant Cell, 29(10), 2465-2477.
13. Saidi, A., &Hajibarat, Z. (2020). In-silico analysis of eukaryotic translation initiation factors (eIFs) in response to environmental stresses in rice (Oryza sativa). Biologia, 75(10), 1731-1738.
14. Shkolnik, D., Finkler, A., Pasmanik-Chor, M., & Fromm, H. (2019). CALMODULIN-BINDING TRANSCRIPTION ACTIVATOR 6: A key regulator of Na+ homeostasis during germination. Plant physiology, 180(2), 1101-1118.
15. Chen, Z. J., Scheffler, B. E., Dennis, E., Triplett, B. A., Zhang, T., Guo, W., ...& Paterson, A. H. (2007). Toward sequencing cotton (Gossypium) genomes. Plant physiology, 145(4), 1303-1310.
16. Voorrips, R. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. Journal of heredity, 93(1), 77-78.
17. Bork, P., Doerks, T., Springer, T. A., &Snel, B. (1999). Domains in plexins: links to integrins and transcription factors. Trends in biochemical sciences, 24(7), 261-263.
18. Rahman, H., Xu, Y. P., Zhang, X. R., &Cai, X. Z. (2016). Brassica napus genome possesses extraordinary high number of CAMTA genes and CAMTA3 contributes to PAMP triggered immunity and resistance to Sclerotiniasclerotiorum. Frontiers in plant science, 7, 581.
19. Wang G. Zeng H.X.Hu Y.Zhu Y.Chen Shen C.H.Wang,BW Poovaiah, and Du L. (2015). Identification and expression analyses of calmodulin-binding transcription activator genes in soybean. Plant and soil, 1386(1-2):205-21.
20. Saidi, A., Hajibarat, Z., &Hajibarat, Z. (2020). Transcriptome analysis of Phytophthorainfestans and Colletotrichumcoccodes in tomato to reveal resistance mechanisms. Asia-Pacific Journal Molecular Biology and Biotechnology.
21. Yang, T., Peng, H., Whitaker, B. D., & Conway, W. S. (2012). Characterization of a calcium/calmodulin-regulated SR. CAMTA gene.
22. Llorca, C. M., Potschin, M., &Zentgraf, U. (2014). bZIPs and WRKYs: two large transcription factor families executing two different functional strategies. Frontiers in plant science, 5, 169.
23. Xie, Z., Nolan, T. M., Jiang, H., & Yin, Y. (2019). AP2/ERF transcription factor regulatory networks in hormone and abiotic stress responses in Arabidopsis. Frontiers in plant science, 10, 228.
24. Pandey, G. K., Grant, J. J., Cheong, Y. H., Kim, B. G., Li, L., & Luan, S. (2005). ABR1, an APETALA2-domain transcription factor that functions as a repressor of ABA response in Arabidopsis. Plant Physiology, 139(3), 1185-1193.
25. Xu, Z. S., Chen, M., Li, L. C., & Ma, Y. Z. (2011). Functions and application of the AP2/ERF transcription factor family in crop improvement F. Journal of integrative plant biology, 53(7), 570-585.
26. Maruyama, K., Todaka, D., Mizoi, J., Yoshida, T., Kidokoro, S., Matsukura, S., & Yamaguchi-Shinozaki, K. (2012). Identification of cis-acting promoter elements in cold-and dehydration-induced transcriptional pathways in Arabidopsis, rice, and soybean. DNA research, 19(1), 37-49.
27. Van Aken, O., Zhang, B., Law, S., Narsai, R., & Whelan, J. (2013). AtWRKY40 and AtWRKY63 modulate the expression of stress-responsive nuclear genes encoding mitochondrial and chloroplast proteins. Plant physiology, 162(1), 254-271.
28. Cao, Z. H., Zhang, S. Z., Wang, R. K., Zhang, R. F., &Hao, Y. J. (2013). Genome wide analysis of the apple MYB transcription factor family allows the identification of MdoMYB121 gene confering abiotic stress tolerance in plants. PLoS One, 8(7), e69955.
29. Höll, J., Vannozzi, A., Czemmel, S., D'Onofrio, C., Walker, A. R., Rausch, T., & Bogs, J. (2013). The R2R3-MYB transcription factors MYB14 and MYB15 regulate stilbene biosynthesis in Vitisvinifera. The Plant Cell, 25(10), 4135-4149.
30. SAIDI, A., HAJIBARAT, Z., & HAJIBARAT, Z. (2020). Identification of responsivegenes and analysis of genes with bacterial-inducible cis-regulatory elements in the promoter regions in Oryza sativa L. ActaagriculturaeSlovenica, 116(1), 115-123.
31. Altenhoff, A. M., &Dessimoz, C. (2009). Phylogenetic and functional assessmentoforthologs inference projects and methods. PLoS computational biology, 5(1), e1000262.
32. Yoon, H. K., Kim, S. G., Kim, S. Y., & Park, C. M. (2008). Regulation of leaf senescence by NTL9-mediated osmotic stress signaling in Arabidopsis. Molecules & Cells (Springer Science & Business Media BV), 25(3).
33. Liu, H. L., Wang, G. C., Feng, Z., & Zhu, J. (2010). Screening of genes associated with dedifferentiation and effect of LBD29 on pericycle cells in Arabidopsis thaliana. Plant growth regulation, 62(2), 127-136.
34. Yang, F., Dong, F. S., Hu, F. H., Liu, Y. W., Chai, J. F., Zhao, H., & Zhou, S. (2020). Genome-wide identification and expression analysis of the calmodulin-binding transcription activator (CAMTA) gene family in wheat (Triticumaestivum L.). BMC genetics, 21(1), 1-10.
35. Mishra, A. K., Choi, J., Rabbee, M. F., &Baek, K. H. (2019). In silico genome-wide analysis of the ATP-binding cassette transporter gene family in soybean (Glycine max L.) and their expression profiling. BioMed research international, 2019.