International Journal of Environment, Agriculture and Biotechnology
Vol-9, Issue-4; Jul-Aug, 2024
Peer-Reviewed International Journal
Journal Home Page Available: https://ijeab.com/
Journal DOI: 10.22161/ijeab

Genome-wide identification, Characterization, and Expression analysis of the *Caffeic Acid O-Methyl Transferase* (*COMT*) Gene Family of *Sorghum Bicolor*

Prashanth Bollempally[1], Vinod Kumar Anumandla[1], Anjana Priyadarshani Kanathala[2], Srinivas Naik Kethavath[2] and Prashant Singam[1]*

[1]Department of Genetics, Osmania University, Hyderabad -500 007, India
[2]Centre for Plant Molecular Biology, Osmania University, Hyderabad -500 007, India
*Corresponding author: Dr. Singam Prashant
Email: prashantsingam@gmail.com
Department of Genetics, Osmania University, Hyderabad, India.

Received: 15 Jul 2024; Received in revised form: 16 Aug 2024; Accepted: 22 Aug 2024; Available online: 26 Aug 2024

Abstract— *Caffeic acid O-methyltransferases (COMTs) are essential enzymes for producing natural products in plants, specifically involved in the phenylalanine metabolic pathway and the monolignol biosynthetic pathway. These enzymes are responsible for the methylation of caffeic acid compounds, which are the building blocks for many plant-derived compounds with various biological activities. The investigation of the evolutionary divergence, expression patterns under diverse abiotic stress conditions, and lignin content-related features of the COMT gene family in Sorghum has not been explored. In this study, forty-eight SbCOMTs were identified in S.bicolor. Based on the examination of evolutionary relationships, 48 SbCOMTs were classified into two distinct categories. The gene characterization and the conserved motif patterns in each group were similar, demonstrating the reliability of the phylogenetic categorization. Chromosomes 5 and 7 have been found as the hotspot of SbCOMTs with 10 and 7 genes respectively. Phylogenetic analysis revealed the conservation of Sorghum COMT genes among Zea mays and Oryza sativa. Investigation of regulatory elements specifies the significant roles that COMT genes play in the monolignol biosynthetic pathway of S. bicolor. Analysis of miRNA, transcription factor binding, and gene expression analysis provides insights to further engineer lignin biosynthetic pathway for better biofuel yield. We found that two SbCOMTs (SbCOMT26& 36) were highly expressed and their relative contents were similar to the variation drift of lignin content under abiotic stress conditions in S. bicolor. These results provide a clue for further study on the roles of SbCOMTs in the development of Sorghum and could favourably be foundations for the cultivation of Sorghum with higher biomass and yield with enhanced abiotic stress tolerance.*

Keywords— ***Caffeic acid O-methyltransferase, Monolignol biosynthesis, Abiotic stress, Biomass, Biofuels.***

Highlights

- Discovery of *COMT* gene family members in Sorghum helps in identification of genes responsible for developmental lignification and their involvement in other metabolic process.
- Cis regulatory analysis, transcription factor prediction and miRNA analysis *SbCOMTs* provide insights into manipulation of these genes for development of crops for better biofuel yield.
- The genomic location and tissue specific expression analysis of *Caffeic acid O-*

methyltransferase (*COMT*) genes under drought and salt stress reveal their critical role in lignification in *Sorghum bicolor.*

I. INTRODUCTION

O-methyltransferases (OMTs) catalyze a wide range of reactions in lignin and flavonoid biosynthesis pathways. COMTs are responsible for lignin biosynthesis and are involved in phenyl-alanine metabolism in plants. According to previous reports on monolignol production, the key methylations of Lignin precursors are primarily facilitated by specific S-adenosyl-L-methionine (SAM)-dependent enzymes, including caffeoyl CoA 3-O-methyltransferase (CCoAOMT; EC 2.1.1.104) and caffeic acid O-methyltransferase (COMT; EC 2.1.1.68) (Louie et al. 2010). The COMTs are grouped in plant type 1 of the SAM-dependent O-methyltransferases family (Noel et al. 2003). These enzymes utilize S-adenosyl-methionine as a methyl group donor and perform methylation of the 5-hydroxyl group of their substrate, 5-hydroxy coniferaldehyde, ultimately leading to the production of S-lignin units.In *Arabidopsis thaliana*, COMT may convert 5-OH coniferaldehyde/5-OH coniferyl alcohol into sinapaldehyde/sinapyl alcohol and caffeic acid into ferulic acid, which results in the synthesis of both G and S units of lignin (Goujon et al., 2003). Previous studies on *Arabidopsis thaliana* by Lee et al. (2015) suggest that COMT is also essential for the conversion of N-acetyl serotonin to melatonin. The COMTs of *sorghum* can methylate flavones such as luteolin and selgin in *sorghum* to aid the synthesis of tricin (Eudes et al. 2017).

Sorghum (*Sorghum bicolor*) is one of the primary staple grains consumed in India, following rice (*Oryza sativa*) and wheat (*Triticum aestivum*), and holds the 5th position in global cereal production. In addition, it is a promising crop for biofuel and a possible source of cellulosic feedstock. The estimated size of its diploid genome is 730 MB, and it has a haploid chromosome number of 10. Plant-based renewable biofuels promise sustainable solutions to food and energy demands. *Sorghum* offers the status of a highly diverse food, feed, and biofuel source globally. *Sorghum* is a useful crop for almost all renewable energy systems that are being developed for green technology and renewable fuels.

Lignin is a polyphenolic polymer enclosed by wood fibres, other tube bundle cells, and thick-walled cell walls. The three major monolignols, p-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol, yield p-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) subunits, respectively. Upon polymerization, these three subunits will form rigid and complex lignin in plants. The composition of these subunits will regulate the physical properties and digestibility of lignin (Baucher et al., 2003). Bugos et al. (1991) reported the first exploration of the COMT gene family in *Populus tremuloides*. Later, the COMT gene family is uncovered in several species, which include seven COMTs in *Eucalyptus grandis* (Carocha et al., 2015), *Catalpa bungei* comprises 23 COMTs (Lu et al., 2019), 92 COMT members found in *blueberries* (Liu et al., 2021), *Populus trichocarpa* (Chiang et al., 2010), *Brassica rapa* L. (Wei et al., 2016), and *Betula pendula* (Chen et al., 2020) harbours 25 COMT candidates and *Soybean* contains 55 COMTs (Zhang et al., 2021). In plants, COMT regulates responses to a variety of stresses, including drought (Yao et al., 2022), salt (Chang et al.,2021), cold (Zhang et al., 2021), and phytohormone signaling.

In the present study, identification of COMT homologs, gene structure, gene characterizations, chromosomal locations, evolutionary relationships, conserved motifs analysis, cellular localization, promoter analysis, protein modeling, protein-protein interactions, miRNA prediction, transcription factor prediction, and expression patterns was mined in *S. bicolor*. These findings would help in the manipulation of the lignin biosynthetic pathway for better biofuel yield and breeding *Sorghum* cultivars with enhanced abiotic and biotic stress tolerance.

II. MATERIALS AND METHODS

Plant material and induction of stress

Seeds of the *Sorghum bicolor*high biomass variety (IS 4698) were obtained from the Indian Institute of Millet Research (IIMR), Rajendranagar, Hyderabad, and sown in pots filled with 4 kg of black soil at the Departmental Farm, Department of Genetics, Osmania University, Hyderabad (India). Seedlings were raised in a glass house environment at 28–20^{0}C day/night temperatures. Sixty-five-day-old seedlings were treated with 200 mM NaCl solution and 200 mM Mannitol solution for 48 hours each. Under comparable conditions, corresponding controls were kept well-watered and without any treatment. Various plant tissues, such as leaves, stems, and roots, were collected from both the treated and control groups. These tissues were then snap-frozen in liquid nitrogen and preserved at -80°C for future use. Three technical and three biological replicates were employed for the qRT-PCR study.

In silico prediction, identification, and characterization of *SbCOMT* genes

For the identification of the *COMT* gene family in *Sorghum*, the protein sequences of the *Sorghum bicolor* genome were retrieved from the Phytozome (http://www.phytozome.net/) plant database to use as the local protein database. Previously characterized

Arabidopsis COMT genes were used to perform a BLASTP search against the local protein database with a threshold of E-value < 1e-5. The PFAM profile was used as the query to search against the local protein database using HMMER 3.0 with a threshold of E-value < 1e-5. Based on the results of HMMER and BLASTP, the redundant sequences were removed. Then, the putative *Sorghum* COMT genes were retrieved from PFAM databases (http://pfam.xfam.org/search) to predict the conserved protein domain, and those containing a complete COMT domain remained as candidates.

Gene structure prediction, conserved motif analysis, sub-cellular localization, and protein parameters

The prediction of gene structure was carried out using GTF annotation files using TB tools. For the prediction of subcellular localization of the proteins, WoLFPSORT was used. Parameters like isoelectric point (pI), molecular weight (MW), GRAVY (grand average of hydropathy), instability, and aliphatic indexes in ProtParam software were employed. For conserved motifs, MEME with parameters like 10 numbers of motifs, 2–20 motif sites, and 6–20 wide motif widths were used. The genes on the chromosomes were mapped based on their physical location using an online phenogram tool.

Phylogenetic analysis, multiple sequence alignment, generation of Synteny maps, and Ka/Ks analysis

A phylogenetic tree was developed using the MEGAv10 program, employing the Neighbour-Joining algorithm (Kumar et al. 2018) with 1000 bootstrap samples based on the amino acid sequences of *Sorghum bicolor* (Sb), *Oryza sativa*(Os), *Zea mays* (Zm), and *Arabidopsis thaliana* (At). SbCOMT 26 and 36 proteins were aligned with orthologs in the above-mentioned species using MEGAv10.0. All the predicted SbCOMT homologs were mapped on *Z. mays* and *O. sativa* genomes, and synteny maps were generated with TB tools (Chen et al. 2020). The synonymous to non-synonymous ratios and time of evolution (MYA) of the SbCOMT paralog pair were calculated by an online Ka/Ks calculator.

Prediction of cis-elements, protein modelling, and protein-protein interactions

Promoter elements were identified for all the SbCOMT genes from the Phytozome database, and the 1500-bp sequence upstream for all the *Sorghum* COMT homologs was extracted and submitted to the Plant CARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) to predict cis-elements. The 3D structures of all the SbCOMT proteins were predicted using the SWISS-MODEL server (https://swissmodel.expasy.org/) (Biasini. et al. 2014). The predicted 3D structures of proteins were evaluated for stability using the Protein Structure Verification Server (PSVS) (https://saves.mbi.ucla.edu/) and Ramachandran plots. The predicted protein-protein interaction (PPI) map of *Sorghum* COMT homologs was generated from the STRING database (https://string-db.org/).

miRNA and Transcription factor analysis

We predicted miRNAs that might target SbCOMTs to control their expression using the Plant psRNA Target tool (https://www.zhaolab.org/psRNATarget/) with default parameters, and all of the *Sorghum* miRNAs were used. The regulatory network of the SbCOMT gene and miRNA was visualized using Cytoscape (https://cytoscape.org/). Transcription factor binding sites of all SbCOMT homologs were predicted by the Plant transcription factor database (PTFDB) (http://planttfdb.gao-lab.org/) and a network built using Cytoscape.

***In silico* expression analysis of *SbCOMT* genes**

The transcriptome data (FPKM) of Sorghum was downloaded from the Gramene database (https://www.gramene.org/). The transcriptome data include baseline expression of SbCOMTs in various organs of *Sorghum* (Davidson et al. 2012), vascular and non-vascular tissue (Turco et al. 2017), stem internodes of bioenergy *Sorghum* (Kebrom et al. 2017), and expression patterns in leaf and root tissue under drought conditions (Varoquaux et al. 2019). The expression patterns were visualized by a heat map built with TB tools. Transcriptome (FPKM) data of *S. bicolor* under osmotic stress and ABA stress (Acc: SRP007361) (Dugas et al. 2011) mined from the Morokoshi *Sorghum* transcriptome database (http://sorghum.riken.jp).

Expression analysis of *SbCOMT* genes by qRT PCR

Total RNA was extracted from stress-exposed and control (without stress) plants using the Trizol reagent method. The purity of the RNA was determined using an Eppendorf Bio photometer. One microgram of RNA was used as a template for first-strand cDNA synthesis with the PrimeScriptTM RT Reagent Kit (Takara, Japan) according to the manufacturer's instructions. 2X SYBR Premix Ex Tag (Tli RNaseH Plus, Takara, Japan) Master Mix with gene-specific primers (Table 5) was used to determine the relative gene expression levels of SbCOMTs.

Thermal cycling conditions of 95^{0}C for 2 min, followed by 40 cycles of 95 ^{0}C for 30 s, 58 ^{0}C for 30 s, and 72 ^{0}C for 30 s, were programmed in the ABI 7500 real-time PCR system (Applied Biosystems, Foster City, CA) for qRT-PCR analysis. SbCOMT gene expression in both treated and control samples was normalized using the EIF4a (Eukaryotic Initiation Factor 4A) reference gene. For each sample, qRT-PCR was performed using three biological

and three technical replicates. The relative amounts (fold change) of each transcript were calculated using the comparative 2^ΔΔCT method.

III. RESULTS

Identification and Characterization of SbCOMT homologs in *Sorghum bicolor*

A total of 48 potential sequences were obtained from the *Sorghum* genome. Then, all 48 candidate sequences were scanned for a methyltransf_2 domain. Forty-eight sequences with a methyltransf_2 domain (Fig. 1.d) were identified in the Sorghum diploid genome. All of them were mapped on pseudochromosomes and renamed from SbCOMT1 to SbCOMT48 (Table 1). The protein parameters (Table 2) and gene structural characteristics were analyzed (Fig. 1.b). The result showed that SbCOMT43 was the shortest protein (100 amino acids), and the longest one was SbCOMT9.

The molecular weight of 48 SbCOMT proteins ranged from 39 to 47 kDa, and the isoelectric point ranged from 4.65 to 7.13. The investigation of the conserved domain and gene structure indicated that all *COMT* genes possessed a catalytic domain at the C-terminus, which was referred to as the Methyltransf_2 domain, encompassing a binding pocket for SAM/SAH and the AdoMet-MTase superfamily domain. Some of them showed a common structure with an N-terminal domain called dimerization. The binding pocket for SAM/SAH exhibited significant conservation, whereas the binding sites for substrates were distinct for proteins belonging to diverse groups. SbCOMT 32, 34, 37, and 38 contain 3:4 of introns: exons; SbCOMT9, 26; and COMT47 consist of two introns and three exons; and SbCOMT 30, 40, and 41 have only one exon without introns. The patterns of the methyltransf_2 domain in SbCOMTs were similar in the same group.

The structural differences in protein sequences across the *Sorghum* COMTs were assessed using the Multiple Expectation Maximisation for Motif Elicitation (MEME) online tools (Fig. 1.c). A total of 10 motifs were found in the sorghum COMT proteins. Most of the motifs were the same in the two groups, and they were in the same order in COMT proteins within the same group. The consensus motif 1 (Methyl transferase-2) and motif 2 (AdoMet-MTase) are found in all SbCOMTs.

Most of the SbCOMT gene homologs are localized in the cytoplasm, followed by the chloroplast, plasma membrane, and mitochondria. Only SbCOMT39 is present in mitochondria; SbCOMT9, SbCOMT10, SbCOMT15, SbCOMT16, SbCOMT23, SbCOMT40, SbCOMT41, SbCOMT44, and SbCOMT45 are localized on the plasma membrane; SbCOMT5, SbCOMT24, SbCOMT26, SbCOMT27, SbCOMT32, SbCOMT34, SbCOMT35, SbCOMT37, SbCOMT38, SbCOMT43, SbCOMT46, SbCOMT47, and SbCOMT48 are localized in the cytoplasm; and the rest of the SbCOMT homologs are found in the cytoplasm (Table 1).

Phylogenetic analysis, multiple sequence alignment, chromosomal location, synteny, and Ka/Ks analysis of SbCOMTs

The phylogenetic tree analysis revealed the evolutionary connection of SbCOMT homologs in *Sorghum bicolor* with *Oryza sativa*, *Zea mays*, and *Arabidopsis thaliana* (Fig. 2.a). A total of 15 SbCOMT paralogs were identified. A neighbourhood joining (NJ) phylogenetic tree created with *Sorghum* COMT protein sequences showed that the proteins were distributed into two groups.

Sorghum showed 12 ortholog pairs, 11 with *Zea mays* (SbCOMT9 and ZmCOMT18, SbCOMT10 and ZmCOMT12, SbCOMT11, and ZmCOMT20, SbCOMT16 and ZmCOMT29, SbCOMT17 and ZmCOMT5 and SbCOMT20 and ZmCOMT6, SbCONT26, and ZmCOMT22, SbCOMT30 and ZmCOMT31, SbCOMT31 and ZmCOMT4, SbCOMT36 and ZmCOMT28, SbCOMT43 and ZmCOMT17) and 1 with *Oryza* (SbCOMT5 and OsCOMT16). The location of SbCOMT homologs was mapped on the *Sorghum* genome (Fig. 2.b). SbCOMT genes are scattered on all 10 chromosomes of *Sorghum*. Synonymous to non-synonymous substitution rates (Ka/Ks) of 9 *Sorghum* paralogs (SbCOMT1 and SbCOMT2, SbCOMT3 and SbCOMT4, SbCOMT8, and SbCOMT12, SbCOMT14 and SbCOMT20, SbCOMT18 and SbCOMT21, SbCOMT19 and SbCOMT22, SbCOMT23 and SbCOMT24, SbCOMT28 and SbCOMT30, SbCOMT32 and SbCOMT34) were calculated (Table 3).

Multiple sequence alignments of SbCOMT26 and 36 with orthologs in other species displayed highly conserved residues, which indicates these genes are conserved among species. All *Sorghum* paralogs showed substitution rates <1. The lowest Ka/Ks (0.05735056) was observed in gene pairs SbCOMT18 and SbCOMT21, and the highest Ka/Ks were observed in gene pairs SbCOMT32 and SbCOMT34, respectively. The selection pressures on the COMTs in *S. bicolor* were explored based on the Ka/Ks ratios. This investigation revealed that the Ka/Ks ratios of SbCOMT paralogs are <1, which indicates that SbCOMTs experienced purifying selection during evolution.

Collinearity analysis of SbCOMT homologs has been performed on the genomes of *Zea mays* and *Oryza sativa* (Figs. 3a and b). *S. bicolor* chromosomes 1, 5, and 9 display 3, 3, and 2 homologs each with the *Zea mays genome*, respectively. *S. bicolor* chromosome 1 shows 2 homologs with the *Zea mays* 9 chromosome and 1

homolog with the *Zea mays* 10 chromosome; *S. bicolor* chromosome 4 displays 1 homolog on the *Zea mays* 4 chromosome; *S. bicolor* chromosome 5 displays 3 homologs on the *Zea mays genome*, 2 homologs on the 4th, and 1 homolog on the 2nd chromosome; and *S. bicolor* chromosome 9 shows two homologs with *Zea mays* chromosomes 6 and 8, respectively. Pink-coloured links represent homologs between two genomes. The *S. bicolor* genome shows a total of 6 homologs in the *O. sativa* genome. *S. bicolor* chromosome 4 shows one homolog with *O. sativa* chromosome 2, and *S. bicolor* chromosome 3 displays 1 homolog with *O. sativa* chromosomes 1 and 5, and *S. bicolor* chromosome 5 shows 1 homolog with *O. sativa* chromosome 1. Sb chromosome 7 displays 1 homolog with O. sativa chromosome 8 and Sb chromosome 9 shows 1 homolog with *O. sativa* chromosome 5.

Cis-regulatory elements analysis of SbCOMTs

The initiation of transcription is a pivotal phase in gene expression, representing a critical juncture where RNA polymerase interacts with regulatory sequences like the promoter, which ultimately impacts the gene expression level. (Liu et al. 2019). Promoter analysis of SbCOMT homologs revealed the occurrence of lignin biosynthesis, abiotic stress, light-responsive, and phytohormone-responsive putative cis-regulatory elements (Fig. 4). Different elements like defense-responsive, wound-responsive, MYB-drought-responsive, MYB-light-responsive, and MYB-flavonoid genes-related cis-regulatory elements are found in the promoter regions of SbCOMT genes. MYB and NAC represented the highest number of elements in all the SbCOMT homologs, indicating their involvement in lignin biosynthesis and stress tolerance. SbCOMT homologs contain defense-responsive elements, indicating their involvement in biotic stress-related defense. Most COMT homologs have phytohormone-responsive elements like ABRE, MeJARE, GARE, SARE, and AURE. MeJARE and SARE, the defense-responsive elements, have been found to have the highest number of elements among the phytohormone-responsive elements and have been identified in all the SbCOMT homologs, indicating their involvement in defense mechanisms. Light-responsive elements are also found in the promoter regions of SbCOMT homologs. This finding indicates that COMT genes in *S. bicolor* may be regulated by light. Similar cis-regulatory elements within homologs may significantly influence similarities among gene expression patterns and gene roles. A large majority of SbCOMTs had ABRE, related to abscisic acid, and MeJRE, related to methyl jasmonate.

3D structures and PPI analysis of SbCOMTs

3D structures of SbCOMT proteins were predicted with the best PDB templates (Fig. 5.a). The template PDB ID, chain, model of the oligomer, and their structure validations are represented in Table. 3D structures of SbCOMT36 displayed 100% identity with the Caffeic acid-O-Methyltransferase of *S. bicolor* (PDB ID-4pgg.1. A) protein, and SbCOMT38 showed 100% identity with the Stilbene-O-Methyltransferase protein (PDB ID-7vb8.1. A). The rest of the SbCOMT homologs displayed identity, with corresponding templates ranging from 31% to 66%. In the predicted PPI map, one of the putatively expressed and characterized SbCOMT (36) proteins exhibited interactions with several proteins (Fig. 5.b). SbCOMT36 protein shows 11 nodes with 38 edges with other proteins. Each protein showed more than one interactant. The proteins that display interactions with SbCOMT (Sb07g003860.1) are CAD (Sb04g005950.1) terminal gene in Lignin biosynthesis; phenylalanine ammonia-lyase (Sb04g026510.1, Sb04g026520.1, Sb01g014020.1, Sb06g022750.1, Sb06g022740.1) is involved in the L-phenylalanine catabolic process, phenylpropanoid biosynthetic process, and phenylpropanoid metabolic process. Probable 4-coumarate-coA ligase 1 (Sb07g007810.1) is involved in the early stages of lignin biosynthesis; F5H (Sb02g002630.1) is involved in the conversion of coniferaldehyde to sinapaldehyde in lignin biosynthesis; and Folylpolyglutamate synthase (Sb01g049840.1) is involved in purine, pyrimidine, and amino acid synthesis.

miRNA and Transcription factor binding site prediction

Additionally, we predicted the miRNAs that might target SbCOMTs to regulate their expression. In total, 19 SbCOMT genes were found to be targeted by 31 miRNAs, and miRNA-SbCOMT interactions were constructed (Fig. 6.a). Combined with the miRNA-SbCOMT relationship and co-regulation modules of SbCOMTs, which provide some insights into the regulation of SbCOMTs expression to control lignin biosynthesis, in silico analysis of SbCOMTs revealed the presence of numerous cis-elements that may assist as binding sites for transcription factors with vital functions in lignin biosynthesis. To further determine this, Plant TFDB (Jin et al. 2014) was used to attain comparative models of transcription factors binding on regulatory regions of SbCOMTs. The model displays interactions with various transcription factors such as Dof, LFY, BES1, MYB-related, E2F, HSF, TCP, ARF, ERF, MICK-MADS, SBP, NAC, MYB, and LBD (Fig. 6.b). Plants comprise an MYB sub-family protein that is characterized by the R2R3-type MYB domain, which plays the role of master regulatory switch in secondary cell wall biosynthesis (McCarthy et al. 2009; Zhong et al. 2012; Kim et al. 2019). They might also directly activate some

lignin genes through the secondary wall MYB-responsive element (SMRE) binding site (consensus motif ACC(A/T) A(A/C) (T/C)) in the promoter region (Zhong et al. 2012). MYB transcription factors function specifically in the regulation of lignin biosynthesis (Stracke et al. 2001).

The NAC family of transcription factors is composed of a vast array of proteins. NAC transcription factors were found to contribute to plant responses to pathogens, viral infections, and environmental stimuli such as drought and salinity conditions (Xie et al., 1999; Ren et al., 2000; Collinge et al., 2001; Kim et al., 2007). Certain NAC transcription factors have been identified as playing a crucial role in controlling cell aging, proliferation, and the development of wood. (Takada et al., 2001; Vroemen et al., 2003; Weir et al., 2004; Zhong et al., 2006; Kim et al., 2007; Yamaguchi et al., 2008).

***In silico* expression analysis of SbCOMTs**

The transcriptome data (FPKM) of SbCOMT genes were analyzed to determine the expression patterns of these genes under natural habitat and in drought, osmotic, and Abscisic acid stress conditions in various tissues and organs of *S. bicolor*. We predicted the expression patterns of SbCOMT homologs in different regions of the stem internodes of bioenergy sorghum (Fig. 7.a). Among all SbCOMTs, SbCOMT36 displayed the highest level of expression patterns in all regions of the stem internodes of bioenergy *sorghum*. SbCOMT11 and 7 showed the highest expression patterns in internode regions 2 and 3. The investigation of expression patterns of SbCOMTs in vascular and non-vascular tissues of s*orghum* (Fig. 7.b) revealed that SbCOMT6, 23, 25, 26, and 27 exhibited the highest expression patterns and the rest of the genes expressed in the medium to very low range. In the baseline expression analysis (Fig. 7.c), at the embryonic stage, SbCOMT26 showed a medium expression level, SbCOMT34 and 30 displayed a low level of expression, and the rest of the genes did not show significant expression patterns. At the flowering stage, SbCOMT36 displayed the highest level of expression, and SbCOMT33 and 17 showed a medium range of expression. In the floral meristem, only SbCOMT26 showed a medium level of expression; none of the SbCOMT homologs displayed significant expression patterns. In meristematic tissue, only SbCOMT26 and 36 showed a medium-range expression pattern. In the shoot, SbCOMT 17 and 36 are highly expressed, and SbCOMT 19, 11, and 26 exhibit an average range of expression. In root tissue, SbCOMT36 and 20 exhibited the highest expression patterns, followed by SbCOMT17 and 3, which displayed the second-highest expression patterns.

In the floral meristem, only SbCOMT26 showed a medium level of expression; none of the SbCOMT homologs displayed significant expression patterns. In meristematic tissue, only SbCOMT26 and 36 showed a medium-range expression pattern. In the shoot, SbCOMT 17 and 36 are highly expressed, and SbCOMT 19, 11, and 26 exhibit an average range of expression. In root tissue, SbCOMT36 and 20 exhibited the highest expression patterns, followed by SbCOMT17 and 3, which displayed the second-highest expression patterns. We analyzed the expression patterns of SbCOMTs in leaf and root tissues under drought conditions at different growth intervals (Fig. 7.d). At 42 days of growth in leaf tissue under drought stress, expression levels of SbCOMT 17, 25, 26, and 30 are highly up-regulated. The expression of SbCOMT 12, 29, and 33 is up-regulated in leaf tissue after 63 days of growth. At 77 days of growth, SbCOMT6, 25, and 26 were up-regulated, and only a few SbCOMT homologs were down-regulated in leaf tissue under drought conditions.

None of the *SbCOMT* genes exhibited significant expression patterns in the rest of the growth stages of *sorghum* in leaf tissue during drought stress responses. When compared with root tissue under drought conditions, SbCOMT1, 2, 20, 21, 37, 38, and 39 displayed a low range of expression and also exhibited constant expression in all stages of growth. The expression of SbCOMT25 is up-regulated in root tissue at 35 days and 77 days under drought, and the expression of SbCOMT6 and 36 displayed the highest level of expression at 77 days of growth under the drought stress response.

Additionally, we explored the expression patterns of SbCOMT homologs in *S. bicolor* shoot and root tissue under osmotic (NaOH), ABA, and PEG stress conditions (Fig. 7.e). SbCOMT-36 was highly up-regulated in root tissue under osmotic stress conditions and displayed moderate expression in root tissue under ABA and PEG stress conditions also, SbCOMT-36 shows moderate expression patterns in shoot tissue, almost in the above-mentioned stress conditions. SbCOMT-20 and SbCOMT-17 displayed medium expression patterns in root tissue under osmotic stress conditions and shoot tissue under PEG treatment, respectively.

qRT PCR analysis of SbCOMT genes

Based on in silico transcriptome analysis, SbCOMT26 and 36 were considered for qRT-PCR analysis in different tissues and organs of *Sorghum bicolor* under control, drought, and salinity stress conditions due to their expression in major lignifying organs of *sorghum*. SbCOMT homologs showed variable gene expression across different tissues (Fig. 8). SbCOMT26 expression was significantly higher during drought stress compared to

salt stress. Among the stress treatments, drought-stressed leaves and salt-stressed stem tissues showed a 12.47-fold and 11.75-fold rise in transcript levels of SbCOMT26, respectively. However, there was no significant increase in SbCOMT26 expression under drought and salt stress conditions in other *sorghum* tissues. Whereas, SbCOMT36 expression was found to be higher in salt stress than in drought stress conditions. SbCOMT36 expression was significantly increased in salt-stressed stems. Under salt-stress conditions, there was a 32-fold rise in SbCOMT36 transcript levels in shoots and a 3-fold rise in drought-stressed roots.

IV. DISCUSSION

Caffeic acid O-methyltransferases (COMTs) are essential enzymes that contribute significantly to the synthesis of lignin and the phenylalanine metabolic pathway in plants. Frequently, attempts are made to manipulate the lignin makeup of genetically modified crops to enhance their digestibility as forage, effectiveness in pulping, and the production of biofuels. L-phenylalanine serves as the preliminary material for the synthesis of monolignols. According to the current understanding of monolignol production, the essential O-methylations of hydroxyl groups on the phenolic ring of monolignol precursors are primarily facilitated by specific S-adenosyl-L-methionine (SAM)-dependent enzymes, including caffeoyl CoA 3-O-methyltransferase (CCoAOMT; EC 2.1.1.104) and caffeic acid O-methyltransferase (COMT; EC 2.1.1.68). (Louie et al. 2010). The COMTs are classified in the plant type-1 family of SAM-dependent O-methyltransferases (Noel et al. 2003). *Sorghum caffeic acid O-methyltransferase* uses S-adenosyl-methionine as a donor of methyl groups and performs methylation of the 5-hydroxyl group of its favored substrate, 5-hydroxyconiferaldehyde, ultimately leading to the production of S-lignin units. O-methyltransferases (OMTs) are responsible for a variety of versatile reactions in the biosynthesis pathways of lignin and flavonoids.

Since COMTs may respond to a variety of substrates, including phenylpropanoids, flavonoids, and alkaloids, they are likely to respond to a variety of stimuli. As a result, they are ubiquitous in plants due to their significance in helping plants adapt to their environment and challenging circumstances (Nomura et al., 2010). The publication of diverse plant genomes has allowed analyses of COMT family genes in several species to be carried out (Barakat et al., 2011; Wu et al., 2013; Liu et al., 2016) and is majorly involved in lignin biosynthesis as lignin provides mechanical strength to plants. *Sorghum bicolor* has been extensively studied for its large amounts of flavonoids, primarily in food crops, forage, and biofuel crops. The *Sorghum* v3 genome was released in 2017, and 48 COMTs have been identified, named SbCOMT1-SbCOMT48. Subsequently, many homologs have been detected in plants. We need to understand which of the homologs performs the crucial processes of plant growth, flavonoid metabolism, phenylalanine metabolism, stress tolerance, and lignin biosynthesis. Four plants, including *Sorghum bicolor*, *Zea mays*, *Oryza sativa*, and *Arabidopsis thaliana*, were examined in this study, and each was shown to have a distinct number of COMTs. COMTs in all these plants comprise the conserved methyltransferase-2 and dimerization domains. Furthermore, we found that the number of COMTs in *S. bicolor* is greater than that in *O. sativa*, *Zea mays*, and *Arabidopsis thaliana*. *Oryza sativa* contains the second-highest number of COMTs (39 COMT homologs), followed by *Zea mays* (32 COMT homologs). *Arabidopsis thaliana* comprises the least number of COMTs (17 COMT homologs) of all the studied species. The conserved domains of identified SbCOMT homologs, i.e., methyltransferase-2 and dimerization domains, correlate with those of other plants (Liu et al., 2021).

Lignin is the key component of vascular tissue and provides plants with structural support to stand upright. COMTs are important enzymes involved in lignin biosynthesis that catalyze the methylation of S-lignin monomers. Evolutionary analysis suggests that these 48 SbCOMTs are grouped into two clades denoted as Group Ia, Group Ib, and Group II. SbCOMT homologs were more closely related to *O. sativa* COMTs than *Arabidopsis thaliana*. All the identified SbCOMT proteins comprise the conserved Methyl transferase-2 (PF00891) domain, which has 207 amino acid residues, including a SAM/SAH binding pocket and a substrate-binding site, and the Dimerization domain (PF08100), which contains 52 amino acid residues. All the discovered SbCOMT homologs displayed conserved AdoMet-MTase superfamily domains.

In the present research, all the identified *Sorghum* homologs comprised conserved domains such as methyl transferase-2 and dimerization domains, and they were involved in numerous functions. About 20–30% of SbCOMT homologs belong to the Iso flavone-O-Methyltransferase family. These methyltransferases were involved in secondary metabolite biosynthesis and iso-flavonoid biosynthesis (BRENDA: EC2.1.1.46). Some of them SbCOMTs belong to the ZPR3 and ZPR4 families, which encode O-methyltransferase and might be complicated in suberin biosynthesis (Held et al. 1993). SbCOMT5, belonging to Trans-resveratrol di-O-methyltransferase and Resveratrol O-methyltransferase, plays vital roles in biotic (Sambangi et al. 2016) and abiotic stress responses (Chiron et al. 2000). SbCOMT 34,

37, and 38 functionally belonged to the iso-eugenol O-methyltransferase family and regulated the biosynthesis of secondary metabolites and phenylpropanoid biosynthesis (BRENDA: EC2.1.1.146).

Collinearity analysis of SbCOMTs with other species, including *O. sativa* and *Z. mays*, revealed varied collinearity with each species. One of the SbCOMT genes shows two collinearity blocks on the *Z. mays* genome. This observation suggests that *Z. mays* has undergone two rounds of whole-genome duplication. WGD and TD are the key forces behind gene expansion in *Populus* (Chiang et al., 2010; Barakat et al., 2011). COMTs of *maize*, *rice*, and *foxtail millet* have similar gene copy numbers (Liu et al. 2019). The SbCOMT homolog gene pairs had Ka/Ks ratios of <1, indicating that the SbCOMTs had undergone significant purifying selection. The cis-regulatory elements existing in the promoter regions were the binding sites of the COMTs gene with other proteins, which play an essential role in regulating gene transcription. There were a huge number of light-responsive elements, phytohormone-responsive elements, which involve plant defense mechanisms and growth, drought stress-responsive elements, and regulatory elements that promote lignin synthesis (Sega et al., 2020).

COMT expression was upregulated in plants when stressed or exposed to hormones. (Asif et al., 2014; Zhang et al., 2015; Li et al., 2016; FU et al., 2019). According to in silico expression analysis, SbCOMT26 and SbCOMT36 are highly expressed in all tissues of sorghum under natural conditions and also in leaf and root tissues under drought stress responses at different growth intervals. This study demonstrates the relationship between SbCOMTs, a crucial enzyme in the biosynthesis of monolignol, and the methods by which sorghum adapts to drought stress. Zhang et al. (2021) reported that the COMT gene family plays a significant role in plant defense to abiotic stress and lignification under drought conditions. Under drought conditions, lignin concentration increased considerably in the stems of *Eucalyptus urograndis* and *Eucalyptus globulus* (Moura-Sobczak et al. 2011). SbCOMTs are implicated in salt stress responses. Under salt stress, the contents of S and G units of lignin are raised in Coffea arabica (de Lima et al. 2014).

In many abiotic stress conditions, such as drought stress, the majority of the genes of the monolignol biosynthesis pathway are usually upregulated. This helps plants resist water loss by fortifying their cell walls.We also found that SbCOMT26 and 36 are highly stimulated under drought and salt stress; hence, these are potential targets for manipulation of lignin biosynthesis in sorghum to engineer biomass for better biofuel yield and enhanced abiotic stress tolerance.

CONCLUSION

In the present research, we identified COMT48 genes from *Sorghum bicolor*. Based on a phylogenetic investigation of COMTs, we divided the COMTs into two groups, which specified the existence of two ancestor genes. Gene characterization, conserved domains, motif identification, localization, and phylogenetic analysis revealed a close relationship between *Sorghum bicolor* COMT gene homologs and its relative *Oryza sativa* and *Zea mays*. The Ka/Ks ratios for the COMTs from *Sorghum* were less than one, indicating that the COMTs have undergone strong purifying selection. Identification of cis-acting elements and transcription factor prediction would be helpful to explore further and manipulate SbCOMT genes to design better biofuel crops. The miRNA prediction and elucidation of expression patterns under diverse abiotic stress conditions would help in the regulation of SbCOMT genes to engineer the lignin composition, further improving the biomass and enhancing abiotic stress tolerance in *Sorghum bicolor*.

ACKNOWLEDGMENTS

Prashanth Bollempally (PB) is thankful to CSIR, New Delhi for Junior Research Fellowship (JRF: File no: 09/132(0885)/2019-EMR-I) and Senior Research Fellowship (SRF). S. Prashant (SP) is thankful to DST-PURSE Programme-II for financial assistance to his lab.

AUTHOR CONTRIBUTION STATEMENT

PB and SP have designed the experiments and the structure of the article. PB has prepared the first draft. All others have added lateral text in the manuscript and refined it. APK, and VKA have prepared the figures. SP, SNK, PB, APK, and VKA have revised the manuscript. All authors have read and approved it.

AVAILABILITY OF DATA AND MATERIALS

The data that support the findings of this study are available from the corresponding author upon reasonable request.

STATEMENTS AND DECLARATIONS

Competing Interests:

The authors declare that they have no competing interests.

Funding This study was supported by the Council of Scientific and Industrial Research (CSIR), India.

Table 1: Characterization of SbCOMT homologs.

Gene Name	Transcript ID	Chr	CDS bp	Introns/Exons	A.a	Domain	Mw (kDa)	Localization
>SbCOMT-1	Sobic.010G230800.1	10	1125	01:02	375	Methyl transferase-2	41.366	Cytoplasmic
>SbCOMT-2	Sobic.010G231000.2	10	1287	01:02	429	Methyl transferase-2	47.265	Cytoplasmic
>SbCOMT-3	Sobic.010G234500.1	10	1149	01:02	383	Methyl transferase-2	42.007	Cytoplasmic
>SbCOMT-4	Sobic.010G234400.1	10	1152	01:02	384	Methyl transferase-2	42.131	Cytoplasmic
>SbCOMT-5	Sobic.003G298500.1	3	1152	01:02	384	Methyl transferase-2	41.293	Chloroplast
>SbCOMT-6	Sobic.009G197600.4	9	1089	01:02	363	Methyl transferase-2	39.438	Cytoplasmic
>SbCOMT-7	Sobic.009G197600.5	9	1089	01:02	363	Methyl transferase-2	39.438	Cytoplasmic
>SbCOMT-8	Sobic.009G197800.1	9	711	01:02	237	Methyl transferase-2	25.973	Cytoplasmic
>SbCOMT-9	Sobic.009G043900.1	9	1299	02:03	433	Methyl transferase-2	47.186	Plasma Membrane
>SbCOMT-10	Sobic.009G197400.1	9	1086	01:02	362	Methyl transferase-2	38.919	Plasma Membrane
>SbCOMT-11	Sobic.009G197000.1	9	1068	01:02	356	Methyl transferase-2	38.689	Cytoplasmic
>SbCOMT-12	Sobic.009G198000.1	9	1080	01:02	360	Methyl transferase-2	39.136	Cytoplasmic
>SbCOMT-13	Sobic.005G129100.1	5	1140	01:02	380	Methyl transferase-2	40.75	Cytoplasmic
>SbCOMT-14	Sobic.005G110451.1	5	690	01:02	230	Methyl transferase-2	25.578	Cytoplasmic
>SbCOMT-15	Sobic.005G086600.1	5	1125	01:02	375	Methyl transferase-2	40.754	Plasma Membrane
>SbCOMT-16	Sobic.005G045600.1	5	1116	01:02	372	Methyl transferase-2	39.886	Plasma Membrane
>SbCOMT-17	Sobic.005G101900.1	5	1098	01:02	366	Methyl transferase-2	40.51	Cytoplasmic
>SbCOMT-18	Sobic.005G216100.1	5	1092	01:02	364	Methyl transferase-2	38.627	Cytoplasmic
>SbCOMT-19	Sobic.005G224400.1	5	1119	01:02	373	Methyl transferase-2	40.898	Cytoplasmic
>SbCOMT-20	Sobic.005G107900.1	5	1101	01:02	367	Methyl transferase-2	39.939	Cytoplasmic
>SbCOMT-21	Sobic.005G216200.1	5	1092	01:02	364	Methyl transferase-2	38.641	Cytoplasmic
>SbCOMT-22	Sobic.005G224300.1	5	1119	01:02	373	Methyl transferase-2	41.077	Cytoplasmic
>SbCOMT-23	Sobic.008G014000.1	8	1182	01:02	394	Methyl transferase-2	42.377	Plasma Membrane
>SbCOMT-24	Sobic.008G013900.1	8	807	01:02	269	Methyl transferase-2	29.537	Chloroplast
>SbCOMT-25	Sobic.004G083500.1	4	1098	01:02	366	Methyl transferase-2	39.589	Cytoplasmic
>SbCOMT-26	Sobic.004G351400.1	4	1134	02:03	378	Methyl transferase-2	40.294	Chloroplast
>SbCOMT-27	Sobic.004G083401.1	4	933	02:03	311	Methyl transferase-2	33.646	Chloroplast
>SbCOMT-28	Sobic.004G341600.1	4	1176	01:02	392	Methyl transferase-2	41.801	Cytoplasmic
>SbCOMT-29	Sobic.004G128400.1	4	1089	01:02	363	Methyl transferase-2	39.287	Cytoplasmic
>SbCOMT-30	Sobic.004G341500.1	4	1209	00:01	403	Methyl transferase-2	43.573	Cytoplasmic
>SbCOMT-31	Sobic.007G099400.1	7	1092	01:02	364	Methyl transferase-2	40.032	Cytoplasmic
>SbCOMT-32	Sobic.007G058600.1	7	1110	03:04	370	Methyl transferase-2	40.347	Chloroplast
>SbCOMT-33	Sobic.007G170500.1	7	1107	01:02	369	Methyl transferase-2	39.719	Cytoplasmic
>SbCOMT-34	Sobic.007G058400.1	7	1131	03:04	377	Methyl transferase-2	40.846	Chloroplast
>SbCOMT-35	Sobic.007G074800.1	7	1125	01:02	375	Methyl transferase-2	40.628	Chloroplast
>SbCOMT-36	Sobic.007G047300.1	7	1089	01:02	363	Methyl transferase-2	39.59	Cytoplasmic
>SbCOMT-37	Sobic.007G058800.1	7	1125	03:04	375	Methyl transferase-2	41.064	Chloroplast
>SbCOMT-38	Sobic.007G059100.1	7	1134	03:04	378	Methyl transferase-2	41.492	Chloroplast
>SbCOMT-39	Sobic.001G354400.1	1	1098	01:02	366	Methyl transferase-2	39.583	Mitochondrial
>SbCOMT-40	Sobic.001G246700.1	1	852	00:01	284	Methyl transferase-2	33.125	Plasma Membrane

>SbCOMT-41	Sobic.001G246700.2	1	930	00:01	310	Methyl transferase-2	30.032	Plasma Membrane
>SbCOMT-42	Sobic.001G354200.1	1	1116	01:02	372	Methyl transferase-2	40.284	Cytoplasmic
>SbCOMT-43	Sobic.001G456650.1	1	300	03:04	100	Methyl transferase-2	39.602	Chloroplast
>SbCOMT-44	Sobic.006G008000.1	6	1125	01:02	375	Methyl transferase-2	40.857	Plasma Membrane
>SbCOMT-45	Sobic.006G007900.1	6	1125	01:02	375	Methyl transferase-2	40.884	Plasma Membrane
>SbCOMT-46	Sobic.002G079500.1	2	1140	01:02	380	Methyl transferase-2	40.46	Chloroplast
>SbCOMT-47	Sobic.002G079500.2	2	1032	02:03	344	Methyl transferase-2	36.41	Chloroplast
>SbCOMT-48	Sobic.002G077700.1	2	1188	01:02	396	Methyl transferase-2	42.752	Chloroplast

Table 2: SbCOMT protein parameters.

Gene Name	Protein length (A.A)	Protein Molecular Weight(kDa)	pI	GRAVY
SbCOMT-1	375	41.366	5.46	0.133
SbCOMT-2	429	47.265	5.51	0.036
SbCOMT-3	383	42.007	5.08	0.012
SbCOMT-4	384	42.131	5.21	0.026
SbCOMT-5	384	41.293	5.6	0.22
SbCOMT-6	363	39.438	5.42	0.15
SbCOMT-7	363	39.438	5.42	0.15
SbCOMT-8	237	25.973	4.65	0.248
SbCOMT-9	433	47.186	5.3	0.168
SbCOMT-10	362	38.919	5.45	0.204
SbCOMT-11	356	38.689	5.43	0.175
SbCOMT-12	360	39.136	5.8	0.231
SbCOMT-13	380	40.75	5.56	0.132
SbCOMT-14	230	25.578	5.46	-0.065
SbCOMT-15	375	40.754	5.13	0.2
SbCOMT-16	372	39.886	4.86	0.218
SbCOMT-17	366	40.51	5.61	0.056
SbCOMT-18	364	38.627	4.91	0.184
SbCOMT-19	373	40.898	5.38	0.117
SbCOMT-20	367	39.939	5.75	0.109
SbCOMT-21	364	38.641	4.91	0.185
SbCOMT-22	373	41.077	5.35	0.066
SbCOMT-23	394	42.377	6	0.081
SbCOMT-24	269	29.537	6.5	-0.065
SbCOMT-25	366	39.589	5.45	0.139
SbCOMT-26	378	40.294	5.32	-0.011
SbCOMT-27	311	33.646	5.2	0.106
SbCOMT-28	392	41.801	5	0.195
SbCOMT-29	363	39.287	5.39	0.172
SbCOMT-30	403	43.573	5.23	0.08
SbCOMT-31	364	40.032	5.39	0.086
SbCOMT-32	370	40.347	5.39	0.008
SbCOMT-33	369	39.719	5.84	0.194
SbCOMT-34	377	40.846	4.93	0.117
SbCOMT-35	375	40.628	7.13	0.086

SbCOMT-36	363	39.59	5.46	-0.015
SbCOMT-37	375	41.064	5.05	0.007
SbCOMT-38	378	41.492	5.15	0.052
SbCOMT-39	366	39.583	5.77	0.166
SbCOMT-40	284	33.125	5.72	0.205
SbCOMT-41	310	30.032	4.97	0.253
SbCOMT-42	372	40.284	5.76	0.174
SbCOMT-43	100	39.602	5.54	0.033
SbCOMT-44	375	40.857	5.13	0.199
SbCOMT-45	375	40.884	5.15	0.199
SbCOMT-46	380	40.46	4.94	0.181
SbCOMT-47	344	36.41	5.29	0.029
SbCOMT-48	396	42.752	5.55	-0.03

pI, isoelectric point, GRAVY, Grand average of hydropathicity index

Table 3: Non-synonymous and synonymous substitution rates of sorghum COMT paralog genes

Gene-1	**Gene-2**	**Ka**	**Ks**	**Ka_Ks**	**T(MYA)**
SbCOMT1	SbCOMT2	0.093790365	0.279215	0.335907	7.148655862
SbCOMT3	SbCOMT4	0.02093362	0.166845	0.125468	1.59555029
SbCOMT28	SbCOMT30	0.099433025	0.390486	0.254639	7.578736631
SbCOMT23	SbCOMT24	0.114992761	0.302893	0.379648	8.764692129
SbCOMT8	SbCOMT12	0.01094577	0.081252	0.134714	0.834281284
SbCOMT19	SbCOMT22	0.065428074	0.459634	0.142348	4.98689589
SbCOMT14	SbCOMT20	0.139548407	0.696896	0.200243	10.63631153
SbCOMT32	SbCOMT34	0.066990967	0.080131	0.836019	5.106018846
SbCOMT18	SbCOMT21	0.001836361	0.03202	0.057351	0.139966528

Ks-synonymous substitution; Ka-non-synonymous substitution; T(MYA)-Evolution time in Million years ago. Time calculated based on T=Ks/2x where x is $6.56x10^{-9}$ formula.

Table 4: COMT genes in the four different genomes sequenced

Species	Total no of protein coding genes	Predicted no of COMT genes	Genome size	Reference
Sorghum bicolor v3	34129	48	730Mb	Phytozome
Arabidopsis thaliana TAIR10	27416	17	135Mb	Ensembl
Oryza sativa v7	42189	39	500Mb	Ensembl
Zea mays v4	39498	32	2.13Gb	Phytozome

Table 5: Primers used for q-RT PCR study

S.no	Name	Sequence	Len	Tm	GC%
1	SbPAL3-FP	5-GGTCTTGTCCGCTCCCTGAAC-3	21	62.96	61.90
	SbPAL3-RP	5-TCGCGCCCTGGATCTTCAC-3	19	62.37	63.16
2	SbPAL8-FP	5- CTCGTCTCCGCCAGGAAGA-3	19	61.05	63.16
	SbPAL8-RP	5-GACGGGTTCATGGTCAGCAC-3	20	61.30	60.00
3	SbC4H2-FP	5- AACCTGATGTCCCTCGCCAA-3	20	61.49	55.00
	SbC4H2-RP	5-GGCCTTTCCCCGTGAAGATG-3	20	61.03	60.00
4	Sb4CL4-FP	5- TGCAGACCTACTGCTTCGGG-3	20	61.89	60.00
	Sb4CL4-RP	5-AGTTGCGGAGCAGGTTCATC-3	20	60.67	55.00
5	SbHCT2-FP	5- GACGACTACGGTGACTTCGC-3	20	60.79	60.00
	SbHCT2-RP	5-CCAGACATGCCATCCGCTAC-3	20	60.88	60.00
6	SbC3H1-FP	5- GGAGCACGCAAAGTCTCTCA-3	20	60.32	55.00
	SbC3H1-RP	5-TCTGCCATTGCCCACTCAAC-3	20	60.90	55.00
7	SbCCoAOMT3-FP	5- CAGTGGGGGTTCATGCAGTC-3	20	60.96	60.00
	SbCCoAOMT3-RP	5-TACTCCCTGCTCACGTCGAA-3	20	60.61	55.00
8	SbCCoAOMT1-FP	5-CGGAGGACGGCACGATCT-3	18	60.26	58.00
	SbCCoAOMT1-RP	5- CGAAGTCGAACGACCCGTG-3	19	59.30	59.69
9	SbCCR1-FP	5- GACCTGGGATTGGAGTTCCG-3	20	60.11	60.00
	SbCCR1-RP	5- CACGCACGGATGGCGATT-3	18	61.20	61.11
10	SbF5H1-FP	5- CATGGACGTGATGTTTGGCG-3	20	60.18	55.00
	SbF5H1-RP	5-TGAGGAAGGGGAGCTTGTCC-3	20	61.20	60.00
11	SbCAD2-FP	5- CGTCCGAGAGGAAGGTGGTC-3	20	61.94	65.00
	SbCAD2-RP	5-GGGTACTTTGAAGCCCCGAG-3	20	60.39	60.00

Tools & Database used in this study:

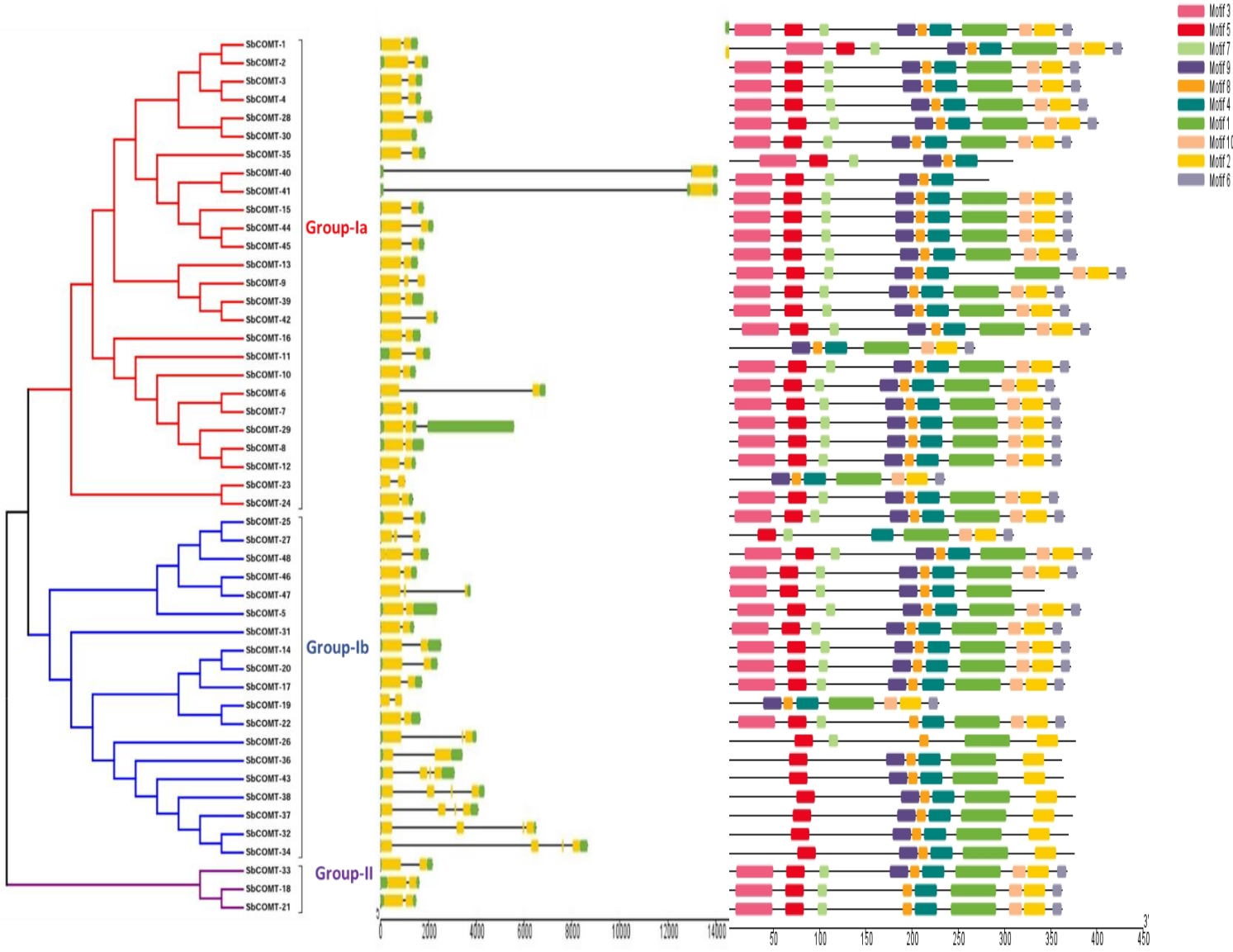
1. **MEGA v 7.0 (http://www.megasoftware.net)**
2. **TBtools (https://github.com/CJ-Chen/TBtools/releases)**
3. **Wolfpsort (https://wolfpsort.hgc.jp)**
4. **Phytozome (http://www.phytozome.net/)**
5. **MEME (http://meme-suite.org/tools/meme)**
6. **Protparam (https://web.expasy.org/protparam/)**
7. **SMART (http://smart.embl-heidelberg.de/)**
8. **PFAM database (http://pfam.xfam.org/ search)**
9. **Ka/Ks calculator (http://services.cbu.uib.no/tools/kaks)**

10. **PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/)**
11. **SWISS-MODEL SERVER (https://swissmodel.expasy.org/)**
12. **Protein structure verification server (SAVES v6.0) (https://saves.mbi.ucla.edu/).**
13. **STRING database (https://string-db.org/).**
14. **A Plant Small RNA Target Analysis Server (https://www.zhaolab.org/psRNATarget/)**
15. **Cytoscape (https://cytoscape.org/).**
16. **Plant transcription factor database (PTFDB)**

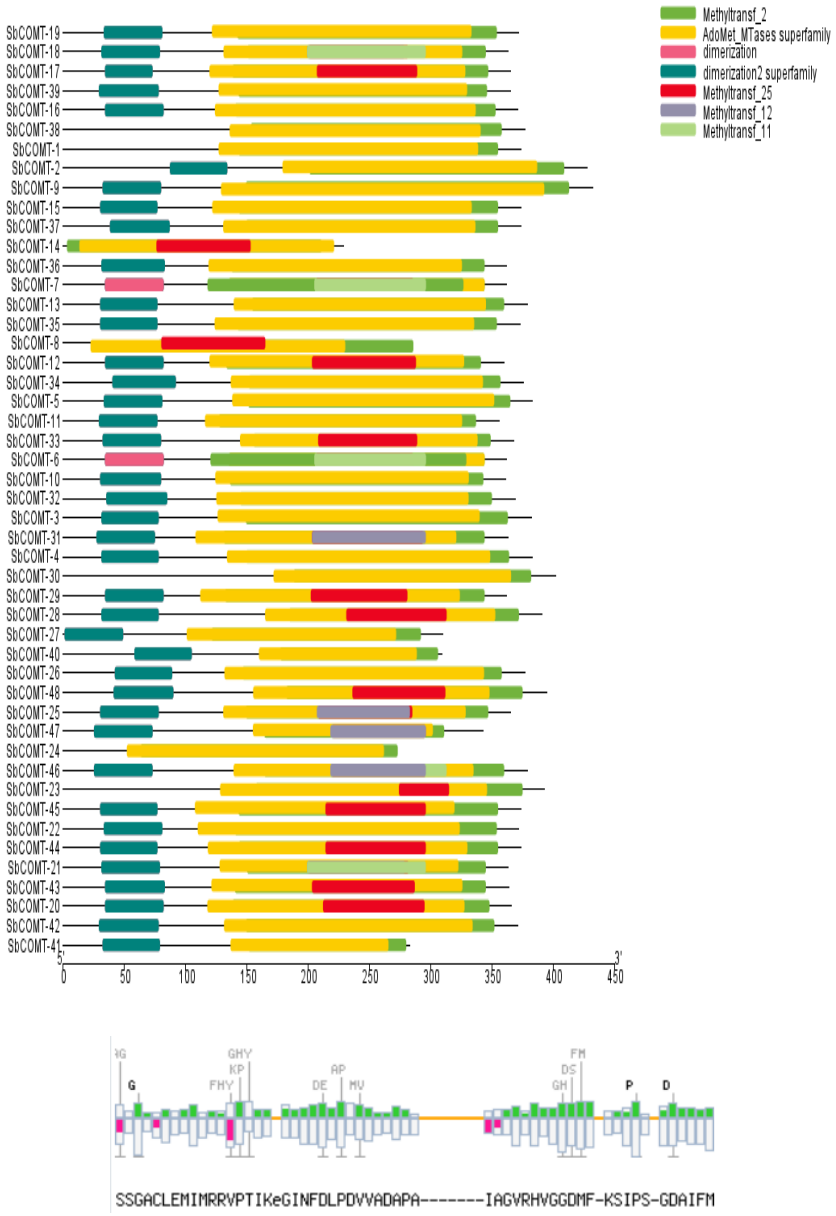
(http://planttfdb.gao-lab.org/)

17. **Gramene database (https://www.gramene.org/).**
18. **MOROKOSHI Sorghum transcriptome database** (http://sorghum.riken.jp).

a. Evolutionary tree **b. Gene structure** **c. Motif pattern**



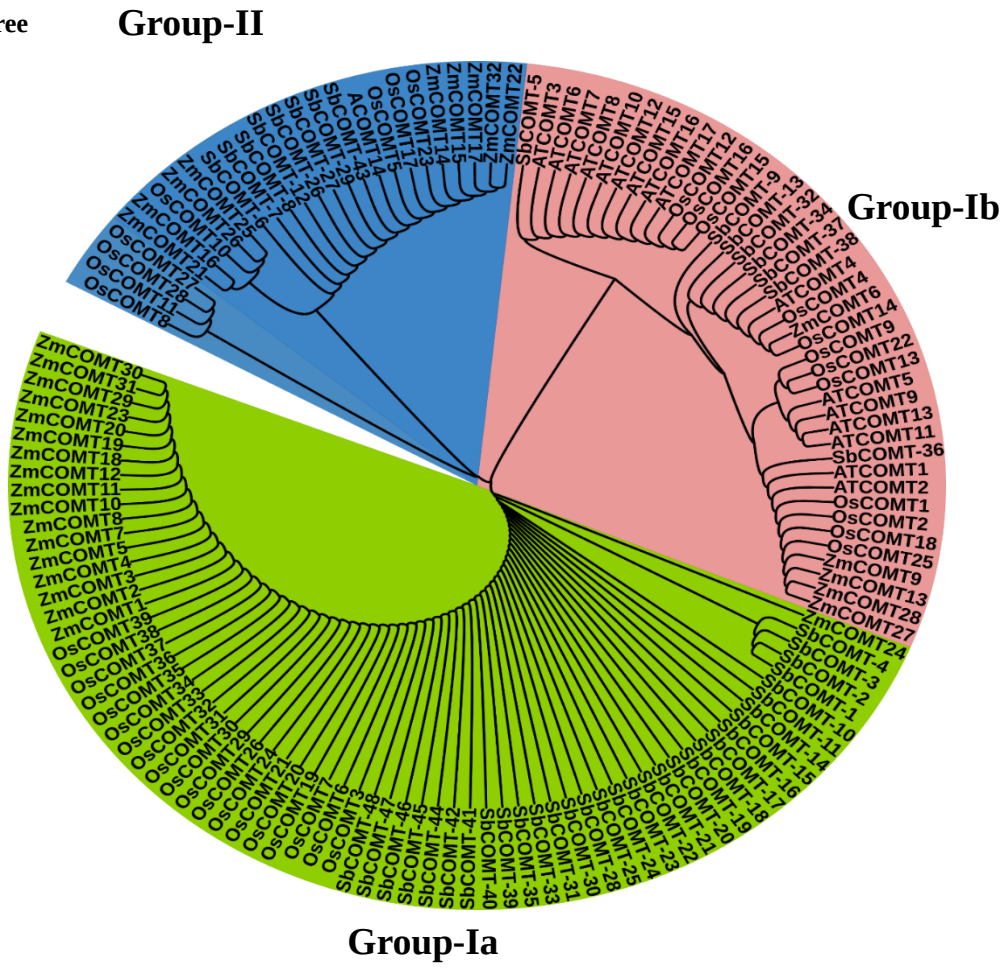
d.Conserved domains



e. Methyl transferase_2 domain (O-Methyltransferase)

Fig 1: The evolutionary relationship, gene structure, and motif analysis of the 48 SbCOMTs from sorghum bicolor. a. The phylogenetic tree was constructed by MEGA v10.0 with the NJ method. b Structures of the 48 putative SbCOMT genes. c Motif distribution of SbCOMTs proteins, d Conserved domains, and e. Methyl transferase_2 domain. The different motifs are designated by different colours.

a. Phylogenetic tree



b. Physical mapping of sorghum COMT gene homologs.

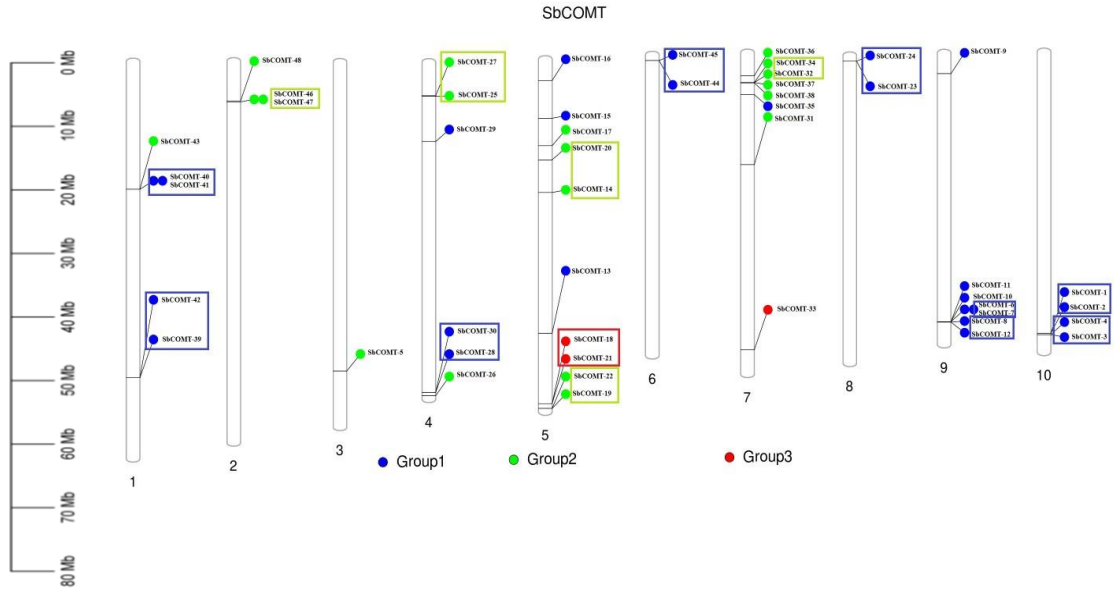
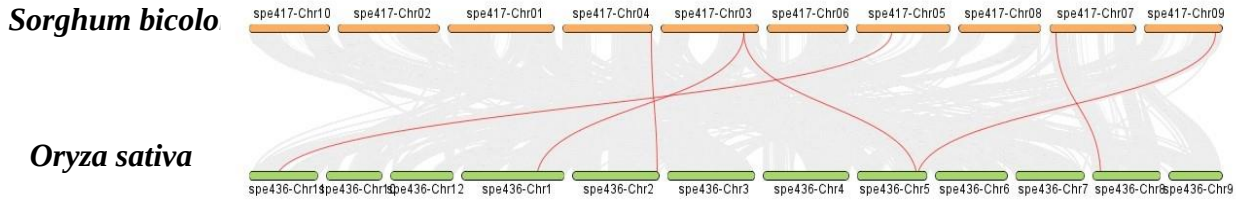
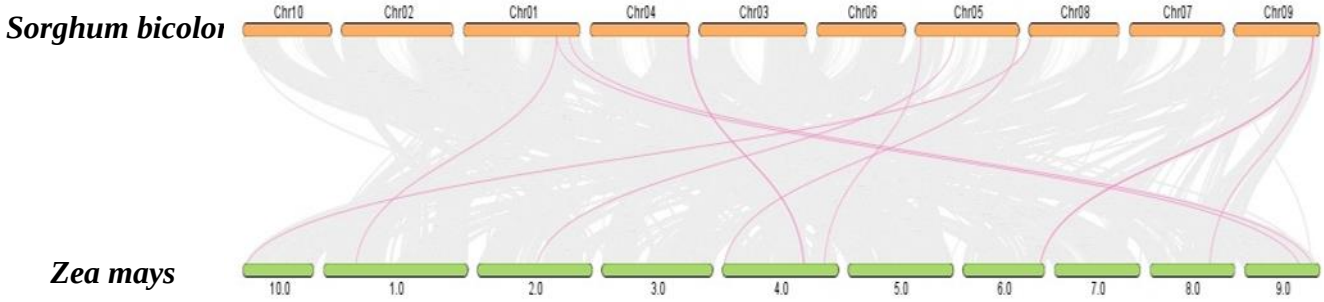


Fig 2:a. Phylogenetic tree representing the evolutionary relationship of COMTs from Sorghum bicolor, Oryza sativa, Zea mays, and Arabidopsis thaliana. and b. Physical mapping of sorghum COMT gene homologs. The 15 Paralog gene pairs are

represented in coloured boxes. The blue, green, and red coloured boxes represent Group Ia, Group Ib, and Group II SbCOMTs respectively



a. Synteny analysis of COMT genes between Sorghum and Oryza



b. Synteny analysis of COMT genes between Sorghum and Maize

Fig 3: a. Synteny analysis of SbCOMT genes between Sorghum and Oryza sativa and b. Sorghum and Zea mays. Gray lines in the background indicate the collinear blocks within sorghum and other plant genomes. The pink colour lines represent COMTs with collinearity in different genomes.

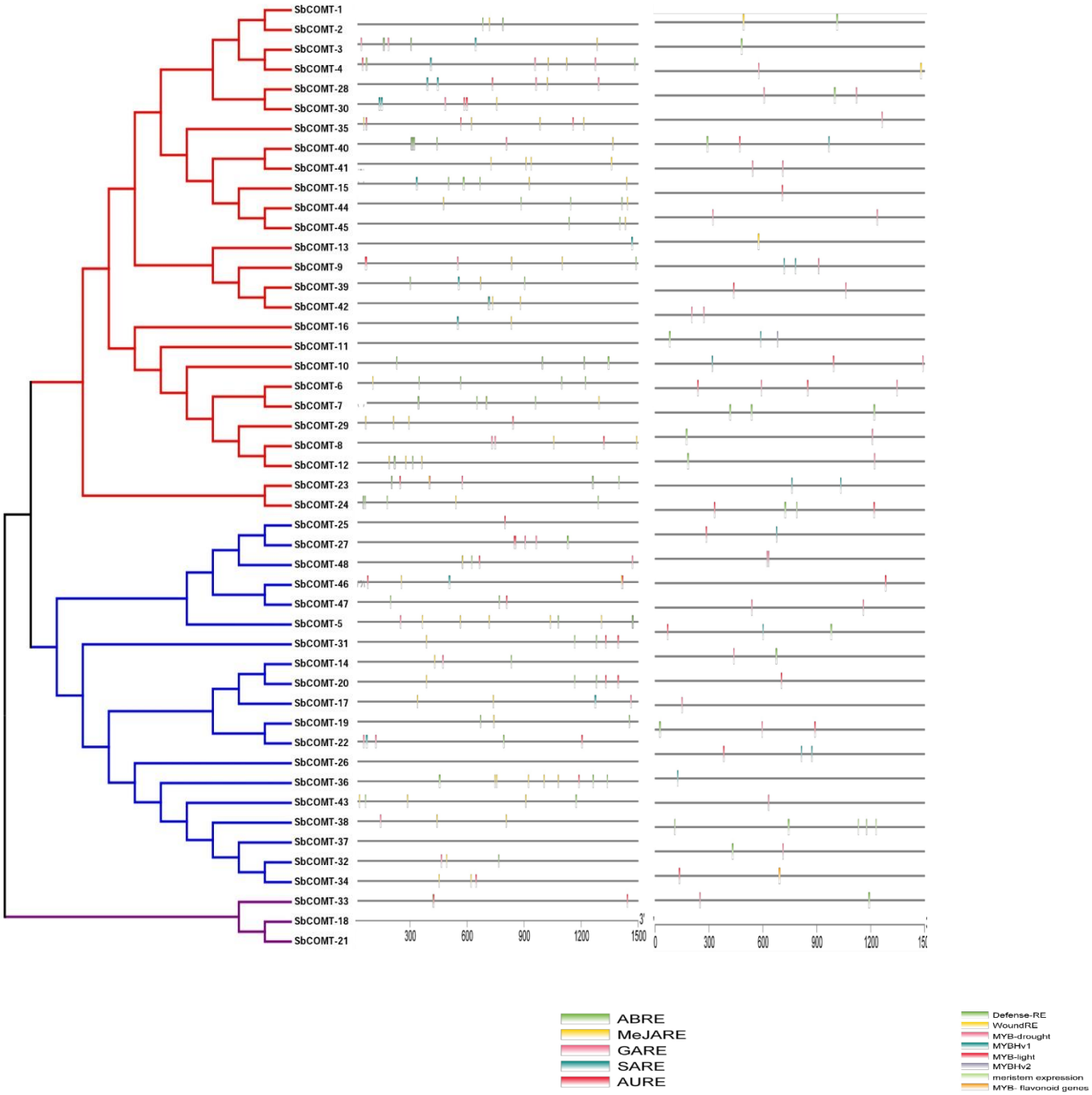
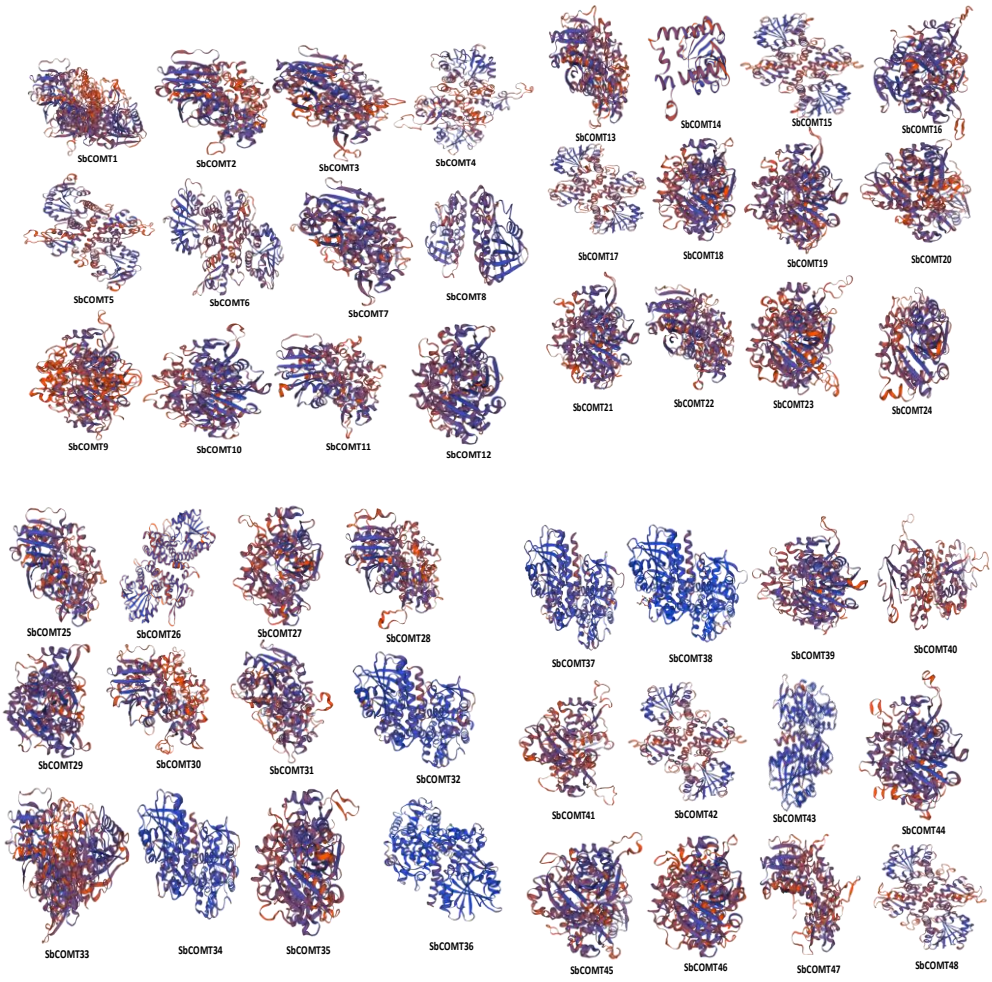
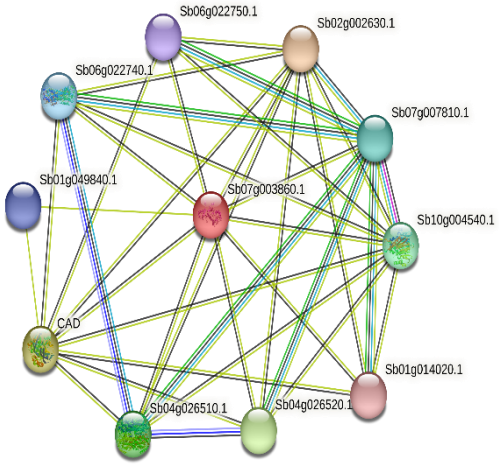


Fig 4: Predicted cis-regulatory elements in the promoter regions of Sb COMT genes



a.

3D structure analysis of SbCOMT homologs

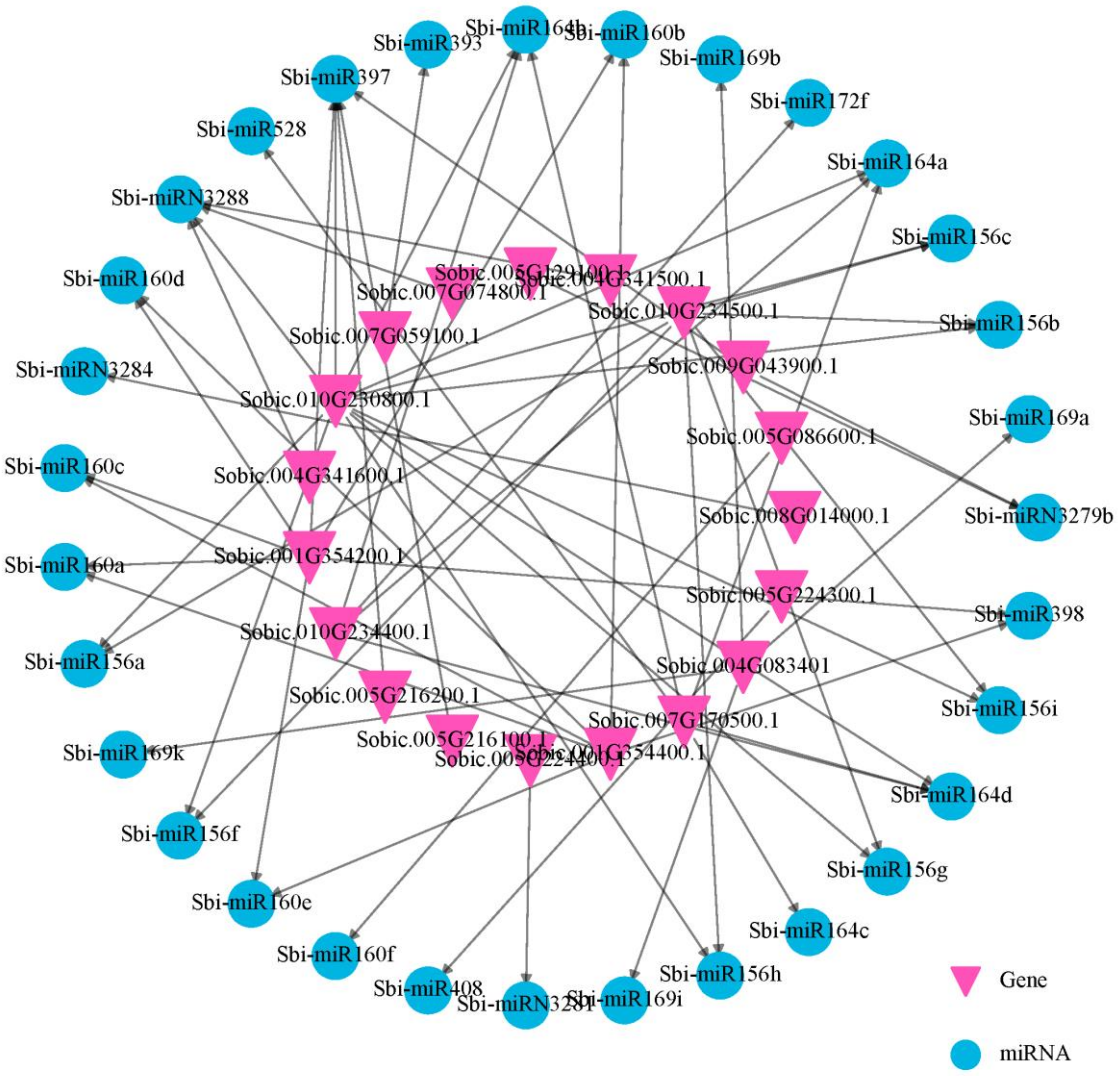


b. String analysis

Fig 5:a. Structural analysis of 48 modelled sorghum bicolor SbCOMT proteins. b. String analysis of sorghum COMT. The SbCOMT protein exhibited interaction with various lignin biosynthetic pathways and secondary metabolite partners.SbCOMT (Sb07g003860.1), CAD (Sb04g005950.1), Phenylalanine ammonia-lyase (Sb04g026510.1, Sb04g026520.1, Sb01g014020.1,

Sb06g022750.1, Sb06g022740.1), Probable 4-coumarate-CoA ligase 1(Sb07g007810.1) F5H(Sb02g002630.1) Folylpolyglutamate synthase (Sb01g049840.1) SbCOMT (Sb07g003860.1)

a



b.

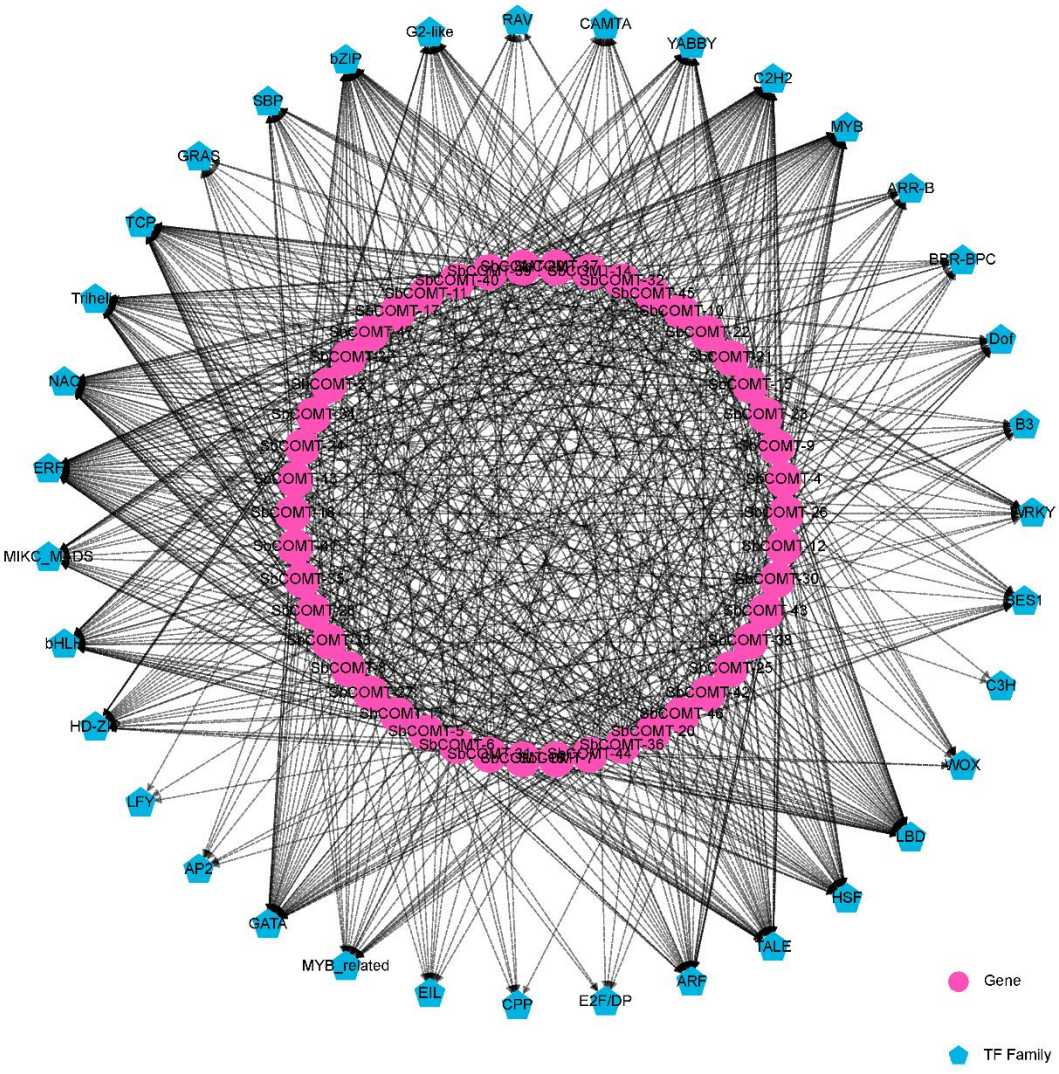


Fig 6: a. miRNA prediction of SbCOMT genes. The blue colour indicates predicted miRNA targets and the yellow colour boxes represent candidate SbCOMT homologs; b. Transcription factor prediction analysis of SbCOMT homologs. NAC, MYB, and MYB-related TFs are indicated in blue colour. The network is built with Cytoscape.

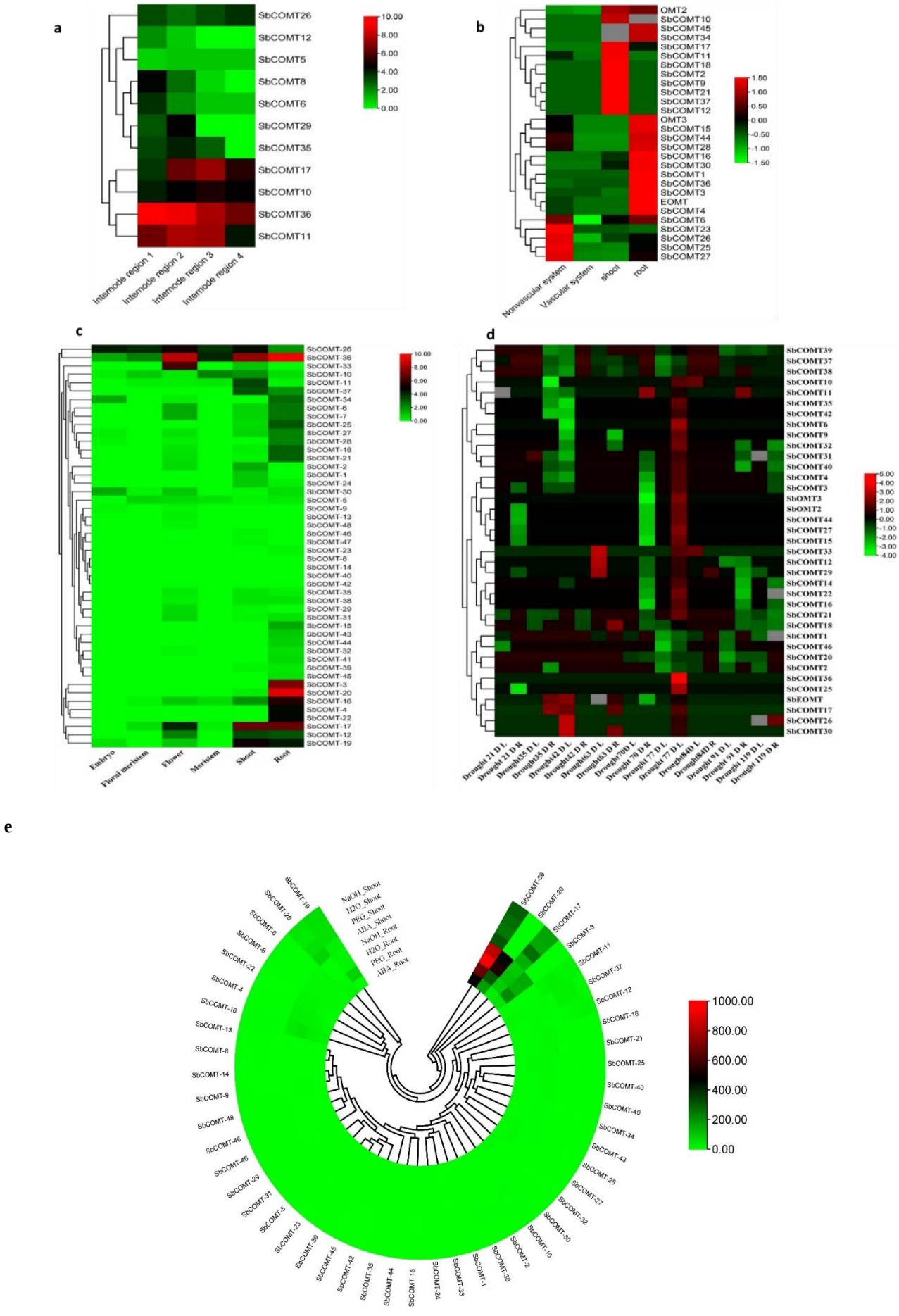

a
SbCOMT26
SbCOMT12
SbCOMT5
SbCOMT8
SbCOMT6
SbCOMT29
SbCOMT35
SbCOMT17
SbCOMT10
SbCOMT36
SbCOMT11
10.00
8.00
6.00
4.00
2.00
0.00
Internode region 1
Internode region 2
Internode region 3
Internode region 4
b
OMT2
SbCOMT10
SbCOMT45
SbCOMT34
SbCOMT17
SbCOMT11
SbCOMT18
SbCOMT2
SbCOMT9
SbCOMT21
SbCOMT37
SbCOMT12
OMT3
SbCOMT15
SbCOMT44
SbCOMT28
SbCOMT16
SbCOMT30
SbCOMT1
SbCOMT36
SbCOMT3
EOMT
SbCOMT4
SbCOMT6
SbCOMT23
SbCOMT26
SbCOMT25
SbCOMT27
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
Nonvascular system
Vascular system
shoot
root
c
SbCOMT-26
SbCOMT-36
SbCOMT-33
SbCOMT-10
SbCOMT-11
SbCOMT-37
SbCOMT-34
SbCOMT-6
SbCOMT-7
SbCOMT-25
SbCOMT-27
SbCOMT-28
SbCOMT-18
SbCOMT-21
SbCOMT-2
SbCOMT-1
SbCOMT-24
SbCOMT-30
SbCOMT-5
SbCOMT-9
SbCOMT-13
SbCOMT-48
SbCOMT-46
SbCOMT-47
SbCOMT-23
SbCOMT-8
SbCOMT-14
SbCOMT-40
SbCOMT-42
SbCOMT-35
SbCOMT-38
SbCOMT-29
SbCOMT-31
SbCOMT-15
SbCOMT-43
SbCOMT-44
SbCOMT-32
SbCOMT-41
SbCOMT-39
SbCOMT-45
SbCOMT-3
SbCOMT-20
SbCOMT-16
SbCOMT-4
SbCOMT-22
SbCOMT-17
SbCOMT-12
SbCOMT-19
10.00
8.00
6.00
4.00
2.00
0.00
Embryo
Floral meristem
Flower
Meristem
Shoot
Root
d
SbCOMT39
SbCOMT37
SbCOMT38
SbCOMT10
SbCOMT11
SbCOMT35
SbCOMT42
SbCOMT6
SbCOMT9
SbCOMT32
SbCOMT31
SbCOMT40
SbCOMT4
SbCOMT3
SbOMT3
SbOMT2
SbCOMT44
SbCOMT27
SbCOMT15
SbCOMT33
SbCOMT12
SbCOMT29
SbCOMT14
SbCOMT22
SbCOMT16
SbCOMT21
SbCOMT18
SbCOMT1
SbCOMT46
SbCOMT20
SbCOMT2
SbCOMT36
SbCOMT25
SbEOMT
SbCOMT17
SbCOMT26
SbCOMT30
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
Drought 21 D L
Drought 21 D R
Drought35 D L
Drought35 D R
Drought42 D L
Drought42 D R
Drought63 D L
Drought63 D R
Drought70 D L
Drought 70 D R
Drought 77 D L
Drought 77 D R
Drought84 D L
Drought84 D R
Drought 91 D L
Drought 91 D R
Drought 119 D L
Drought 119 D R
e
NaOH_Shoot
H2O_Shoot
PEG_Shoot
ABA_Shoot
NaOH_Root
H2O_Root
PEG_Root
ABA_Root
1000.00
800.00
600.00
400.00
200.00
0.00

Fig 7: In silico expression analysis of SbCOMTs; a. different regions of stem internodes, b. vascular and non-vascular system, c. baseline expression patterns in various tissues and organs, d. expression patterns under drought stress conditions and e. expression patterns under ABA, PEG, and NaOH stress conditions.

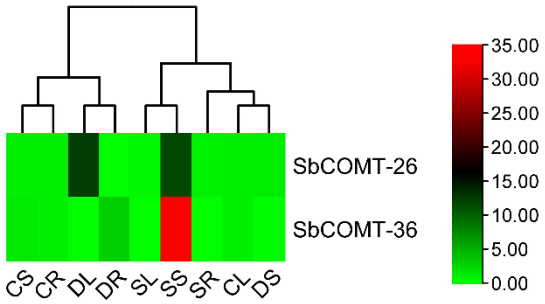


Fig 8: Heat map representing the expression patterns (Fold change) of SbCOMT genes in S. bicolor analyzed by qRT PCR. CL-control leaf, CS-control stem, CR-control root, DL-drought leaf, DS-drought stem, DR-drought root, SL-salt leaf, SS-salt stem, and SR-salt root. (p=0.05)

REFERENCES

[1] Asif, M. H., Lakhwani, D., Pathak, S., Bhambhani, S., Bag, S. K., & Trivedi, P. K. (2014). Genome-wide identification and expression analysis of the mitogen-activated protein kinase gene family from banana suggest involvement of specific members in different stages of fruit ripening. *Functional & integrative genomics*, *14*, 161-175.

[2] Barakat, A., Choi, A., Yassin, N. B. M., Park, J. S., Sun, Z., & Carlson, J. E. (2011). Comparative genomics and evolutionary analyses of the O-methyltransferase gene family in Populus. *Gene*, *479*(1-2), 37-46.

[3] Baucher, M., Halpin, C., Petit-Conil, M., &Boerjan, W. (2003). Lignin: genetic engineering and impact on pulping. Critical Reviews in Biochemistry and Molecular Biology, 38(4), 305-350.

[4] Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., ... &Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic acids research*, *42*(W1), W252-W258.

[5] Bugos, R. C., Chiang, V. L., & Campbell, W. H. (1991). cDNA cloning, sequence analysis and seasonal expression of lignin-bispecific caffeic acid/5-hydroxyferulic acid O-methyltransferase of aspen. *Plant molecular biology*, *17*, 1203-1215.

[6] Carocha, V.; Soler, M.; Hefer, C.; Cassan-Wang, H.; Fevereiro, P.; Myburg, A.A.; Paiva, J.A.; Grima-Pettenati, J. Genome-wideanalysis of the lignin toolbox of Eucalyptus grandis. New Phytol. 2015, 206, 1297–1313.

[7] Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., & Xia, R. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular plant*, *13*(8), 1194-1202.

[8] Chen, S., Zhao, Y., Zhao, X., & Chen, S. (2020). Identification of putative lignin biosynthesis genes in Betula pendula. *Trees*, *34*, 1255-1265.

[9] Chiang, V.L. Towards a systems approach for lignin biosynthesis in Populus trichocarpa: Transcript abundance and specificity ofthe monolignol biosynthetic genes. Plant Cell Physiol. 2010, 51, 144–163.

[10] Chiron, H., Drouet, A., Lieutier, F., Payer, H. D., Ernst, D., &Sandermann Jr, H. (2000). Gene induction of stilbene biosynthesis in Scots pine in response to ozone treatment, wounding, and fungal infection. *Plant Physiology*, *124*(2), 865-872.

[11] Collinge, M., &Boller, T. (2001). Differential induction of two potato genes, Stprx2 and StNAC, in response to infection by Phytophthora infestans and to wounding. *Plant molecular biology*, *46*, 521-529.

[12] Davidson, R. M., Gowda, M., Moghe, G., Lin, H., Vaillancourt, B., Shiu, S. H., ... & Robin Buell, C. (2012). Comparative transcriptomics of three Poaceae species reveals patterns of gene expression evolution. The Plant Journal, 71(3), 492-502.

[13] de Lima RB, dos Santos TB, Vieira LGE, de Lourdes Lúcio FerrareseM, Ferrarese-Filho O, et al. 2014. Salt stress alters the cell wallpolysaccharides and anatomy of coffee (Coffea arabica L.) leafcells. Carbohydrate Polymers 112:686–94

[14] Dugas, D. V., Monaco, M. K., Olson, A., Klein, R. R., Kumari, S., Ware, D., & Klein, P. E. (2011). Functional annotation of the transcriptome of Sorghum bicolor in response to osmotic stress and abscisic acid. *BMC genomics*, *12*, 1-21.

[15] Eudes, A., Dutta, T., Deng, K., Jacquet, N., Sinha, A., Benites, V. T., ... &Loqué, D. (2017). SbCOMT (Bmr12) is involved in the biosynthesis of tricin-lignin in sorghum. *PLoS One*, *12*(6), e0178160.

[16] FU, L. P., XIAO, Y. G., Jun, Y. A. N., LIU, J. D., WEN, W. E., Zhang, Y., ... & HE, Z. H. (2019). Characterization of TaCOMT genes associated with stem lignin content in common wheat and development of a gene-specific marker. *Journal of Integrative Agriculture*, *18*(5), 939-947.

[17] Goujon, T., Sibout, R., Eudes, A., MacKay, J., &Jouanin, L. (2003). Genes involved in the biosynthesis of lignin precursors in Arabidopsis thaliana. Plant Physiology and Biochemistry, 41(8), 677-687.

[18] Held, B. M., Wang, H., John, I., Wurtele, E. S., & Colbert, J. T. (1993). An mRNA putatively coding for an O-methyltransferase accumulates preferentially in maize roots and is located predominantly in the region of the endodermis. *Plant Physiology*, *102*(3), 1001-1008.

[19] Jin, J., Zhang, H., Kong, L., Gao, G., & Luo, J. (2014). PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucleic acids research*, *42*(D1), D1182-D1187.

[20] Kebrom, T. H., McKinley, B., & Mullet, J. E. (2017). Dynamics of gene expression during development and expansion of vegetative stem internodes of bioenergy sorghum. Biotechnology for Biofuels, 10(1), 1-16.

[21] Kim, M. H., Cho, J. S., Jeon, H. W., Sangsawang, K., Shim, D., Choi, Y. I., ... & Ko, J. H. (2019). Wood transcriptome profiling identifies critical pathway genes of secondary wall biosynthesis and novel regulators for vascular cambium development in Populus. *Genes*, *10*(9), 690.

[22] Kim, S. G., Kim, S. Y., & Park, C. M. (2007). A membrane-associated NAC transcription factor regulates salt-responsive flowering via FLOWERING LOCUS T in Arabidopsis. *Planta*, *226*, 647-654.

[23] Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution*, *35*(6), 1547.

[24] Lee, H. Y., Byeon, Y., Tan, D. X., Reiter, R. J., & Back, K. (2015). Arabidopsis serotonin N-acetyltransferase knockout mutant plants exhibit decreased melatonin and salicylic acid levels resulting in susceptibility to an avirulent pathogen. *Journal of Pineal Research*, *58*(3), 291-299.

[25] Li, W., Lu, J., Lu, K., Yuan, J., Huang, J., Du, H., & Li, J. (2016). Cloning and phylogenetic analysis of Brassica Napus L. Caffeic acid O-Methyltransferase 1 gene family and its expression pattern under drought stress. *PLoS One*, *11*(11), e0165975.

[26] Liu, X., Luo, Y., Wu, H., Xi, W., Yu, J., Zhang, Q., & Zhou, Z. (2016). Systematic analysis of O-methyltransferase gene family and identification of potential members involved in the formation of O-methylated flavonoids in Citrus. *Gene*, *575*(2), 458-472.

[27] Liu, Y., Wang, Y., Pei, J., Li, Y., & Sun, H. (2021). Genome-wide identification and characterization of COMT gene family during the development of blueberry fruit. *BMC Plant Biology*, *21*(1), 1-16.

[28] Liu, Z., Fan, M., Li, C., & Xu, J. H. (2019). Dynamic gene amplification and function diversification of grass-specific O-methyltransferase gene family. *Genomics*, *111*(4), 687-695.

[29] Louie, G. V., Bowman, M. E., Tu, Y., Mouradov, A., Spangenberg, G., & Noel, J. P. (2010). Structure-function analyses of a caffeic acid O-methyltransferase from perennial ryegrass reveal the molecular basis for substrate preference. *The Plant Cell*, *22*(12), 4114-4127.

[30] Lu, N.; Ma, W.; Han, D.; Liu, Y.; Wang, Z.; Wang, N.; Yang, G.; Qu, G.; Wang, Q.; Zhao, K.; et al. Genome-wide analysis of theCatalpa bungei caffeic acid O-methyltransferase (COMT) gene family: Identification and expression profiles in normal, tension,and opposite wood. PeerJ 2019, 7, 6520.

[31] McCarthy, R. L., Zhong, R., & Ye, Z. H. (2009). MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant and Cell Physiology*, *50*(11), 1950-1964.

[32] Moura-Sobczak, J.; Souza, U.; Mazzafera, P. Drought Stress and Changes in the Lignin Content and Composition in Eucalyptus. In Proceedings of the BMC Proceedings, Porto Seguro, Brazil, 26 June–2 July 2011; Springer: Berlin/Heidelberg, Germany, 2011; Volume 5, p. 1.

[33] Noel, J. P., Dixon, R. A., Pichersky, E., Zubieta, C., & Ferrer, J. L. (2003). Chapter two Structural, functional, and evolutionary basis for methylation of plant small molecules. In *Recent advances in phytochemistry* (Vol. 37, pp. 37-58). Elsevier.

[34] Nomura, T., &Kutchan, T. M. (2010). Three new O-methyltransferases are sufficient for all O-methylation reactions of ipecac alkaloid biosynthesis in root culture of Psychotria ipecacuanha. *Journal of Biological Chemistry*, *285*(10), 7722-7738.

[35] Ren, T., Qu, F., & Morris, T. J. (2000). HRT gene function requires interaction between a NAC protein and viral capsid protein to confer resistance to turnip crinkle virus. *The Plant Cell*, *12*(10), 1917-1925.

[36] Sambangi, P., & Rani, P. U. (2016). Physiological effects of resveratrol and coumaric acid on two major groundnut pests and their egg parasitoidbehavior. *Archives of Insect Biochemistry and Physiology*, *91*(4), 230-245.

[37] Sega, P., Kruszka, K., Szewc, Ł., Szweykowska-Kulińska, Z., &Pacak, A. (2020). Identification of transcription factors that bind to the 5′-UTR of the barley PHO2 gene. *Plant molecular biology*, *102*, 73-88.

[38] Stracke, R., Werber, M., &Weisshaar, B. (2001). The R2R3-MYB gene family in Arabidopsis thaliana. *Current opinion in plant biology*, *4*(5), 447-456.

[39] Takada, S., Hibara, K. I., Ishida, T., &Tasaka, M. (2001). The CUP-SHAPED COTYLEDON1 gene of Arabidopsis regulates shoot apical meristem formation. *Development*, *128*(7), 1127-1135.

[40] Turco, G. M., Kajala, K., Kunde-Ramamoorthy, G., Ngan, C. Y., Olson, A., Deshphande, S., ... & Brady, S. M. (2017). DNA methylation and gene expression regulation associated with vascularization in Sorghum bicolor. New Phytologist, 214(3), 1213-1229. Exp

[41] Varoquaux, N., Cole, B., Gao, C., Pierroz, G., Baker, C. R., Patel, D., ... &Purdom, E. (2019). Transcriptomic analysis of field-droughted sorghum from seedling to maturity reveals biotic and metabolic responses. Proceedings of the National Academy of Sciences, 116(52), 27124-27132.

[42] Vroemen, C. W., Mordhorst, A. P., Albrecht, C., Kwaaitaal, M. A., & de Vries, S. C. (2003). The CUP-SHAPED COTYLEDON3 gene is required for boundary and shoot meristem formation in Arabidopsis. *The Plant Cell*, *15*(7), 1563-1577.

[43] Wei, L.; Lu, J.; Lu, K.; Yuan, J.; Li, J. Cloning and phylogenetic analysis of Brassica napus L. Caffeic acid O-methyltransferase 1 genefamily and its expression pattern under drought stress. PLoS ONE 2016, 11, e0165975.

[44] Weir, I., Lu, J., Cook, H., Causier, B., Schwarz-Sommer, Z., & Davies, B. (2004). CUPULIFORMIS establishes lateral organ boundaries in Antirrhinum.

[45] Wu, X., Wu, J., Luo, Y., Bragg, J., Anderson, O., Vogel, J., & Gu, Y. Q. (2013). Phylogenetic, molecular, and biochemical characterization of caffeic acid o-methyltransferase gene family in Brachypodiumdistachyon. *International journal of plant genomics*, *2013*.

[46] Xie, Q., Sanz-Burgos, A. P., Guo, H., García, J. A., & Gutiérrez, C. (1999). GRAB proteins, novel members of the NAC domain family, isolated by their interaction with a geminivirus protein. *Plant molecular biology*, *39*, 647-656.

[47] Yamaguchi, M., Kubo, M., Fukuda, H., &Demura, T. (2008). VASCULAR-RELATED NAC-DOMAIN7 is involved in the differentiation of all types of xylem vessels in Arabidopsis roots and shoots. *The Plant Journal*, *55*(4), 652-664.

[48] Zhang, J., Li, Y., Jia, H. X., Li, J. B., Huang, J., Lu, M. Z., & Hu, J. J. (2015). The heat shock factor gene family in Salix suchowensis: a genome-wide survey and expression profiling during development and abiotic stresses. *Frontiers in Plant Science*, *6*, 748.

[49] Zhang, X., Chen, B., Wang, L., Ali, S., Guo, Y., Liu, J., ... & Zhang, Q. (2021). Genome-Wide Identification and Characterization of Caffeic Acid O-Methyltransferase Gene Family in Soybean. Plants, 10(12), 2816.

[50] Zhong, R., & Ye, Z. H. (2012). MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant and Cell Physiology*, *53*(2), 368-380.

[51] Zhong, R., Demura, T., & Ye, Z. H. (2006). SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of Arabidopsis. *The Plant Cell*, *18*(11), 3158-3170.

[52] Yao, Z., Zhang, X., Liang, Y., Zhang, J., Xu, Y., Chen, S. and Zhao, D., 2022. NtCOMT1 responsible for phytomelatonin biosynthesis confers drought tolerance in Nicotiana tabacum. Phytochemistry, 202, p.113306.

[53] Chang, J., Guo, Y., Yan, J., Zhang, Z., Yuan, L., Wei, C., Zhang, Y., Ma, J., Yang, J., Zhang, X. and Li, H., 2021. The role of watermelon caffeic acid O-methyltransferase (ClCOMT1) in melatonin biosynthesis and abiotic stress tolerance. Horticulture Research, 8.