Journal of Soil Science and Plant Nutrition
https://doi.org/10.1007/s42729-025-02964-1

ORIGINAL PAPER

Computational Analysis of Three-Amino Acid Loop Extension (TALE) Proteins in *Cicer arietinum* L. And Evaluation of Their Transcript Levels Under Drought Stress Conditions

Serdar Coşkun[1] · Aybüke Okay[1] · İlker Büyük[1] · Emine Sümer Aras[1]

Received: 21 May 2025 / Accepted: 20 December 2025
© The Author(s) 2026

Abstract
Chickpea is an agronomically and nutritionally important legume. The *TALE* transcription factor gene family plays a key role in plant growth and development. However, no comprehensive study has yet characterized the *TALE* gene family in chickpea. Therefore, the main objective of this research was to identify and characterize members of the *TALE* gene family in chickpea for the first time. A total of 27 *TALE* gene family members were identified in chickpea. Based on bioinformatic analyses, these genes were classified into two subfamilies: BEL1-like (BLH/BELL homeodomain) and KNOX (KNOTTED-like homeodomain). Various in silico analyses were performed, including intron–exon structure analysis, conserved motif identification, chromosomal localization, gene duplication assessment, cis-element and miRNA target prediction. Gene expression patterns were analyzed using RNA-seq drought datasets and validated by qPCR. The bioinformatic and expression analyses revealed that *CaTALE* genes play distinct regulatory roles under drought stress conditions. The structural and functional characteristics identified in this study highlight the evolutionary diversification of the *TALE* gene family in chickpea. Overall, the findings on gene structure, conserved domains, duplication events, and abiotic stress responses provide valuable insights into the evolution and functional roles of *TALE* genes in chickpea. This study serves as the first comprehensive resource for future functional and evolutionary studies of the *TALE* gene family in chickpea.

Keywords Chickpea · (*C. arietinum* L.) · Gene duplications · Bioinformatic analyses · Chromosomal distribution

1 Introduction

Chickpea (*Cicer arietinum* L.) is the sole domesticated species among the 43 species within the *Cicer* genus. It is classified as an annual diploid legume (2n = 16) that primarily undergoes self-pollination (Rasheed et al. 2021). Abiotic stressors represent one of the major threats to agriculture, causing damage to agricultural activities, degradation of arable land and reductions in crop yields. Abiotic stress factors rank first among the causes of yield loss in agricultural areas and are known to decrease yields by more than 50% in the most productive crop species (Wang et al. 2004). Therefore, to understand stress mechanism in plants in detail, it is essential to describe abiotic stress factors and plant responses as systematically as possible. Understanding the genetic and molecular basis of plant responses to abiotic stress is essential for improving crop tolerance and productivity.

Transcription factors play a pivotal role in regulating various cellular processes, including cell development, differentiation, and responses to environmental signals. By modulating gene expression, they contribute significantly to the proper growth and functioning of an organism (Lambert et al. 2018). Among the diverse families of transcription factors, homeobox genes represent one of the most evolutionarily conserved and functionally significant

✉ İlker Büyük
buyuki@ankara.edu.tr

Serdar Coşkun
bio.serdarcoskun@gmail.com

Aybüke Okay
aokay@ankara.edu.tr

Emine Sümer Aras
aras@science.ankara.edu.tr

[1] Department of Biology, Graduate School of Natural and Applied Sciences, Ankara University, Ankara, Turkey

Published online: 02 February 2026

groups. The regulation of transcription factor (TF) activity can occur through changes in the abundance of its active form via transcription, translation, and post-translational modifications or through alterations in the accessibility of binding sites, which involve epigenetic mechanisms and cell type-specific chromatin configurations. Once transcription factors (TFs) bind to DNA, they can either facilitate chromatin opening to enable the binding of additional transcription factors, inhibit the binding of certain factors, or regulate gene expression by either activating or repressing transcription (Filtz et al. 2014). Homeobox genes are characterized by a highly conserved 180 bp DNA sequence that encodes a 60-amino-acid DNA-binding domain known as the "homeodomain." These genes function as transcription factors by recognizing specific DNA sequences and modulating the expression of target genes. The discovery of the first homeobox genes occurred in *Drosophila*, where they were identified as crucial regulators of embryonic development, guiding the formation of various structures (Quiring et al. 1994).

The three- amino acid loop extension (*TALE*) superfamily genes are widely present in plants and play crucial roles in regulating plant growth, development, responses to abiotic stress (Zhao et al. 2019). Early studies classified plant homeobox genes into 11 distinct groups, including LD, NDX, DDT, PHD, SAWADEE, PINTOX, HD-ZIP, WOX, PLINC, KNOTTED-like homeodomain (KNOX), BLH/BELL homeodomain (BEL1-like) (Bürglin 2016). The KNOX and BEL1-like homeodomains possess an additional three amino acid residues, forming an extended loop that enables specific DNA-binding interactions. These domains form the three-amino-acid loop extension (*TALE*) superclass homeodomain, which plays a vital role in regulating various plant biological processes (Hamant and Pautot 2010).

The *TALE* superfamily genes have been extensively studied across various plant species; however, their structural features, physiological roles, and dynamic responses to abiotic stress remain insufficiently understood, particularly in *Cicer arietinum*. In this study, we conducted a comprehensive bioinformatic characterization of the *TALE* gene family in chickpea and investigated the expression profiles of *CaTALE* genes under drought stress using RNA-seq data and qPCR validation. The findings presented here provide the first integrated insight into the structural organization, evolutionary diversification, and stress-responsive regulation of *TALE* genes in chickpea, establishing a foundation for future functional studies and crop improvement efforts.

2 Materials and Methods

2.1 Identification of the *TALE* Genes in *C. arietinum*

To identify TALE proteins in *Cicer arietinum*, previously published TALE protein sequences from *Arabidopsis thaliana* and *Phaseolus vulgaris* were used as query sequences in a BLASTP search against the entire *C. arietinum* proteome available in the Phytozome v13 database (Goodstein et al. 2012). Candidate sequences were selected if their E-value was $\leq 1e^{-5}$ and sequence identity was $\geq 50\%$. Redundant sequences were then removed using the Decrease Redundancy Tool (https://web.expasy.org/decrease_redundancy/) and the remaining putative TALE proteins were examined for the presence of the TALE domain using the Simple Modular Architecture Research Tool (SMART) and InterPro databases (Letunic et al. 2021; Paysan-Lafosse et al. 2023).

The confirmed putative CaTALE proteins were subsequently used as queries in the NCBI-BLASTP searches for further characterization. Information about the *CaTALE* genes, including amino acid length and chromosomal coordinates was obtained from Phytozome v13 through Pfam domain searches (PF03791-KNOX2 domain, PF03790-KNOX1 domain, PF03789-ELK domain and PF05920-Homeobox KN domain) (Goodstein et al. 2012). Various physicochemical properties of the CaTALE proteins were calculated using the Expasy ProtParam tool (Gasteiger et al. 2003).

2.2 Phylogenetic Analysis

To elucidate the evolutionary relationships among CaTALE proteins, the full length TALE protein sequences from *A. thaliana* and *P. vulgaris* were aligned using ClustalW. A phylogenetic tree was subsequently constructed using the maximum likelihood (ML) method implemented in MEGA11 (Tamura et al. 2021). The ML method was applied with 1000 bootstrap replications, and the Poisson model was used for amino acid substitution.

The Maximum Likelihood approach was selected because Tamura et al. (2021) demonstrated it to be one of the most accurate and statistically robust methods for constructing phylogenetic trees based on peptide sequences. The resulting phylogenetic tree was visualized and color-coded using iTOL (Interactive Tree of Life) (Letunic and Bork 2021).

2.3 Evolutionary Duplication and Chromosomal Localization Analyses

The Phytozome v13 database was used to identify the chromosomal locations of *CaTALE* genes and the resulting data were visualized using TBtools. Gene duplication events and

orthologs of *CaTALE* genes were also analyzed through Phytozome v13 and TBtools. In addition, the Phytozome v13 database was used to examine the genetic relationships between *Cicer arietinum* and two other species, *Arabidopsis thaliana* and *Phaseolus vulgaris*.

Genomic annotation (GFF) and genome sequence (FASTA) files for the *TALE* gene families from these plant species were retrieved, and collinearity analysis was performed using the Text Merge tool in TBtools for MCScanX. The results were visualized using Multiple Synteny Plot function in TBtools. Non-synonymous/synonymous substitution (Ka/Ks) ratios were calculated according to the the Yang-Nielsen (2000) model using the KaKs_Calculator 2.0 algorithm (Chen et al. 2020).

2.4 Analysis of Cis-Regulatory Elements and miRNA Prediction for the *CaTALE* Genes

To investigate the cis-acting elements of the *CaTALE* genes, 1500 bp of genomic sequence upstream of the start codon (ATG) of each gene was retrieved from the Phytozome v13 database. PlantCARE was used to identify and annotate potential cis-acting elements within the promoter regions of the *CaTALE* genes (Lescot et al. 2002).

All known *C. arietinum* miRNA sequences were obtained from PmiREN database (Guo et al. 2022). Potential miRNA–target interactions were predicted using the psRNATarget web server (Dai et al. 2018) with an expectation threshold of $\leq 1e^{-10}$. The in silico–predicted miRNA–CaTALE interactions were visualized as a regulatory network using Cytoscape v3.9.1 (Shannon et al. 2003).

2.5 Plant Growth and Stress Treatment

Chickpea seeds of the Menemen-92 and İzmir-92 cultivars were obtained from the Aegean Agricultural Research Institute. The plant growth and stress treatment protocol was adapted from the methods described by Büyük et al. (2022).

First, the seeds were surface-sterilized and sown into pots containing equal amounts of soil. The plants were then germinated and grown under controlled conditions at 24 ± 2 °C with 70% relative humidity, a 16/8 h light/dark photoperiod, and a light intensity of 550–600 µmol m^{-2} s^{-1}. Plants were maintained in a climate-controlled growth chamber and irrigated with Hoagland nutrient solution every two days. Chickpea seedlings were grown for eight days until they reached the three-leaf stage. Afterward, the plants were divided into two groups: a control group and a drought-stress group. Watering was continued for the control plants, while irrigation was withheld for the drought-stress group. Pot weights were measured daily, and drought stress was imposed until the pot weight decreased by approximately 30% compared to the initial weight, corresponding to day 8, when visible drought symptoms such as leaf wilting were observed.

2.6 RNA Isolation, cDNA synthesis, and qPCR Analysis

Total RNA was extracted from harvested chickpea tissues using the Favorgen Total Plant RNA Isolation Mini Kit (Biotech Corp., Ping-Tung, Taiwan). The concentration and purity of the extracted RNA were assessed using a NanoDrop Lite spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). Complementary DNA (cDNA) was synthesized using the Transcriptor One-Step cDNA Synthesis Kit (Roche Diagnostics, Basel, Switzerland) following the manufacturer's instructions.

Gene-specific primers were designed using the PrimerQuest Tool. Quantitative PCR was performed using the iTaq Universal SYBR Green Supermix (Bio-Rad, Hercules, CA, USA) according to the protocol described by Büyük et al. (2022). The qPCR amplification program was as follows: initial denaturation at 95 °C for 30 s, followed by 40 cycles of denaturation at 95 °C for 15 s and annealing/extension at 60 °C for 30 s. A melting-curve analysis was subsequently performed from 60 °C to 95 °C to verify amplification specificity.

Each expression analysis included three biological replicates per treatment and tissue, and each qPCR reaction was performed in triplicate (technical replicates). Primer specificity was confirmed by both melt-curve analysis and agarose gel electrophoresis, while amplification efficiency (95–103%) was validated using standard curves with R^2 values greater than 0.99. The expression stability of the Actin reference gene across control and drought-stressed samples was evaluated using geNorm and NormFinder, both confirming its suitability ($M < 0.5$). All qPCR experiments were conducted in accordance with MIQE guidelines to ensure reproducibility and data reliability.

The *Actin* gene (NCBI Accession Code: EU529707) was used as the internal reference, and relative gene expression levels were calculated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen 2001). Statistical analyses were performed using GraphPad Prism v9.2. Two-way ANOVA followed by Fisher's least significant difference (LSD) test was applied, and differences were considered statistically significant at $p < 0.05$.

3 Results

3.1 Characteristics of the *CaTALE* Gene Family Members in *C. arietinum*

As a result of bioinformatic analyses performed on chickpea (*C. arietinum*) genome, 27 *CaTALE* genes were identified and named. These 27 *CaTALE* genes were distributed across eight different chromosomes. The main features of the *CaTALE* gene family and its members, including molecular weight (MW), protein length (aa, amino acid), isoelectric point (pI), conserved domains, chromosome locations, instability indices, NCBI accession numbers, were determined and are summarized in Supplementary Table 1.

The molecular weights of CaTALE proteins ranged from 39857.7 KDa (CaTALE-16) to 92046.79 KDa (CaTALE-21). The isoelectric point (pI) values varied between 4.71 (CaTALE- 23) to 9.04 (CaTALE-26), indicating that 19 of the CaTALE proteins were acidic, seven were basic and one was neutral. Based on the instability index, three of the 27 members had values below 40, whereas the remaining 24 proteins had values above 40. This suggests that the majority of the proteins encoded by *CaTALE* genes are predicted to be unstable. The protein lengths of the *CaTALE* gene family ranged from 123 amino acids (CaTALE-23) to 514 amino acids (CaTALE-21).

3.2 Phylogenetic Analysis of the CaTALE Proteins

The peptide sequences of 27 TALE proteins from chickpea (*Cicer arietinum*), 22 from *Arabidopsis thaliana*, and 35 from common bean (*Phaseolus vulgaris*) were aligned, and a phylogenetic tree was constructed using the TBtools software to examine the evolutionary relationships among these plant species. The resulting phylogenetic tree was color-coded using the Interactive Tree of Life (iTOL) platform.

A total of 84 TALE proteins from *C. arietinum*, *A. thaliana*, and *P. vulgaris* were grouped into two major subfamilies: KNOX and BELL1-like. Further analysis revealed that the KNOX subfamily could be divided into three clades: KNOX-I, KNOX-II, and KNOX-III. In chickpea, 12 TALE proteins belonged to the KNOX subfamily, compared to 16 in bean and 9 in *A. thaliana*. Similarly, 15 CaTALE proteins were classified as BELL1-like, along with 19 in bean and 13 in *A. thaliana*. The phylogenetic relationships among the 84 TALE proteins from these species are illustrated in Fig. 1.

3.3 Gene Structure and Conserved Protein Motifs of *CaTALE* Genes

The protein sequences of *CaTALE* genes were retrieved from the Phytozome v13 database (https://phytozome-next.jgi.doe.gov/). Conserved motifs within CaTALE proteins were identified using the MEME online analysis tool (https://meme-suite.org/meme/tools/meme) (Bailey et al. 2009). The analysis revealed that members of the *CaTALE* family, categorized into subfamilies as shown in Fig. 2, exhibited motif similarities within each family/subfamily. Specifically, the BEL1-like family contained motifs 1, 2, 3, 4, 5, 6, 7, and 10, with motif 1 being conserved across all proteins.

Among all classes of the *KNOX* gene family, motif 1 was identified as a common feature shared by KNOX1, KNOX2, and KNOX3 members. Motif 4 was found to be present in all members of the KNOX1 class. In the KNOX2 class, no single motif was universally conserved; however, motifs 1, 2, 3, 4, 5, 6, and 7 were detected. Domain analysis showed that the KNOX1 and KNOX2 domains are conserved in all KNOX1 class members within the *KNOX* gene family. Additionally, the PBX, ELK, and HOMEBOX domains were identified in both the KNOX1 and KNOX2 classes.

Most members of the BEL1-like gene family were characterized by the presence of PBX domains, with additional HOMEBOX domains also identified. The gene structure, functional attributes, and evolutionary relationships of the *CaTALE* gene family were analyzed using the Gene Structure Display Server (GSDS) database. Subsequently, data from MEME v5.4.1, Motif Finder, the NCBI Conserved Domain Database (CDD), the JGI Phytozome v13, and GSDS 2.0 were processed and visualized using TBtools.

3.4 Chromosomal Distribution of *CaTALE* Genes

The chromosomal distribution of *CaTALE* genes, visualized using the TBtools software, is presented in Fig. 3. The nomenclature of *TALE* genes in chickpea was designated as "*CaTALE*," derived from the initial letters of the Latin name of the species (*Cicer arietinum*) and the abbreviation of the *TALE* gene family. The *CaTALE* genes were numbered according to their sequential arrangement on the chromosomes.

It was determined that *CaTALE* genes are distributed across eight chromosomes of *C. arietinum*. Among these, chromosome 4 (Chr04) harbors the highest number of *CaTALE* genes (seven), whereas chromosomes 1, 3, and 10 (Chr01, Chr03, and Chr10) each contain only a single *CaTALE* gene.

3.5 Analysis of Duplication Events and Synteny Within the *CaTALE* Gene Family

Gene duplication plays a critical role in increasing gene numbers and promoting functional diversification. These duplications are generally categorized into two types: tandem duplications and segmental duplications. Segmental

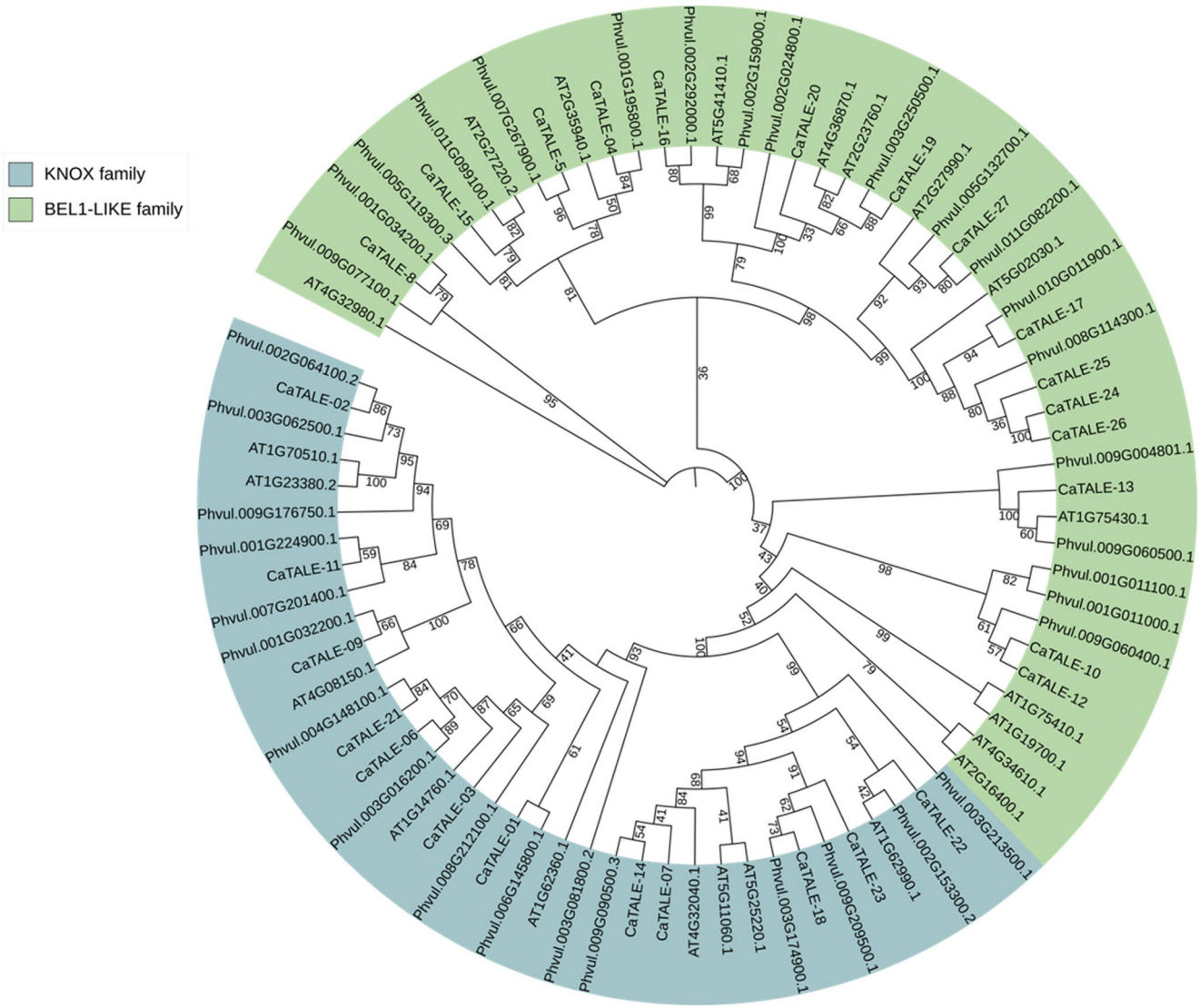

Fig. 1 Phylogenetic tree constructed using the peptide sequences of 84 TALE proteins from *Cicer arietinum*, *Phaseolus vulgaris*, and *Arabidopsis thaliana*

duplication refers to the duplication of a gene sequence that is subsequently transferred to a different chromosomal region, whereas tandem duplication occurs when the duplicated gene remains adjacent to the original locus on the same chromosome.

Analysis of *CaTALE* genes revealed seven segmental duplication events but no tandem duplications. Specifically, seven segmentally duplicated gene pairs were identified: *CaTALE-08*/*CaTALE-14*, *CaTALE-20*/*CaTALE-22*, *CaTALE-14*/*CaTALE-19*, *CaTALE-06*/*CaTALE-17*, *CaTALE-04*/*CaTALE-17*, *CaTALE-06*/*CaTALE-04*, and *CaTALE-11*/*CaTALE-12*. The duplication results were processed using TBtools and visualized in a Circos plot (Figs. 4, 5, 6).

Genes that are present in different species but share a common ancestral origin and retain structural and functional similarities are referred to as orthologous genes. The term synteny was originally defined as homologous genes located on the same chromosome; however, it is now broadly used to describe orthologous genes across species that maintain conserved genomic arrangements based on sequence similarity rather than function.

Syntenic relationships between chickpea (*Cicer arietinum*), *Arabidopsis thaliana*, and common bean (*Phaseolus vulgaris*) were analyzed. The results indicated that the orthologous relationships between chickpea and common bean were considerably stronger than those between chickpea and *A. thaliana*.

To investigate the evolutionary dynamics of these genes, the nonsynonymous substitution rate (Ka), synonymous substitution rate (Ks), and Ka/Ks ratio were calculated for *TALE* orthologous gene pairs in chickpea, common bean,

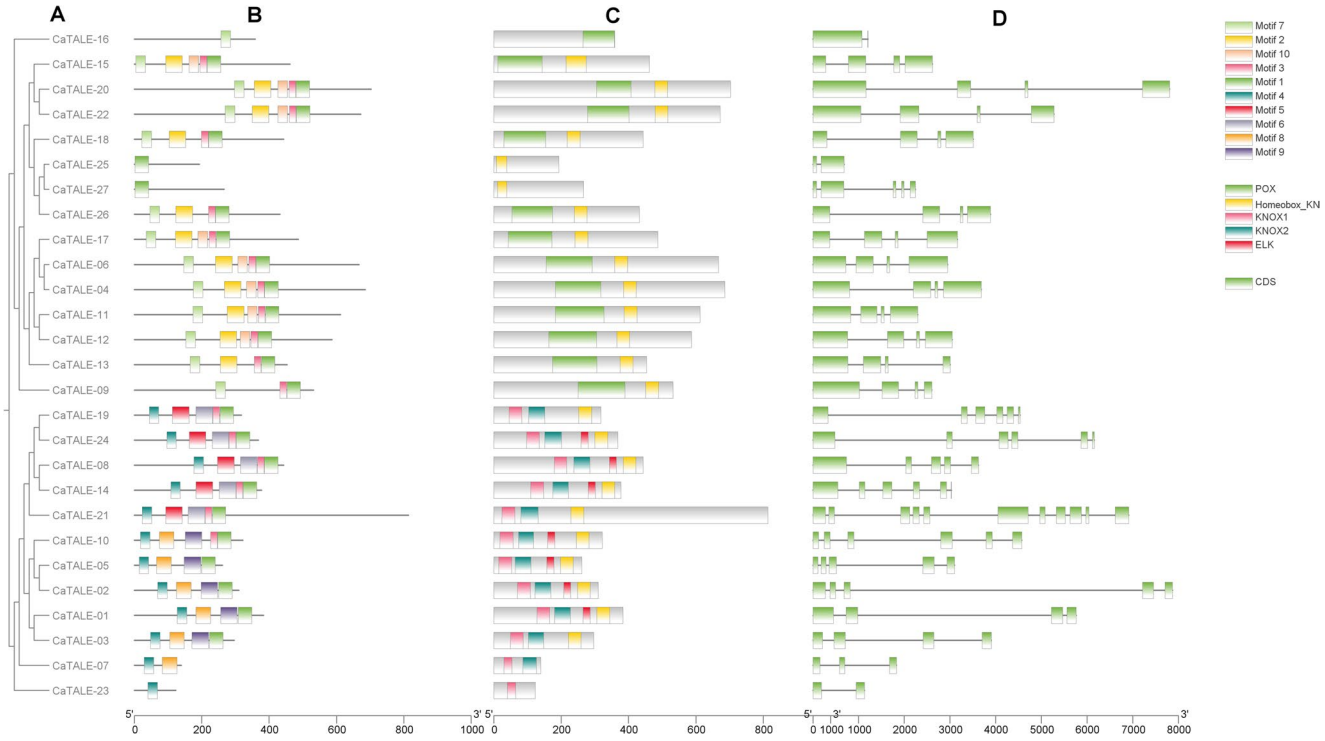

Fig. 2 Conserved motifs, protein domains, and exon–intron structures of *TALE* genes in *Cicer arietinum*. (**A**) Phylogenetic tree of TALE proteins in *C. arietinum*. (**B**) Conserved motifs of TALE proteins. Each TALE protein contains distinct motif arrangements represented by different colors, and the order of motifs corresponds to their positions within the respective protein sequences. (**C**) Schematic representation of TALE protein domains. (**D**) Exon–intron organization of *TALE* genes

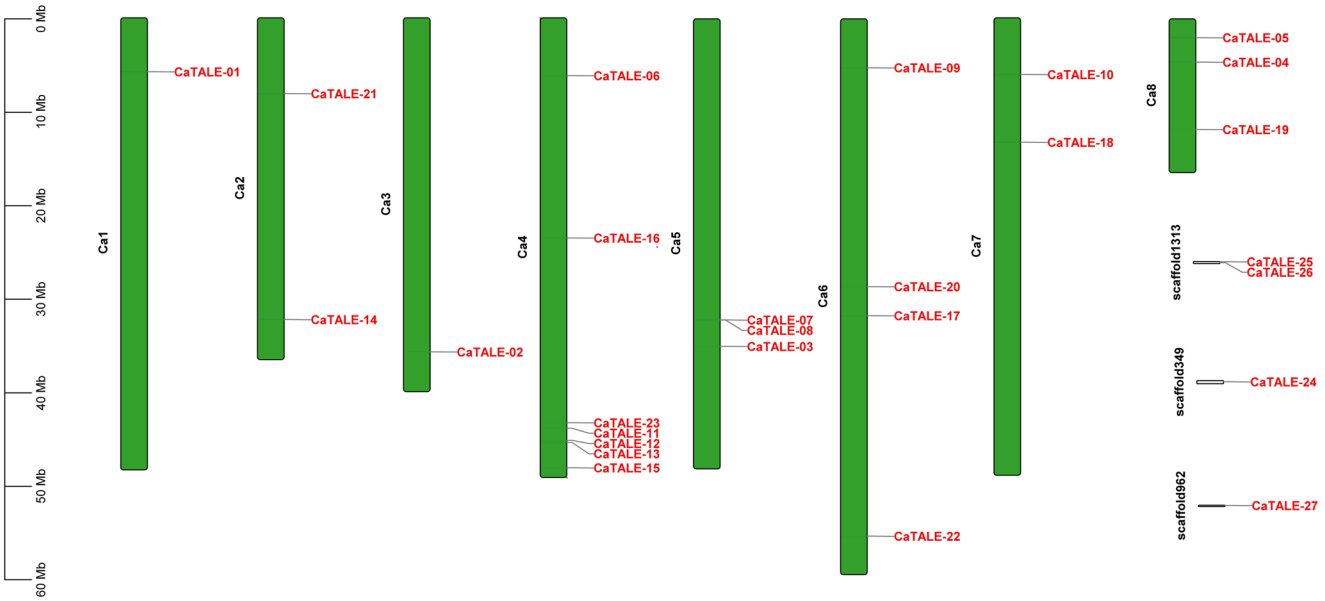

Fig. 3 Chromosomal distribution of *CaTALE* genes across the eight chromosomes of *Cicer arietinum*. The scale is shown in megabases (Mb)

and *A. thaliana* using TBtools. The results are presented in Supplementary Tables 2 and 3. Instances with no valid results were recorded as NaN. The Ka/Ks values were interpreted as follows: (1) Ka/Ks < 1: indicates strong purifying (negative) selection, suggesting that TALE genes are under evolutionary constraint, (2) Ka/Ks > 1: suggests positive selection, implying adaptive evolution, (3) Ka/Ks ≈ 1: reflects neutral evolution, meaning that mutations accumulate randomly without selective pressure.

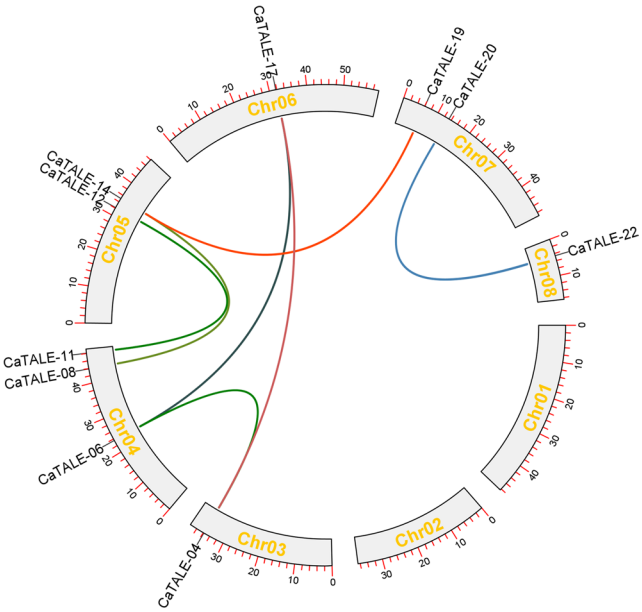

Fig. 4 Circos plot illustrating the segmental duplication relationships among *CaTALE* genes in *Cicer arietinum*

3.6 Analysis of Cis-Elements in the *CaTALE* Genes

Cis-elements play an essential role in the transcriptional regulation of gene expression. To identify the cis-regulatory elements in the promoter regions of *CaTALE* genes, the PlantCARE database was used, and the results were visualized using TBtools. A total of 103 cis-elements were identified across all *CaTALE* genes.

The promoter regions of the *CaTALE* gene family were analyzed using 1500 bp upstream sequences of each gene to identify cis-regulatory elements. The analysis revealed cis-elements associated with gibberellin sensitivity, photosensitivity, abscisic acid (ABA) responsiveness, drought inducibility, auxin sensitivity, defense and stress responses, seed-specific regulation, methyl jasmonate (MeJA) responsiveness, circadian regulation, anaerobic induction, and salicylic acid (SA) sensitivity.

These findings imply that the *CaTALE* gene family may participate in various biological processes, particularly in abiotic stress responses and plant hormone regulation. A schematic representation of the identified cis-elements in the promoter regions is presented in Fig. 7.

3.7 MiRNA Analysis of the *CaTALE* Gene Family

MicroRNAs (miRNAs) are short non-coding RNAs, typically 20–24 nucleotides in length, that play a crucial role in gene regulation in both plants and animals. They regulate gene expression either by promoting the post-transcriptional degredation of target mRNA or by inhibiting their translation.

In this study, miRNAs targeting *TALE* genes were identified in chickpea (*Cicer arietinum*). Chickpea-specific miRNA sequences were retrieved from the PmiREN online database, and potential miRNA–*CaTALE* gene interactions were predicted using the psRNATarget online tool. The interactions between miRNAs and *CaTALE* genes were visualized using Cytoscape v3.9.1 (Fig. 8). In the network visualization, turquoise-colored boxes represent miRNAs, whereas yellow-colored boxes denote *CaTALE* genes.

These results provide valuable insights into the post-transcriptional regulatory network of the *CaTALE* gene family in chickpea.

3.8 RT-qPCR Analysis of *CaTALE* Genes During Drought Stress

Representative *CaTALE* genes were selected for RT-qPCR assays based on the following criteria: (1) genes exhibiting relatively high expression levels in the analyzed plant tissues, as identified through RNA-seq data, and representing

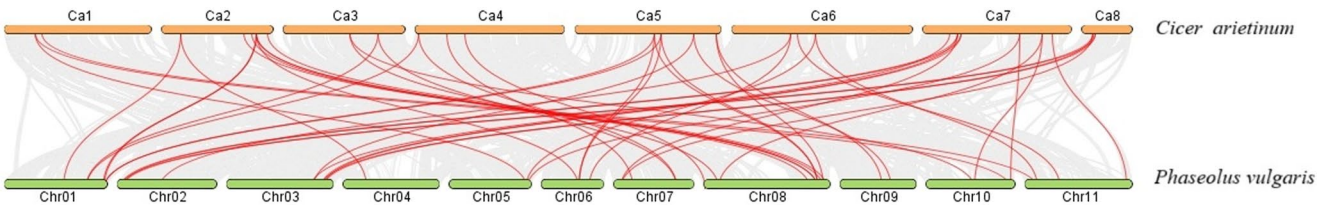

Fig. 5 Synteny analysis of *CaTALE* genes between *Cicer arietinum* and *Phaseolus vulgaris*

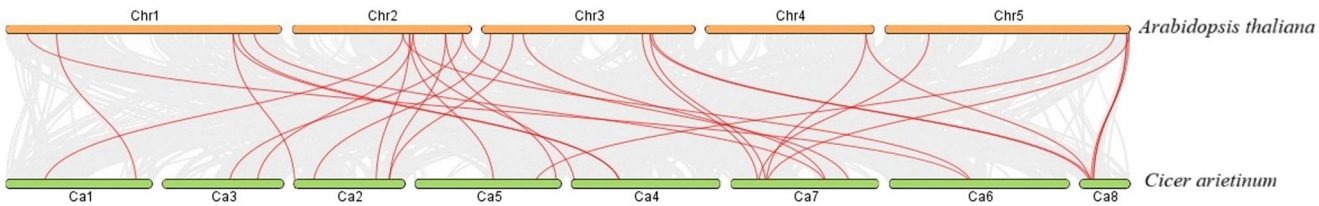

Fig. 6 Synteny analysis of *CaTALE* genes between *Cicer arietinum* and *Arabidopsis thaliana*

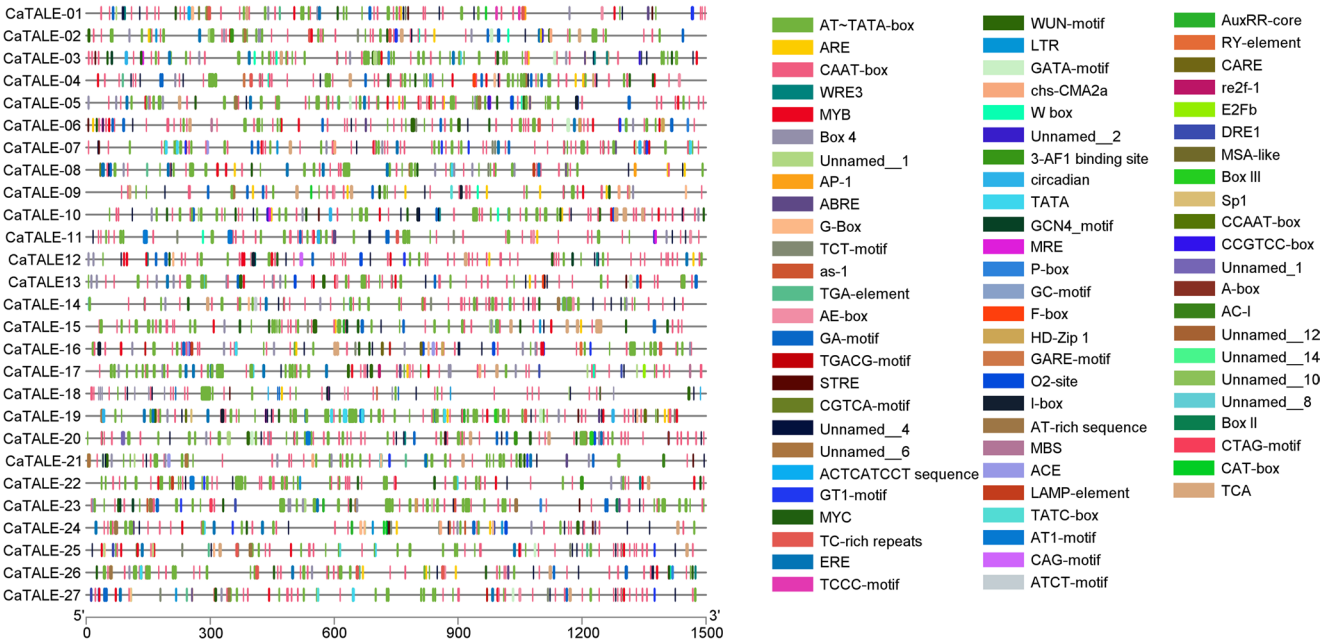

Fig. 7 Distribution of predicted cis-acting regulatory elements within the 1.5 kb upstream regions of *CaTALE* genes in *Cicer arietinum*. The colored boxes represent different cis-elements as indicated in the legend on the right

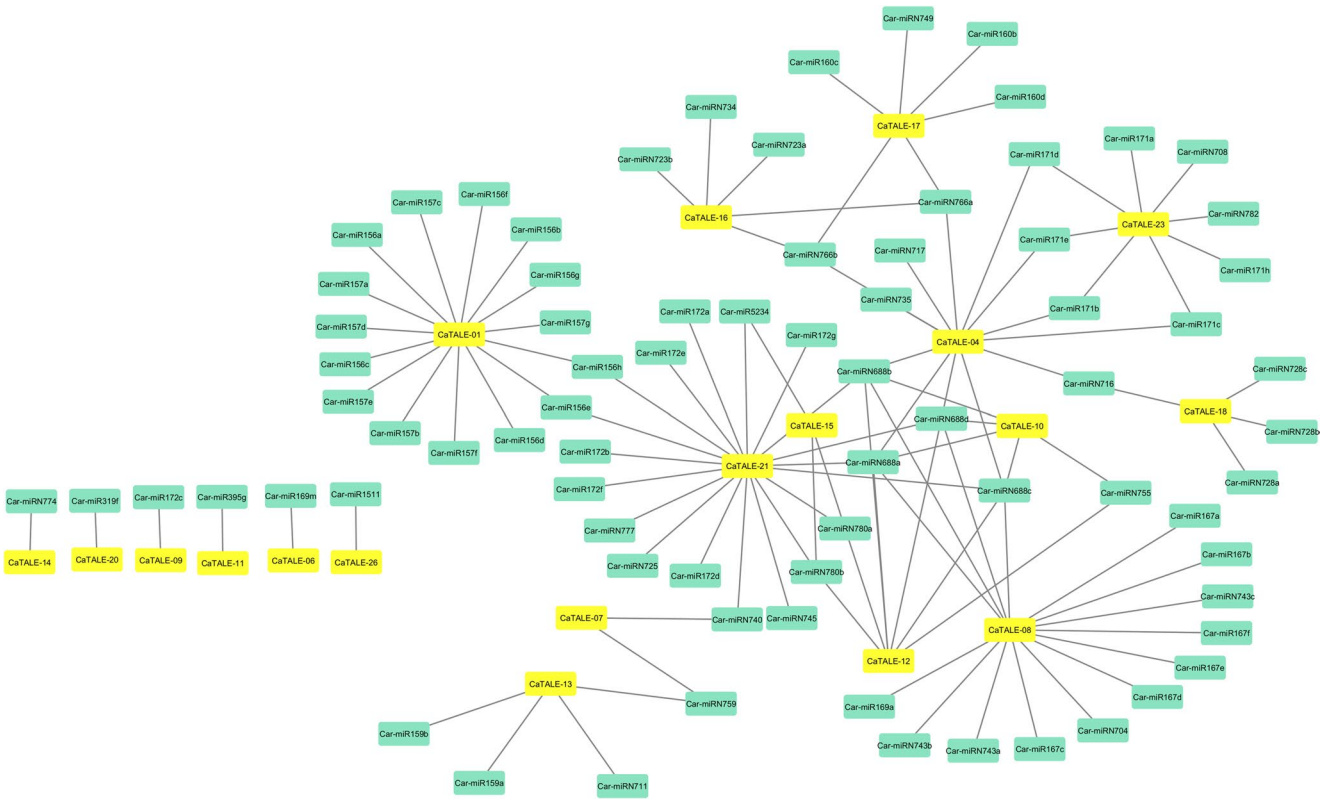

Fig. 8 Interaction network between *CaTALE* gene family members and their associated miRNAs in *Cicer arietinum*. Turquoise-colored boxes represent miRNAs, while yellow-colored boxes represent *CaTALE* genes

various *CaTALE* subfamilies or subgroups; (2) genes showing differential expression in response to drought stress, as indicated by RNA-seq analysis; and (3) genes containing cis-elements previously reported in the literature to be associated with stress responses.

Four of the selected genes belong to the KNOX subfamily (*CaTALE*−01, *CaTALE*−10, *CaTALE*−18, and *CaTALE*−03), while five belong to the BEL1-like subfamily (*CaTALE*−02, *CaTALE*−09, *CaTALE*−13, *CaTALE*−21, and *CaTALE*−27). Primers were designed for these genes (Supplementary Table 4), and mRNA expression levels were analyzed via RT-qPCR using cDNAs synthesized from RNA extracted from leaf and root tissues subjected to drought stress (Fig. 9).

According to the results, in the *İzmir-92* cultivar, an increase in the mRNA expression levels of *CaTALE*−01, *CaTALE*−02, *CaTALE*−09, and *CaTALE*−10 was observed in leaf tissues under drought conditions. No significant changes were detected in the expression of other genes in leaves in response to drought stress. In root tissues, the highest expression changes under drought stress were also recorded for *CaTALE*−01, *CaTALE*−02, *CaTALE*−09, and *CaTALE*−10, with *CaTALE*−10 showing the greatest increase among leaf samples (Fig. 10).

In the *Menemen-92* cultivar, expression changes in leaf tissues under drought stress were detected for *CaTALE*−01, *CaTALE*−02, *CaTALE*−03, *CaTALE*−10, *CaTALE*−18, *CaTALE*−21, and *CaTALE-27*, whereas no increase or decrease was observed for *CaTALE-09* and *CaTALE-13*. In root tissues of *Menemen-92*, a significant upregulation was found for *CaTALE*−03, *CaTALE*−10, and *CaTALE*−18, while *CaTALE*−02 showed no noticeable change in expression. Overall, these findings indicate that *CaTALE*−01, *CaTALE*−10, *CaTALE*−18, *CaTALE*−21, and *CaTALE*−27 exhibit cultivar- and tissue-specific expression patterns in response to drought stress.

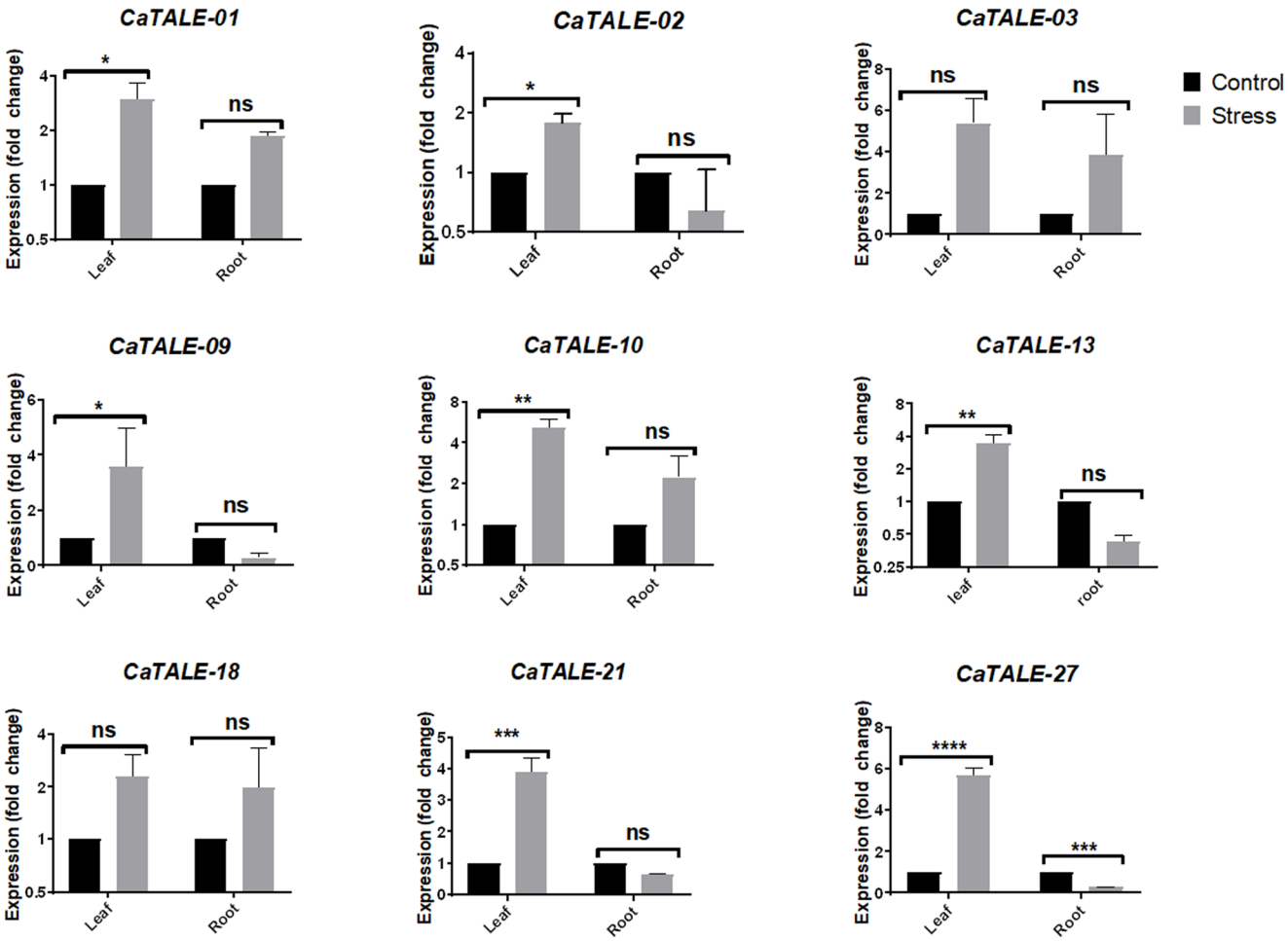

Fig. 9 The mRNA expression changes of certain *CaTALE* genes in response to drought stress in the Izmir-92 chickpea cultivars are shown. The resulting graphics were normalized with the *CaActin* gene. Vertical bars in the graphs represent the standard deviation. Asterisks in the graphs indicate the level of statistical significance (*$P<0.0332$, **$P<0.0021$, ***$P<0.0002$, ****$P<0.0001$ for the respective genes)

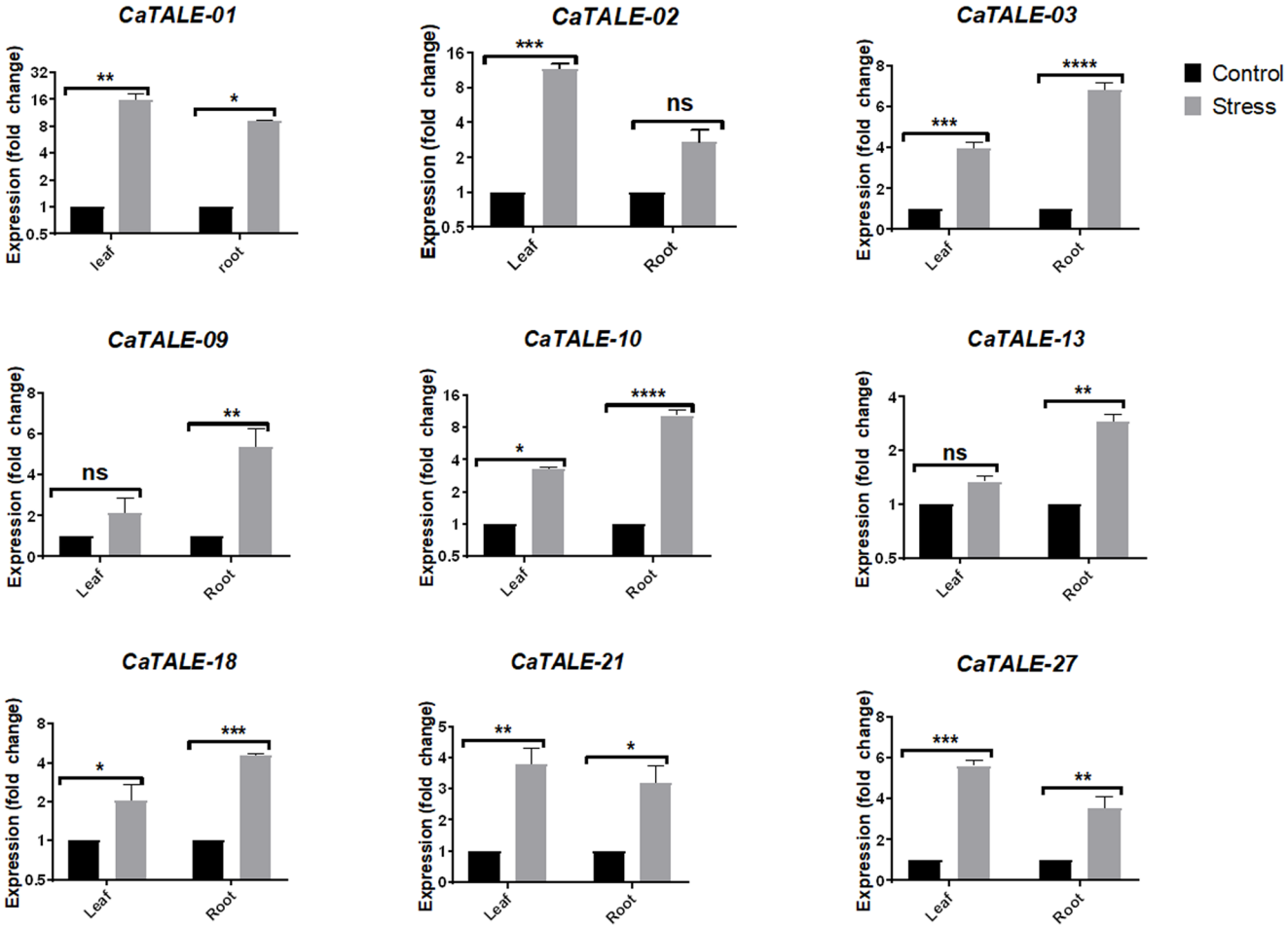

Fig. 10 The mRNA expression changes of certain *CaTALE* genes in response to drought stress in the Menemen-92 chickpea cultivars are shown. The resulting graphics were normalized with the *CaActin* gene. Vertical bars in the graphs represent the standard deviation. Asterisks in the graphs indicate the level of statistical significance (*$P<0.0332$, **$P<0.0021$, ***$P<0.0002$, ****$P<0.0001$ for the respective genes)

4 Discussion

The *TALE* gene family, widely distributed across eukaryotic organisms, has been a major focus of research since the early 1900s. Members of this superfamily are abundant in plant genomes and play crucial roles in the regulation of growth, development, and environmental stress responses. Although genome-wide analyses of *TALE* genes have been reported in several plant species, such as *Arabidopsis thaliana* (Hamant and Pautot 2010), poplar (Zhao et al. 2019), cotton (Ma et al. 2019), pomegranate (Wang et al. 2020), and soybean (Wang et al. 2021), a comprehensive study in chickpea (*Cicer arietinum*) has not yet been performed. Therefore, the present study aimed to systematically identify and characterize the *TALE* gene family in chickpea, investigate their evolutionary relationships, and examine their expression profiles under drought stress.

In this study, 27 *CaTALE* genes were identified in the chickpea genome. The number of *TALE* genes differs among plant species, ranging from 18 in pineapple (*Ananas comosus*) (Ali et al. 2019) to 94 in cotton (*Gossypium hirsutum*) (Ma et al. 2019). This variation suggests that gene family expansion is associated with genome size and ploidy level (Ma et al. 2019). The relatively smaller number of *TALE* genes in diploid chickpea may therefore reflect its simpler genome organization compared with polyploid species.

Most CaTALE proteins were predicted to localize in the nucleus, consistent with previous findings in other plant species such as citrus (Peng et al. 2022) and barley (Liao et al. 2024). Only two proteins, CaTALE-06 and CaTALE-22, were predicted to localize in the mitochondria and Golgi apparatus, respectively. This high level of nuclear localization across plant species indicates that TALE proteins are evolutionarily conserved as nuclear transcription factors.

The chromosomal mapping of *CaTALE* genes revealed that eight genes were distributed across eight chromosomes, while four genes could not be assigned to specific chromosomes. This pattern is comparable to that observed in other species, including soybean (Wang et al. 2021) and

strawberry (Kabak et al. 2021). Such distribution may result from incomplete genome annotation, chromosomal rearrangements, or structural variations, which could contribute to the diversification and functional evolution of *TALE* genes. The dispersed pattern also suggests that *TALE* genes may have undergone redistribution throughout the genome during evolution, thereby facilitating their functional diversification.

Phylogenetic analysis classified the 27 *CaTALE* proteins into two major subfamilies: KNOX (12 members) and BELL1-like (15 members). This distribution aligns with previous studies in *Arabidopsis thaliana*, *Phaseolus vulgaris*, and *Oryza sativa* (Avcı et al. 2024), supporting the evolutionary conservation of these subfamilies. Each subfamily displayed distinct domain and motif compositions, which are critical for DNA binding and protein–protein interactions. For instance, POX domains unique to plant proteins are frequently associated with homeobox domains (Liu et al. 1999). BEL1-like proteins containing POX domains are known to interact with KNAT2 and KNAT5 proteins, influencing plant development. In chickpea, 10 out of 15 POX domain-containing members belong to the BEL1 subfamily. Similarly, the ELK domain, consisting of conserved glutamic acid, leucine, and lysine residues, may serve as a nuclear localization signal or participate in protein interactions (Nagasaki et al. 2001; Ito et al. 2002; Kerstetter et al. 1994). Seven *CaTALE* genes contain the ELK domain, while the KNOX1 and KNOX2 domains, together forming the MEINOX region, play roles in transcriptional repression and homodimerization (Nagasaki et al. 2001). The motif analysis showed that certain motifs, such as Motif 1, were highly conserved across all KNOX classes, while others were specific to either the KNOX or BEL1 subfamilies, reflecting subfamily-specific functional divergence (Guo et al. 2022; Peng et al. 2022).

Gene duplication analysis revealed seven segmental duplication events among *CaTALE* genes, with no evidence of tandem duplications. This pattern is consistent with findings in other species, including bean (Avcı et al. 2024), soybean (Wang et al. 2021), and walnut (Guo et al. 2022), where segmental duplication is a dominant mechanism driving gene family expansion. Interestingly, six of the seven duplicated gene pairs consisted of members from different subfamilies, while only one pair (CaTALE4–17) belonged to the same subfamily. These results suggest that segmental duplications have contributed to the functional diversification of the *CaTALE* family and may play a role in the adaptive evolution of chickpea (Zhu et al. 2014; Du et al. 2023).

Promoter region analysis identified 103 cis-regulatory elements across *CaTALE* genes, including motifs associated with light response, hormonal regulation, and environmental stress. The majority of these elements were related to stress responses (61.39%), followed by hormone-responsive elements (20.64%). This pattern suggests that *CaTALE* genes may participate in stress adaptation and developmental regulation. Previous studies have reported that members of the TALE family are involved in plant responses to heat, hormonal signals, drought, and salinity (Zhao et al. 2019; Peng et al. 2022; Wang et al. 2021; Han et al. 2022; Rathour et al. 2022; Yang et al. 2022; Liao et al. 2024).

The regulatory potential of *CaTALE* genes was further supported by predicted interactions with miRNAs. In particular, miRNA156a and miRNA156b were found to target *CaTALE-01* and *CaTALE-21*, while miRNA171 was predicted to regulate *CaTALE-04* and *CaTALE-23*. miRNA156 is known to modulate developmental transitions and floral traits by targeting SQUAMOSA promoter-binding protein-like (SPL) genes (Wang et al. 2019), whereas miR171 regulates vegetative-to-reproductive phase transitions through the miR156 pathway (Curaba et al. 2013). These results imply that *CaTALE* genes are integrated into complex transcriptional and post-transcriptional regulatory networks that coordinate development and stress responses.

Expression analysis under drought conditions revealed that several *CaTALE* genes were differentially expressed in chickpea tissues, particularly in roots. RNA-seq and qRT-PCR analyses in two chickpea cultivars (İzmir and Menemen-92) demonstrated increased expression of *CaTALE* genes in response to drought, with cultivar-specific differences. Among the examined genes, *CaTALE-10*, *CaTALE-13*, and *CaTALE-27* (BEL1-like subfamily) and *CaTALE-1*, *CaTALE-3*, *CaTALE-9*, *CaTALE-18*, and *CaTALE-21* (KNOX subfamily) exhibited pronounced induction. These findings are consistent with reports in other species linking *TALE* genes to abiotic stress tolerance and hormone signaling. For example, overexpression of *TaKNOX11-A* in wheat enhanced drought and salt tolerance (Han et al. 2022), and *PagKNAT2/6b* in poplar suppressed gibberellin biosynthesis to improve drought resistance (Song et al. 2021). Similarly, transcriptome studies in *Camellia japonica* and *Dendrobium huoshanense* have identified several *KNOX* genes as stress-responsive (Dai et al. 2023; Li et al. 2023). In maize, *ZmTALE1*, *ZmTALE37*, and *ZmTALE38*—homologs of *AtKNAT7*—regulate lignin biosynthesis, a key process in drought tolerance (Qin et al. 2020; Yan et al. 2018).

5 Conclusion

This study provides the first comprehensive genome-wide characterization of the TALE transcription factor family in chickpea (*Cicer arietinum*), identifying 27 *CaTALE* genes grouped into KNOX and BEL1-like subfamilies. Comparative and duplication analyses revealed that segmental

duplication, rather than tandem duplication, played a dominant role in shaping the expansion and diversification of this family. Promoter and expression analyses indicated that several *CaTALE* genes, particularly *CaTALE-01*, *CaTALE-10*, *CaTALE-18*, *CaTALE-21*, and *CaTALE-27*, are transcriptionally activated under drought stress and may function as key regulators of hormonal signaling and stress adaptation.

These genes represent promising candidates for functional validation and targeted manipulation in chickpea breeding programs aimed at enhancing drought resilience. Overall, the structural, regulatory, and expression insights obtained here lay the groundwork for elucidating the molecular mechanisms by which TALE transcription factors modulate developmental and stress-related pathways in legumes.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42729-025-02964-1.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by SC and AO. The final draft of the manuscript was edited by IB and SA. All authors read and approved the final manuscript.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK).

Declarations

Ethical Approval Not applicable.

Consent To Participate Informed consent was obtained from all individual participants included in the study.

Competing interests The authors declare that there are no conflicts of interest. All authors have read, reviewed, and agreed to submit the review article to JOURNAL OF PLANT NUTRITION AND SOIL SCIENCE.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Ali H, Liu Y, Azam SM (2019) Genome wide identification and expression profiles of *TALE* genes in pineapple (*Ananas comosus* L). Trop Plant Biol 12:304–317. https://doi.org/10.1007/s12042-019-09232-4

Avcı FN, Kırlıoğlu T, Okay A, Aras ES, Büyük İ (2024) In silico analysis of three-amino acid loop extension (TALE) proteins in *Phaseolus vulgaris* L. and investigation of their transcript levels in response to drought stress. Turk J Agric For 48:911–926. https://doi.org/10.55730/1300-011X.3230

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37:202–208. https://doi.org/10.1093/nar/gkp335

Bürglin TR (2016) Homeodomain subtypes and functional diversity. Subcell Biochem 116:95–122. https://doi.org/10.1007/978-3-319-30550-8_5

Büyük İ, Okay A, Aras S (2022) Identification and characterization of SRS genes in *phaseolus vulgaris* genome and their responses under salt stress. Biochem Genet 60:482–503. https://doi.org/10.1007/s10528-021-10108-0

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R (2020) TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant 13:1194–1202. https://doi.org/10.1016/j.molp.2020.06.009

Curaba J, Talbot M, Li Z, Helliwell C (2013) Over-expression of microRNA171 affects phase transitions and floral meristem determinancy in barley. BMC Plant Biol. https://doi.org/10.1186/1471-2229-13-6

Dai H, Zheng S, Zhang C, Huang R, Yuan L, Tong H (2023) Identification and expression analysis of the *KNOX* genes during organogenesis and stress responseness in *Camellia sinensis* (L.) O. Kuntze. Mol Genet Genomics 298:1559–1578. https://doi.org/10.1007/s00438-023-02075-6

Dai X, Zhuang Z, Zhao PX (2018) psRNATarget: a plant small RNA target analysis server. Nucleic Acids Res 46:49–54. https://doi.org/10.1093/nar/gky316

Du L, Ma Z, Mao H (2023) Duplicate Genes Contribute to Variability in Abiotic Stress Resistance in Allopolyploid Wheat. Plants 12:2465. https://doi.org/10.3390/plants12132465

Filtz TM, Vogel WK, Leid M (2014) Regulation of transcription factor activity by interconnected post-translational modifications. Trends Pharmacol Sci 35:76–85. https://doi.org/10.1016/j.tips.2013.11.005

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31:3784–3788. https://doi.org/10.1093/nar/gkg563

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 40:1178–1186. https://doi.org/10.1093/nar/gkr944

Guo C, Quan S, Zhang Z, Kang C, Liu J (2022) Genome-wide identification, characterization and expression profile of *TALE* gene family in (*Juglans regia* L.). Sci Hortic 297:110945. https://doi.org/10.1016/j.scienta.2022.110945

Hamant O, Pautot V (2010) Plant development: a TALE story. CR Biol 333:371–381. https://doi.org/10.1016/j.crvi.2010.01.015

Han Y, Zhang L, Yan L, Xiong X, Wang W (2022) Genomewide analysis of TALE superfamily in *Triticum aestivum* reveals *TaKNOX11-A* is involved in abiotic stress response. BMC Genomics 23:1–21. https://doi.org/10.1186/s12864022-08324-y

Ito Y, Hirochika H, Kurata N (2002) Organ-specific alternative transcripts of KNOX family class 2 homeobox genes of rice. Gene 288:41–47. https://doi.org/10.1016/s0378-1119(02)00460-2

Kabak G, Şehsuvar S, Turgut S, Gökdemir Ş (2021) Genome-wide analysis of Fragaria vesca three-amino-acid-loop-extension

(*TALE*) genes. Biotech Stud 30:79–85. https://doi.org/10.38042/biotechstudies.977788

Kerstetter R, Vollbrecht E, Lowe B, Veit B, Yamaguchi J, Hake S (1994) Sequence analysis and expression patterns divide the maize knotted1-like homeobox genes into two classes. Plant Cell 6:1877–1887. https://doi.org/10.1105/tpc.6.12.1877

Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR, Weirauch MT (2018) The human transcription factors. Cell 172:650–665. https://doi.org/10.1016/j.cell.2018.01.029

Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, de Van Peer Y, Rouzé P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res 30:325–327. https://doi.org/10.1093/nar/30.1.325

Letunic I, Bork P (2021) Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res 49:293–296. https://doi.org/10.1093/nar/gkab301

Letunic I, Khedkar S, Bork P (2021) SMART: recent updates, new developments and status in 2020. Nucleic Acids Res 49:458–460. https://doi.org/10.1093/nar/gkaa937

Li G, Manzoor MA, Wang G, Chen C, Song C (2023) Comparative analysis of *KNOX* genes and their expression patterns under various treatments in *Dendrobium huoshanense*. Front Plant Sci 14:1258533. https://doi.org/10.3389/fpls.2023.1258533

Liao TJ, Huang T, Xiong HY, Duo JC, Ma JZ, Du MY, Duan RJ (2024) Genome-wide identification, characterization, and evolutionary analysis of the barley TALE gene family and its expression profiles in response to exogenous hormones. Front Plant Sci 15:1421702. https://doi.org/10.3389/fpls.2024.1421702

Liu L, White MJ, MacRae TH (1999) Transcription factors and their genes in higher plants functional domains, evolution and regulation. Eur J Biochem 262:247–257. https://doi.org/10.1046/j.1432-1327.1999.00349.x

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2<Superscript>−ΔΔCT</Superscript> method. Methods 25:402–408. https://doi.org/10.1006/meth.2001.1262

Ma Q, Wang N, Hao P, Sun H, Wang C (2019) Genome-wide identification and characterization of TALE superfamily genes in cotton reveals their functions in regulating secondary cell wall biosynthesis. BMC Plant Biol 19:1–20. https://doi.org/10.1186/s12870-019-2026-1

Nagasaki H, Sakamoto T, Sato Y, Matsuoka M (2001) Functional analysis of the conserved domains of a rice KNOX homeodomain protein, OSH15. Plant Cell 13:2085–2098. https://doi.org/10.1105/tpc.010113

Paysan-Lafosse T, Blum M, Chuguransky S, Grego T, Pinto BL, Salazar GA, Bileschi ML, Bork P, Bridge A, Colwell L, Gough J, Haft DH, Letunić I, Marchler-Bauer A, Mi H, Natale DA, Orengo CA, Pandurangan AP, Rivoire C, Sigrist CJA, Bateman A (2023) InterPro in 2022. Nucleic Acids Res 51:418–427. https://doi.org/10.1093/nar/gkac993

Peng W, Yang Y, Xu J, Peng E, Dai S (2022) TALE transcription factors in sweet orange (*Citrus sinensis*): genome-wide identification, characterization, and expression in response to biotic and abiotic stresses. Front Plant Sci 12:3215. https://doi.org/10.3389/fpls.2021.814252

Qin W, Yin Q, Chen J, Zhao X, Yue F, He J, Yang L, Liu L, Zeng Q, Lu F, Mitsuda N, Ohme-Takagi M, Wu AM (2020) The class II KNOX transcription factors KNAT3 and KNAT7 synergistically regulate monolignol biosynthesis in Arabidopsis. J Exp Bot 71:5469–5483. https://doi.org/10.1093/jxb/eraa266

Quiring R, Walldorf U, Kloter U, Gehring WJ (1994) Homology of the eyeless gene of drosophila to the small eye gene in mice and aniridia in humans. Science 265:785–789. https://doi.org/10.1126/science.7914031

Rasheed A, Gill RA, Hassan MU, Mahmood A, Qari S, Zaman QU, Ilyas M, Aamer M, Batool M, Li H, Wu Z (2021) A critical review: recent advancements in the use of CRISPR/Cas9 technology to enhance crops and alleviate global food crises. Curr Issues Mol Biol 43:1950–1976. https://doi.org/10.3390/cimb43030135

Rathour M, Shumayla AA, Upadhyay SK (2022) Investigation of roles of *TaTALE* genes during development and stress response in bread wheat. Plants 11:587. https://doi.org/10.3390/plants11050587

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504. https://doi.org/10.1101/gr.1239303

Song X, Zhao Y, Wang J, Lu MZ (2021) The transcription factor KNAT2/6b mediates changes in plant architecture in response to drought via down-regulating GA20ox1 in *Populus alba* × *P. glandulosa*. J Exp Bot 72:5625–5637. https://doi.org/10.1093/jxb/erab201

Tamura K, Stecher G, Kumar S (2021) MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol 38:3022–3027. https://doi.org/10.1093/molbev/msab120

Wang C, Wang Q, Zhu X, Cui M, Jia H, Zhang W, Tang W, Leng X, Shen W (2019) Characterization on the conservation and diversification of miRNA156 gene family from lower to higher plant species based on phylogenetic analysis at the whole genomic level. Funct Integr Genomics 19:933–952. https://doi.org/10.1007/s10142-019-00679-y

Wang L, Yang X, Gao Y, Yang S (2021) Genome-wide identification and characterization of TALE superfamily genes in soybean (*Glycine max* L). Int J Mol Sci 22:4117. https://doi.org/10.3390/ijms22084117

Wang W, Vinocur B, Shoseyov O, Altman A (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci 9:244–252. https://doi.org/10.1016/j.tplants.2004.03.006

Wang Y, Zhao Y, Yan M, Zhao H, Zhang X, Yuan Z (2020) Genome-wide identification and expression analysis of *TALE* gene family in Pomegranate (*Punica granatum* L.). Agronomy 10:829. https://doi.org/10.3390/agronomy10060829

Yan J, Aznar A, Chalvin C, Birdseye DS, Baidoo EEK, Eudes A, Shih PM, Loqué D, Zhang A, Scheller HV (2018) Increased drought tolerance in plants engineered for low lignin and low xylan content. Biotechnol Biofuels 11:195. https://doi.org/10.1186/s13068-018-1196-7

Yang Q, Yuan C, Cong T, Wang J, Zhang Q (2022) Genome-wide identification of three-amino-acid-loop-extension gene family and their expression profile under hormone and abiotic stress treatments during stem development of *Prunus mume*. Front Plant Sci 13:1006360. https://doi.org/10.3389/fpls.2022.1006360

Yang Z, Nielsen R (2000) Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. Mol Biol Evol 17:32–43. https://doi.org/10.1093/oxfordjournals.molbev.a026236

Zhao K, Zhang X, Cheng Z, Yao W, Li R, Jiang T, Zhou B (2019) Comprehensive analysis of the three-amino-acid-loop-extension gene family and its tissue-differential expression in response to salt stress in Poplar. Plant Physiol Biochem 136:1–12. https://doi.org/10.1016/j.plaphy.2019.01.003

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

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.