

---

# Genome-Wide Identification of Mango (*Mangifera indica* L.) MADS-Box Genes Related to Fruit Ripening

---

[Bin Zheng](#) , [Songbiao Wang](#) , [Hongxia Wu](#) , [Xiaowei Ma](#) , [Wentian Xu](#) , Kunliang Xie , [Lingfei Shangguan](#) , [Jinggui Fang](#) \*

Posted Date: 1 November 2023

doi: 10.20944/preprints202311.0050.v1

Keywords: Genome; MADS; gene family; fruit maturation



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

# Genome-Wide Identification of Mango (*Mangifera indica* L.) MADS-Box Genes Related to Fruit Ripening

Bin Zheng<sup>1,2</sup>, Songbiao Wang<sup>2</sup>, Hongxia Wu<sup>2</sup>, Xiaowei Ma<sup>2</sup>, Wentian Xu<sup>2</sup>, Kunliang Xie<sup>2</sup>, Lingfei Shangguan<sup>1</sup>, Jinggui Fang<sup>1,\*</sup>

<sup>1</sup> Horticulture Department, Nanjing Agricultural University, Nanjing 210095, China

<sup>2</sup> Key Laboratory of Tropical Fruit Biology, Ministry of Agriculture and Rural Affairs; South Subtropical Crops Research Institute, Chinese Academy of Tropical Agricultural Sciences, Zhanjiang 524091, China

\* Correspondence: fanggg@njau.edu.cn

**Abstract:** MADS-box genes play a crucial role in fruit ripening, yet limited research has been conducted on mango. Based on the conserved domains of this gene family, 84 MADS-box genes were identified in the mango genome, including 22 type I and 62 type II MADS-box genes. Gene duplication analysis revealed that both tandem duplication and segmental replication significantly contributed to the expansion of MADS-box genes in the mango genome, with purifying selection playing a vital role in the segmental duplication events within the MiMADS gene family. Cis-acting element analysis demonstrated that most MiMADS genes were hormonally regulated and participated in the growth, development, and stress resistance of mango fruit. Moreover, through expression pattern analysis and phylogenetic tree construction, we identified six MiMADS genes belonging to the SEP1 subfamily and two belonging to the AG subfamily as potential candidates involved in mango ripening regulation. Notably, Mi08g17750 and Mi04g18430 from the SEP1 subfamily were identified as key regulators inhibiting mango fruit maturation; their interaction network was also analyzed. These findings provide a foundation for further investigation into the regulatory mechanisms underlying mango ripening.

**Keywords:** genome; MADS; gene family; fruit maturation

## 1. Introduction

MADS-box genes constitute one of the largest families of plant transcription factors (TFs) [1]. Structurally, nearly all MADS-domain proteins possess a highly conserved DNA binding domain called the MADS box at the N-terminus, consisting of approximately 60 amino acid residues. This domain recognizes and binds to CarG boxes (CC[A/T]6GG), thereby regulating downstream gene transcription [2]. From an evolutionary perspective, they can be categorized into two phylogenetically distinct groups: type I and type II [3]. Generally, type I MADS genes exhibit simple structures with 1-2 exons, and limited research has been conducted on their functions. In plants, the type-I MADS TFs are further categorized into three groups based on the M domain of the encoded protein: M $\alpha$ , M $\beta$ , and M $\gamma$ . The type II genes are commonly referred to as MIKC genes due to their distinctive four domain structures: the M domain, which is highly conserved, plays a crucial role in DNA binding, dimerization, and interactions with accessory-factors; the K domain, moderately conserved, facilitates dimerization; the I (intervening) domain, less-conserved but capable of forming dimers through binding; and the variable C (C-terminal) region, which exhibits high variability [4]. Based on differences in their domain structure, MIKC-type MADS-box genes have been divided into MIKC\* and MIKC<sup>C</sup> types [5]. MIKC\* MADS-box genes feature an altered protein domain structure potentially resulting from a duplication of exons encoding the K domain subregion [6]. According to their phylogenetic relationships, MIKC<sup>C</sup>-type MADS-box genes can be further subdivided into distinctive subfamilies, such as AG, AGL6, AGL12, AGL15, ANR1, AP1, AP3, FLF, P1, TT16, SEP1, SOC1, and SVP, based on their phylogeny.

MADS-box proteins are crucial transcription factors involved in the regulation of signal transduction [7], stress responses [8,9], and various aspects of plant development, including vernalization and flowering time control [10,11], floral organ (carpels, stamens, sepals, and petals) development [12–16], carpel identity and pollen development [17,18], reproductive development [19], seed development and pigmentation [20–23], root development [24,25], formation and dehiscence of fruit [26–28], and endosperm development [29]. In recent years, numerous MADS box TFs have been discovered to be associated with fruit ripening [30]. Tomato SiRIN has emerged as a key factor in delaying fruit ripening and has been extensively investigated [31]. It was reported that SiRIN protein influences gene expression by binding to CARG box sequences present in the promoter regions of several ripening-related genes [32,33]. Furthermore, SiRIN interacted with other MADS-box proteins (FUL1, FUL2, TAGL1, TAG1, MBP21, and TDR5) to regulate fruit ripening [34]. Subsequently, some other tomato MADS TFs, such as TAGL1 [35], SiMADS1 [36], SiFYFL [37], FUL [38], FUL1, FUL2 [39], SiMBP8 [40], and SiCMB1 [41], were also confirmed to be able to regulate tomato fruit ripening through fine-tuning ethylene biosynthesis and the expression of ripening-related genes. In strawberries, a non-climacteric fruit, the silencing of a tomato MADS-RIN homologue, FaMADS9, led to the inhibition of normal development and ripening [42]. In addition, FaMADS1a negatively impacted anthocyanin accumulation in strawberry fruit by repressing the expression of FaPAL6, FaCHS, FaDFR, and FaANS, thereby causing a delay in fruit ripening [43]. In bananas, MaMADS1 and MaMADS2 were involved in ethylene-induced fruit ripening [44,45]. Inhibition of MaMADS1 or MaMADS2 expression in transgenic banana plants resulted in delayed fruit ripening and extended shelf-life [46]. In papayas, CpMADS4 potentially functioned in fruit ripening through interacting with CpNAC3 and regulating ethylene signal genes CpERF9 and CpEIL5 [47]. Moreover, CsMADS6 regulated LCYb1 as well as other carotenogenic genes to mediate carotenoid metabolism in citrus fruits [48].

Mango (*Mangifera indica* L.) is a fruit of significant economic importance in the south subtropical regions of China, but it has a limited shelf life after harvest. However, the length of the development period varies greatly among different varieties. To prolong the availability of fresh mangoes, extremely early maturity and extremely late maturity have become key indicators for breeding new mango varieties. Previous studies have reported various biochemical changes during mango fruit ripening, including increased respiration, ethylene production, fruit softening, chlorophyll degradation, carotenoid synthesis, and other metabolic activities leading to changes in carbohydrates, organic acids, lipids, phenolics, and volatile compounds [49]. Nevertheless, the regulation mechanism underlying early-maturing and later-maturing in different mango varieties remains unexplored. MADS-box TFs are known to play a role in regulating fruit ripening. In this study, based on the mango genome data, the MADS TFs of mango were identified and analyzed, and the MADS TFs related to fruit ripening were screened. These findings lay a foundation for further elucidating the molecular mechanisms governing early and late maturity.

## 2. Materials and Methods

### 2.1. Identification and Features Analysis of MADS-Box Family in Mango

We obtained mango genome sequences, proteome sequences, and GFF file from Wang [50], and retrieved the Arabidopsis MADS-box genes from The Arabidopsis Information Resource (TAIR) (<http://www.arabidopsis.org/>) databases. Two strategies were employed to identify the MADS-box transcription factor family. Firstly, the hidden Markov model (HMM) profile of SRF-TF domains (PF00319) and K-box domains (PF01486) from the Pfam database (Pfam 32.0, <http://pfam.xfam.org/>) were used as queries to identify MADS-box sequences with HMMER version 3 [51] against the mango genome with a threshold of e-value  $\leq 1 \times 10^{-5}$ . Additionally, 109 previously identified MADS protein sequences of Arabidopsis were used as queries to search against the mango genome using the BLAST program of Tootools [52] with an e-value of  $1e^{-5}$ . Subsequently, all potential candidate proteins were submitted to the Pfam database and NCBI Conserved Domain Search and Tootools Batch SMART for confirmation of the presence and completeness of the MADS domain. Candidate genes lacking the

MADS domain were manually re-annotated using the BLAST tool of TAIR 10. The physical and chemical properties of mango MADS-box (MiMADS) genes were predicted using the online tool ProtParam from ExPASy (<https://web.expasy.org/protparam/>). Subcellular localization prediction results for MiMADS genes were obtained using the tools from the WoLF PSORT website [53].

## 2.2. Phylogenetic Tree, Gene Structure and Conserved Motif Analysis of MiMADS Genes

In order to gain a comprehensive understanding of the phylogenetic relationships, and further classify type I and type II MADS-box genes, the phylogenetic tree of mango and Arabidopsis MADS proteins was constructed using the maximum likelihood method of MEGA7 software package according to the similarity of full-length amino acid sequence. Based on the relationships between MiMADS and AtMADS proteins and the classification scheme of AtMADS, all the identified MiMADS genes were categorized into distinct groups. The gene structures of MiMADS genes were obtained by comparing the open reading frames (ORFs) with the genomic sequences, which were displayed using the gene structure display server (GSDS). Additionally, conserved motifs of MiMADS proteins were investigated using the MEME program (<http://meme.nbcr.net/meme/cgi-bin/meme.cgi>), with parameters set as follows: site distribution (Zero or one occurrence per sequence), motif number (20), motif width (between 6 and 200 wide). Finally, obtained motifs were annotated using the NCBI CD search program [54].

## 2.3. Chromosomal Distribution, Gene Duplication, and Ka/Ks Analysis of MiMADS Genes

The chromosomal locations of MiMADS proteins were determined according to Wang [50], and the TBtools software package was utilized for constructing linkage maps of MiMADS proteins. Gene duplication events of MADS-box genes in mango were also investigated using Tbtools. Syntenic analysis of MADS genes in mango, *Arabidopsis thaliana* and *Citrus sinensis* was performed using the TBtools software, which embedded MCscan X software under default parameters [52]. The genome of *Arabidopsis thaliana* and *Citrus sinensis* were downloaded from the Arabidopsis (<https://www.arabidopsis.org/>) and citrus (<http://citrus.hzau.edu.cn/orange/>) databases, respectively. Furthermore, the Nei-Gojobori method within TBtools was employed to calculate the non-synonymous substitution rates (Ka) and synonymous substitution rates (Ks) of MiMADS proteins.

## 2.4. Cis-element Analysis of MiMADS Proteins

The 2,000-bp sequences upstream of the start codon of each MiMADS gene were extracted to predict cis-acting elements using the PlantCARE database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) [55]. In this study, we specifically focused on selecting cis-elements associated with hormonal and environmental responses.

## 2.5. Interaction Networks Analysis of MiMADS Proteins

The protein-protein interaction networks of MADS proteins in *Arabidopsis thaliana* and tomato were selected as reference respectively, and the interaction relationship between MiMADS proteins was further analyzed through homologous comparison. The online website STRING 11.0 (<https://string-db.org/>) was used to display the protein-protein interaction network.

## 2.6. Plant Materials Collection and Expression Analyses Based on Publicly Available RNA-seq

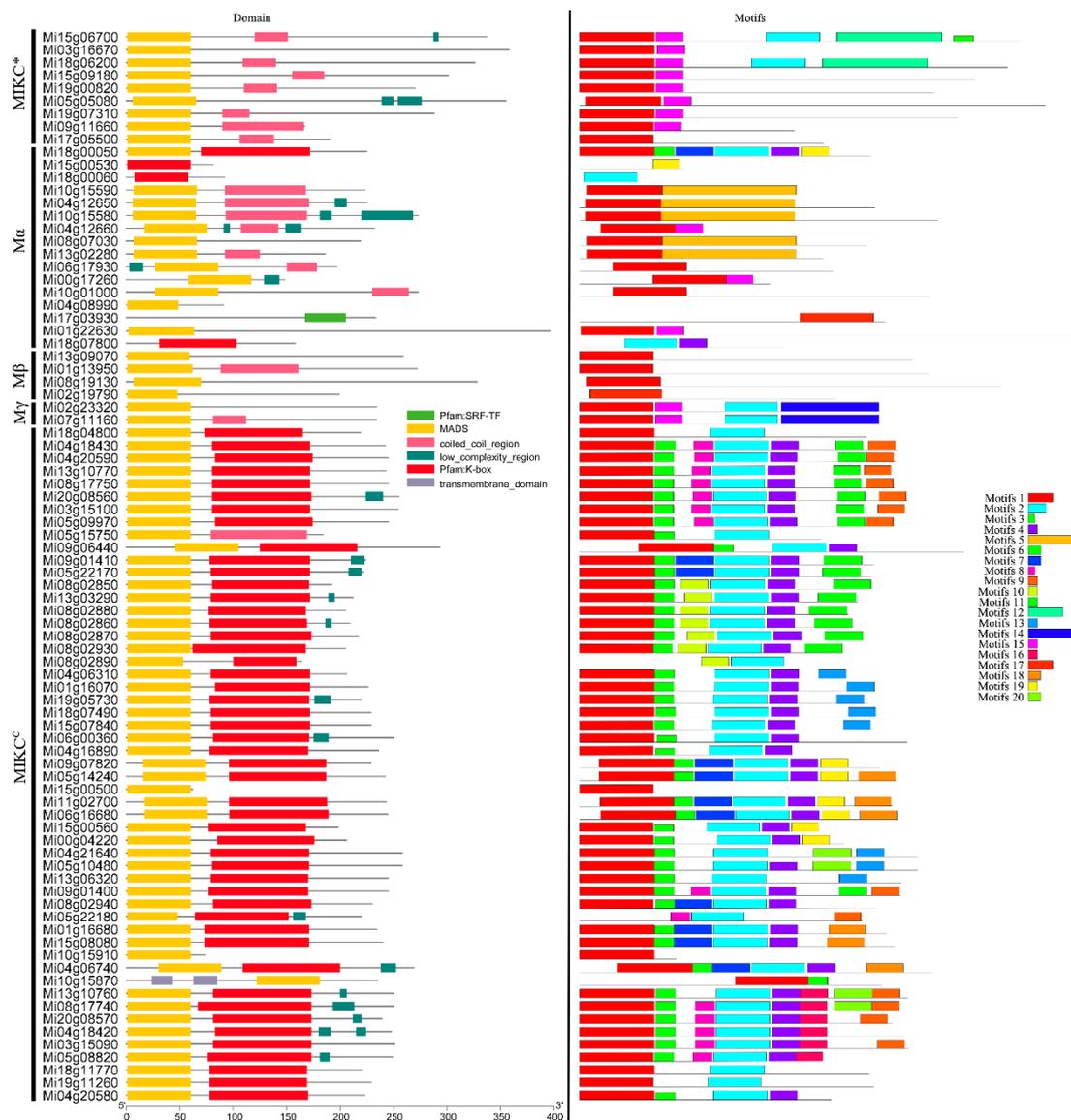
The samples utilized in this experiment conform to Li [56]. Two mango varieties, namely the early maturing Tainong-1 and the late maturing Renong-1, were employed as plant materials. All plants used in this study were cultivated at the South Subtropical Crop Research Institute in Zhanjiang, China (21°10'2"N; 110°16'34"E). Samples of Tainong-1 from four different stages were collected on March 27th (young fruit stage), April 27th (fruit enlargement stage), June 4th (green mature stage), and June 14th (full-ripe stage, one week after harvesting green mature fruits). While samples of late-maturing variety Renong-1 were collected on March 27th (young fruit stage, one week after the green mature were harvest), May 28th (fruit enlargement stage), July 6th (green mature stage), and July 13th (full-ripe stage, one week after harvesting green mature fruits). The Illumina

transcriptome raw data of two varieties at four different stages, each with three biological replicates, were download from NCBI BioProject PRJNA 629065. Clean reads were retrieved after removing low-quality (containing >50% bases with a Phred quality score <20) reads and those with unknown nucleotides (more than 1% ambiguous residues N) using the FastQC tool (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). The Clean Reads were aligned to the reference genome [50] using Hisat2 (v2.1.0, Mihaela Pertea, USA) software with default parameters [57]. Gene expression levels in each sample were estimated using fragments per kilobase of transcript per million fragments mapped (FPKM) values. Multiple hypothesis testing correction of hypothesis test probability (P value) by Benjamini-Hochberg method was also needed to obtain false discovery rate (FDR). Genes meeting the criteria of p-value < 0.05 and an absolute log<sub>2</sub> (Fold Change) ≥ 1 were identified as differentially expressed genes.

### 3. Results

#### 3.1. Identification of MADS Proteins in Mango

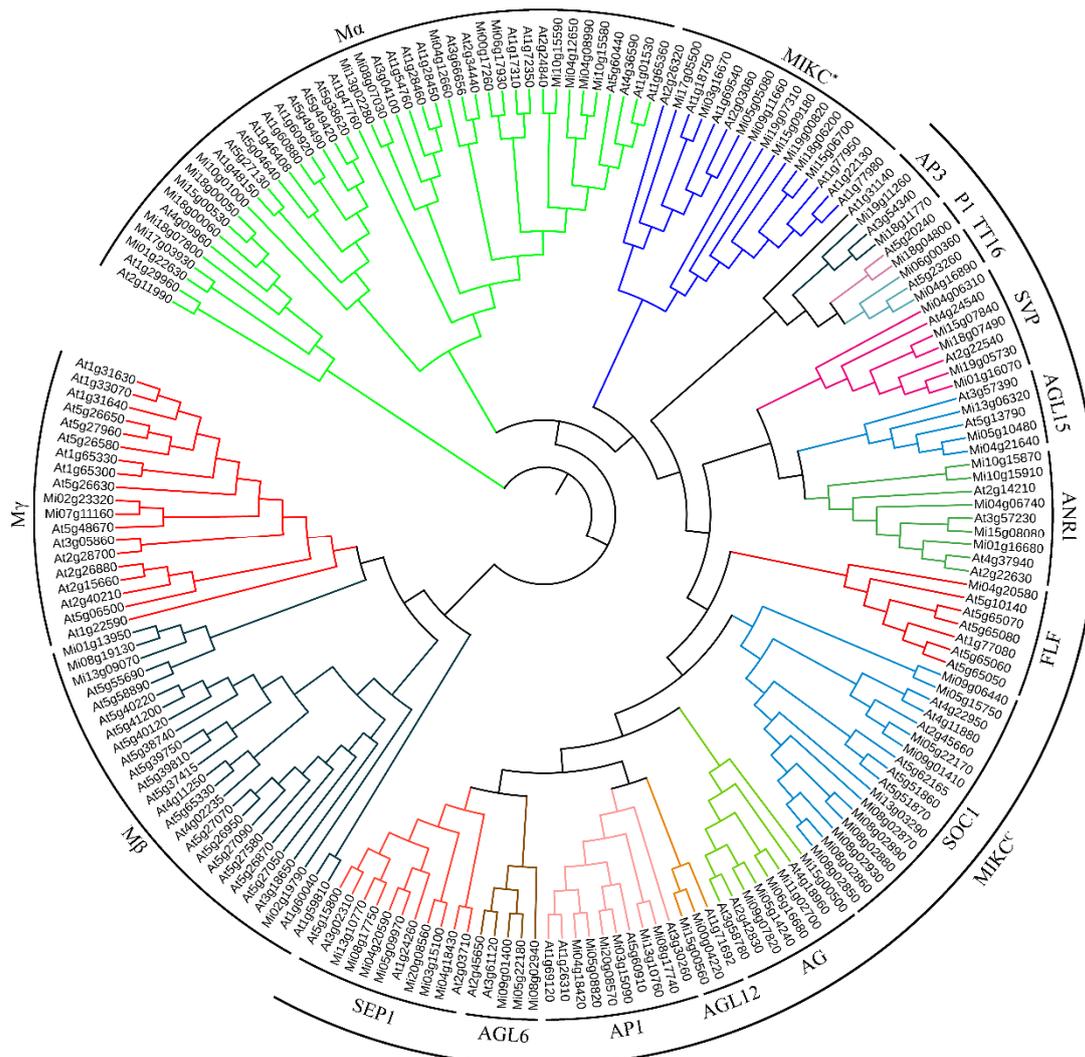
The identification of all potential MADS members in the mango genome was accomplished using two approaches: BlastP search and Hidden Markov Model (HMM) search. As a result, a total of 84 MiMADS proteins were successfully identified. Subsequent analysis confirmed that all identified MiMADS proteins contained conserved domains such as the MADS domain, SRF-TF domain, or K-box domain (Figure 1). Out of the total MiMADS proteins, 24 MiMADS proteins had only MADS or SRF-TF domain, 3 had only K-box domains, and 57 encoded for both domains (Figure 1). The length of these 84 mango MiMADS proteins ranged from 61 (Mi15g00500.1) to 395 (Mi01g22630.1) amino acid residues with relative molecular masses ranging from 6.96 (Mi15g00500.1) to 44.21 (Mi01g22630.1) kDa, while their isoelectric points fell within the range of 4.96-10.68 (Table S1). Out of the 84 MiMADS proteins, only 8 with an instability index below 40 were stable proteins, while the rest were unstable proteins with indices above 40. The aliphatic index (AI) was used to assess protein thermostability and ranged from 74.94-102.21 for MiMADS proteins. Except for Mi10g15870.1, all other MiMADS had a grand average of hydropathicity (GRAVY) less than zero and were hydrophilic proteins. Predicted subcellular localization results indicated that 66 MiMADS proteins localized in the nuclear region, eight in chloroplasts, seven in cytoplasm, and three in mitochondria.



**Figure 1.** Domain and motif analyses of MADS proteins in mango.

### 3.2. Phylogenetic Relationships of the MiMADS Proteins

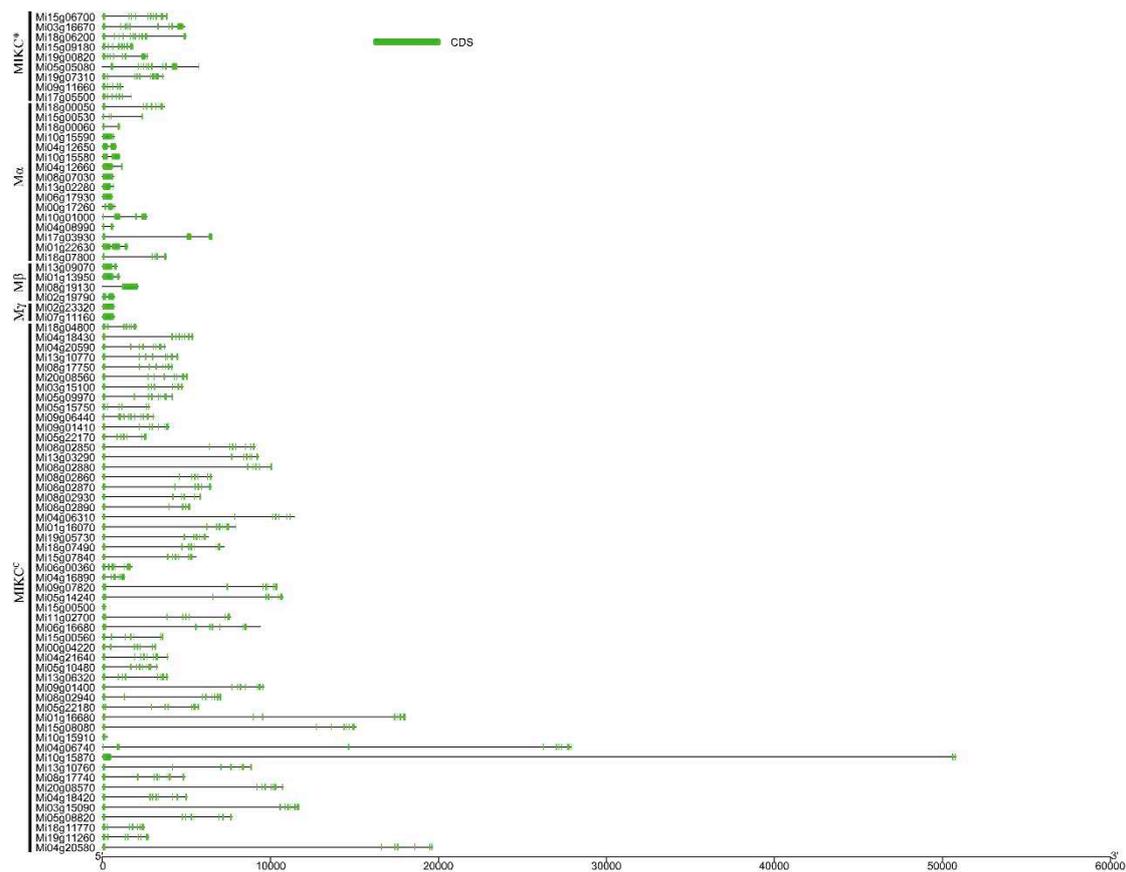
To investigate the phylogenetic relationships between MADS family proteins of mango and those in *Arabidopsis*, an unrooted maximum likelihood tree was constructed using full-length MADS-box proteins from both species with the assistance of MEGA6 software (Figure 2). According to the classification scheme of *Arabidopsis* MADS-box proteins based on previous research [58], all identified MiMADS genes were categorized into two groups: type I (22) and type II (62). Among the 22 type I genes, 16 were classified as members of the M $\alpha$  subgroup, while 4 belonged to the M $\beta$  subgroup, and 2 fell into the category of the M $\gamma$  subgroup. As for type II MiMADS genes, a majority of them (53) were assigned to subgroup MIKC<sup>c</sup>, whereas nine remaining members were classified into subgroup MIKC<sup>\*</sup>. Based on Athaliana's classification system for its own set of MADS-box genes [4], mango MIKC<sup>c</sup> proteins can be further subdivided into thirteen subfamilies along with their corresponding orthologous MADS proteins from *Arabidopsis*. Among these subfamilies, P1 and FLF each contained only one MADS protein; however, SOC1 had eleven MiMADS members, followed by seven for SEP1, six for AP1, five each for SVP, AG, and ANR1; three each for AGL15 and AGL6; two each for TT16, AGL12, and AP3.



**Figure 2.** Phylogenetic tree analysis of mango and *Arabidopsis Thaliana* MADS proteins.

### 3.3. Gene Structure and Conserved Motifs in MiMADS Proteins

The exon-intron organizations of the 84 MiMADS genes were also examined in an attempt to gain a better understanding of their structural evolution. As depicted in Figure 3, only four out of the 84 MiMADS genes consist of a single exon, and all the intronless genes were clustered within the type I group, including two Ma subgroup genes and all the My MADS subgroup genes. On the other hand, the remaining MiMADS genes exhibit varying numbers of exons ranging from two to eight, with most seven to eight-exon-containing genes being grouped under the MIKC<sup>C</sup> classification. This observation suggests that similar exon-intron organizations exist among MiMADS genes within the same group and highlights that gene structure may hold significance for gene evolution and function. Based on previous research conducted on rice, it has been concluded that intron loss occurs at a faster rate than intron gain following segmental duplication [59]. Therefore, it can be inferred that most II-type MADS-box genes likely retain their original structures while many I-type group members appear to have predominantly undergone duplication through segmental duplications followed by subsequent intron loss.



**Figure 3.** Gene structure analysis of MiMADS genes.

The motif distribution of MiMADS proteins was analyzed using the MEME online software tool. A total of 20 conserved motifs, designated as motifs 1-20, were identified in the set of 84 MiMADS proteins (Figure 1). The detected motifs were annotated using NCBI CD search. As anticipated, motif 1 corresponded to the typical MADS-domain and was found in the majority of MiMADS proteins. Motif 2, which specifies the K domain, was prevalent among MIKC<sup>c</sup> type proteins, which were the second most conserved domain and essential for protein-protein interactions among MADS-box transcriptional factors [60]. Additionally, it was observed that most MiMADS proteins within a particular group shared similar motif composition. For instance, motifs 5 were exclusively present in group M $\alpha$ ; motif 14 was solely found in groups M $\gamma$ ; motifs 12 were only detected in group MIKC<sup>\*</sup>; whereas Motif 6, 8, 9, 11, 13, 16, 18, and 20 occurred exclusively within group MIKC<sup>c</sup>. This suggests functional similarity within each respective group.

Previous studies have demonstrated that the presence of a K-box domain is exclusive to type II MADS-box genes [61]. However, our study reveals that four type I genes (Mi18g00050.1, Mi15g00530.1, Mi18g00060.1, Mi18g07800.1) also possess a K-box domain. Additionally, simpler functional roles are associated with lower domain complexities. For instance, limited information exists regarding the impact of MIKC<sup>\*</sup>, M $\gamma$ , M $\alpha$ , and M $\beta$  on plant growth and development due to their relatively uncomplicated structures consisting of one to six motifs. Conversely, the majority (44 out of 53) of the MIKC<sup>c</sup> group exhibit intricate structures with four motifs and play diverse roles in plant growth and development.

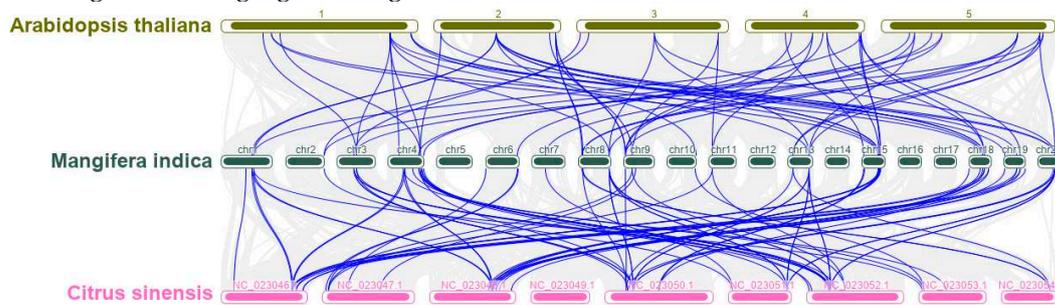
#### 3.4. Chromosomal Localization and Gene Duplication of MiMADS Genes

To investigate the genomic distribution of MADS-box genes and explore their evolutionary patterns in the context of whole genome duplication, we employed mapping coordinates to assign each gene onto the mango genome. Our findings revealed that 82 MiMADS proteins were unevenly localized on 17 chromosomes based on a previously published report of mango [50], while two MADS-box genes were traced on the unanchored scaffolds, which could be attributed to the



duplicated gene pairs. The  $K_s$  values for seven MADS gene pairs were NaN; however, the  $K_a/K_s$  ratios for the remaining 40 MADS gene pairs were less than 1, ranging from 0.08 to 0.64 with an average value of 0.21 (Table S2). These findings indicate that purifying selection played a crucial role in the segmental duplication events within the MiMADS gene family.

To establish the orthologous relationships of MiMADS, we conducted a comparative analysis of the genomic physical locations of MADS genes in *Mangifera indica*, *Citrus sinensis*, and *Arabidopsis thaliana* genomes. Based on the collinearity diagram (Figure 5), we identified 60 orthologous gene pairs between mango and Arabidopsis as well as 89 orthologous gene pairs between mango and sweet orange, indicating significant genetic similarities.



**Figure 5.** Synteny analysis of MADS-box genes between *Mangifera indica* and *Citrus sinensis* and *Arabidopsis thaliana*. The blue line indicates MADS collinearity.

### 3.5. Promoter Analysis of MADS Gene Family

Most of the well-characterized MADS-box genes are involved in plant growth and responses to hormones or environmental stimuli, such as photoperiod and temperature. Analyzing the promoter region features of MiMADS proteins will enhance our understanding of the expression patterns of MiMADS genes.

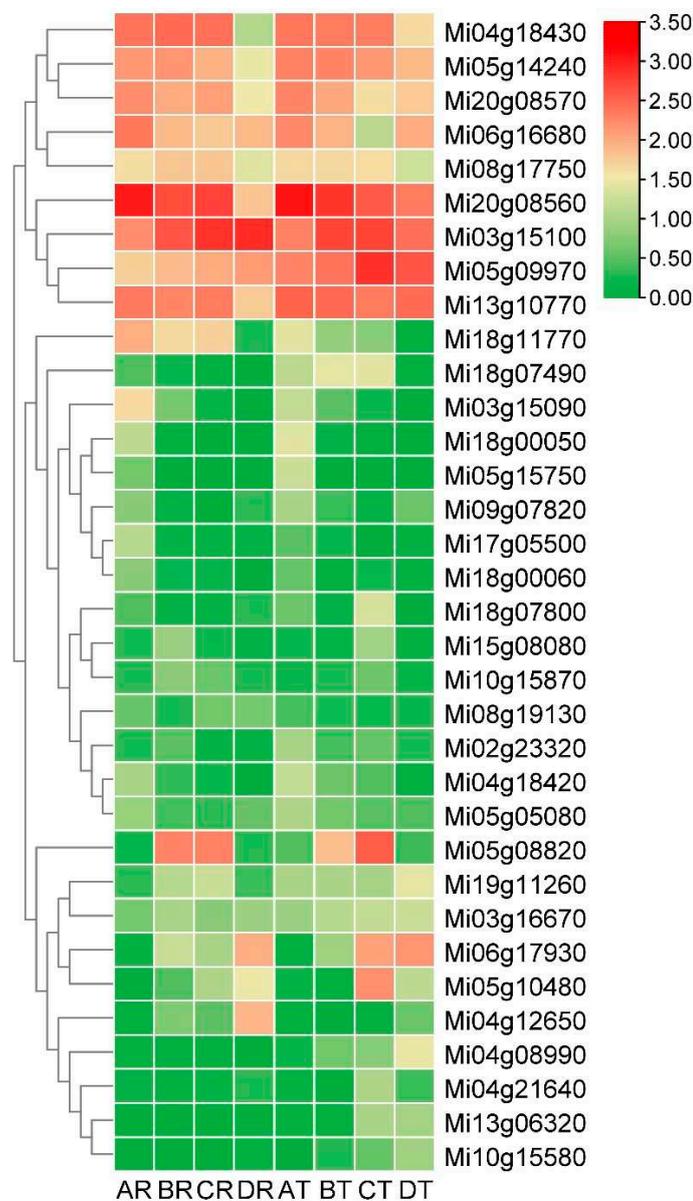
The cis-acting elements within the 2000 bp region upstream of the MADS gene family were investigated using the PlantCARE online tool [55]. Various types of elements were identified: hormone responsive, environmental stress related, promoter related, site binding related, developmental related elements, and others (Figure 6). Among these elements, we statistically analyzed the hormone responsive elements present in MADS gene family of mango and identified as-1 and TCA related to salicylic acid responsiveness in 49 MiMADS proteins. ERE and TGACG-motif related to ethylene and MeJA responsiveness, respectively, were identified in more than 40 MiMADS proteins. ABRE and TGA-element associated with the abscisic acid responsiveness and auxin-responsive element were found in 27 and 13 MiMADS proteins, respectively. Additionally, we attracted attention to the cis-elements responding to environmental stressors and identified stress-responsive (WUN-motif, TC-rich repeats, ARE, AAGAA-motif, LTR and STRE) cis-elements in 77 MiMADS family members, and light-responsive boxes (GT1-motifi, ACE, and G-box) in 54 MiMADS proteins. We also found that 12 and ten MYB binding site MBS and MBR, respectively. In summary, these results suggest that diverse cis-acting elements related to hormones and stress response regulate the functional expression of most MADS genes in mango fruit.



Figure 6. Cis-acting elements analysis of MiMADS genes.

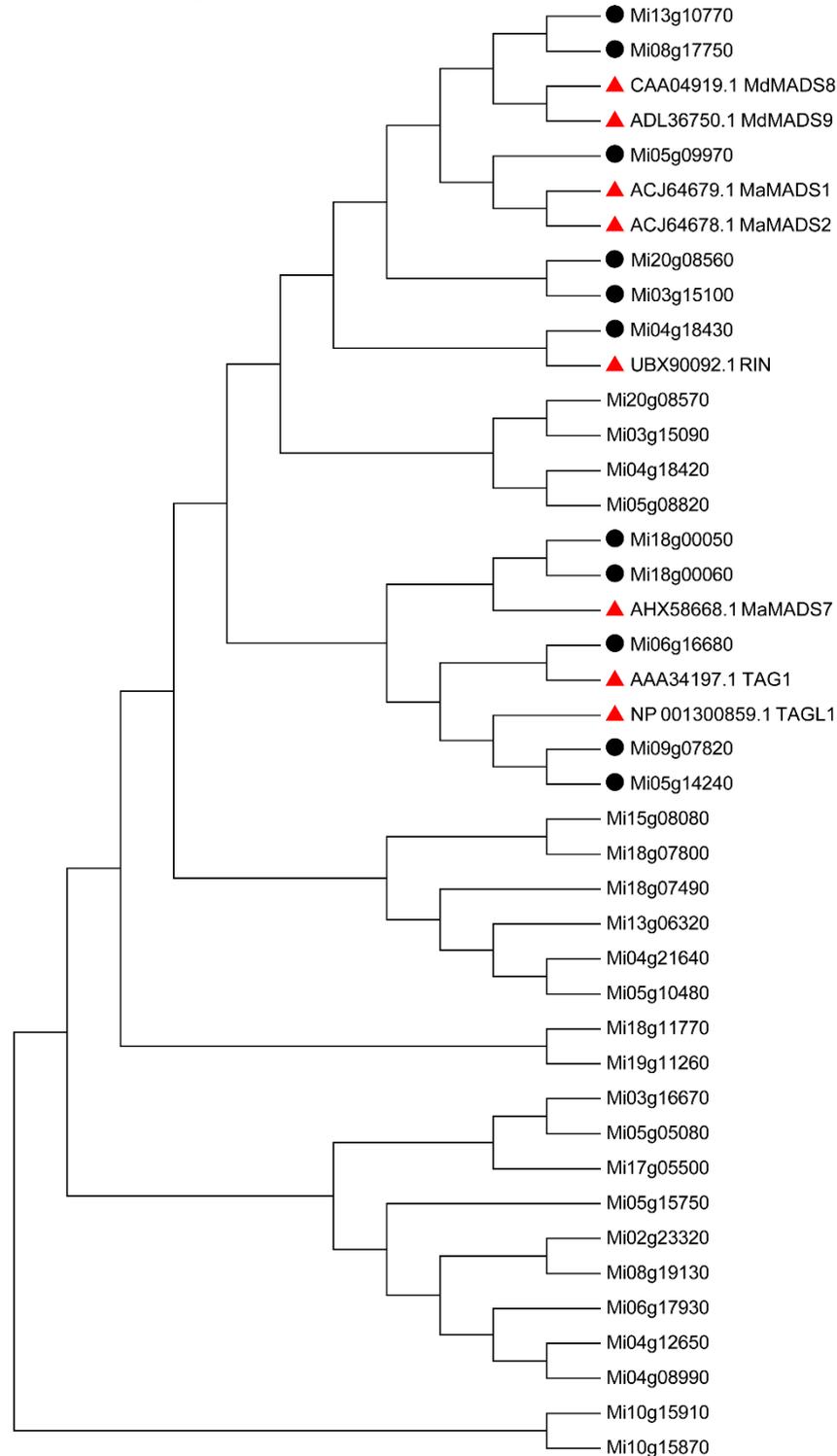
### 3.6. Screening of Early and Late Maturation Related MADS and Expression Patterns Analysis

The RNA-seq data of two mango varieties was remapped to the mango genome. It was observed that out of the 84 MiMADS proteins, 79 genes were expressed in fruit, and 55 genes exhibited differential expression across various stages of fruit development and ripening. Among these 55 differentially expressed MiMADS genes, the average FPKM value of 34 MiMADS genes in two varieties at different stages exceeded 1. Heat map analysis (FPKM value, log<sub>10</sub> scale) was further conducted on these 34 MADS genes, As depicted in Figure 7, distinct variations were observed in the expression patterns of different genes during various developmental stages between the two varieties. Notably, nine MiMADS genes displayed an average FPKM value greater than 100 across different stages in both varieties. Among the nine high-expression differential MADS genes, five MiMADS genes with highest mean FPKM values (Mi20g08560, Mi03g15100, Mi05g09970, Mi13g10770, and Mi04g18430) and Mi08g17750 belonged to the SEP1 subfamily. Additionally, two genes (Mi05g14240 and Mi06g16680) belonged to the AG subfamily. One gene (Mi20g08570) belonged to the AP1 subfamily.



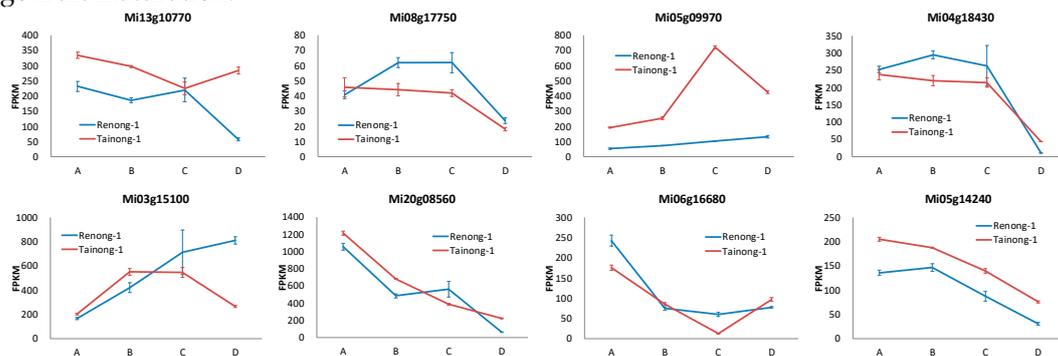
**Figure 7.** Heatmap of MiMADS FPKM value scaled by log<sub>10</sub>. AR, BR, CR, DR and AT, BT, CT, DT represent four different stages (young fruit stage, fruit enlargement stage, green mature stage, and full-ripe stage) of variety Renong-1 and Tainong-1 respectively.

Phylogenetic tree analysis was conducted on these 34 differentially expressed MADS genes and other previously reported MADS genes associated with maturity. The results revealed that Mi13g10770, Mi08g17750, Mi05g09970, Mi20g08560, Mi03g15100, and Mi04g18430 were grouped together with apple's MdMADS8 and MdMADS9, banana's MaMADS1 and MaMADS2, as well as tomato's Rin gene. Additionally, Mi18g00050, Mi18g00060, Mi06g16680, Mi09g07820, and Mi05g14240 were clustered alongside MaMADS7, TAG1, and TAGL1.



**Figure 8.** Phylogenetic tree analysis of the screened differentially expressed MADS proteins and other reported mature-related MADS proteins.

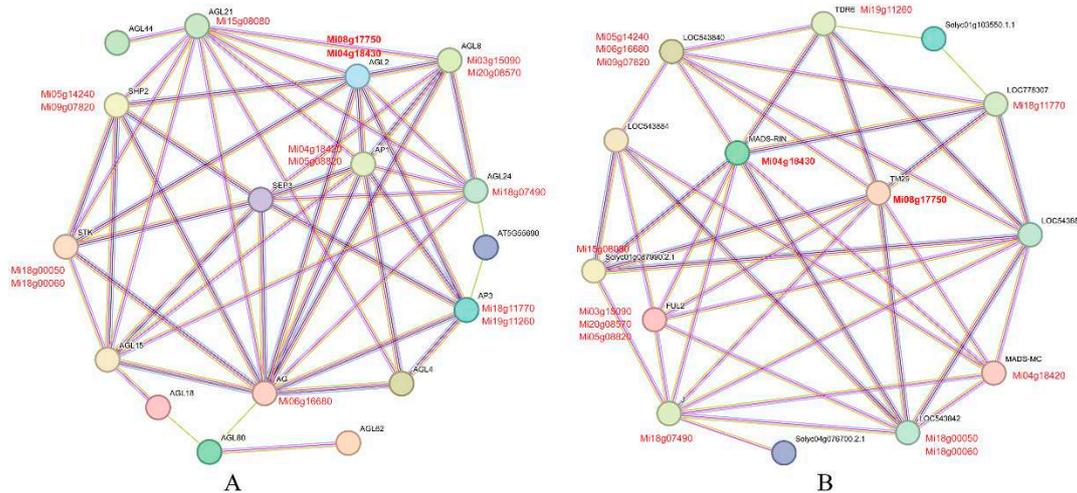
The comprehensive analysis of the phylogenetic tree and the average FPKM value (average FPKM value > 10) suggested that Mi20g08560, Mi03g15100, Mi05g09970, Mi13g10770, Mi04g18430, and Mi08g17750 from the SEP1 subfamily, as well as Mi05g14240 and Mi06g16680 from the AG subfamily, were key MiMADS genes associated with mango ripening. These results were closely aligned with the report that the AG and SEP subfamilies were the key regulators of fruit development and ripening [64]. Among them, two genes (Mi03g15100 and Mi13g10770) exhibited contrasting expression patterns during the postharvest period between Renong-1 and Tainong-1 varieties. Specifically, the expression levels of Mi03g15100 increased while those of Mi13g10770 decreased during the postharvest period of Renong-1 variety. In contrast, they showed a decreasing trend for Mi03g15100 but an increasing trend for Mi13g10770 during the postharvest period of Tainong-1 variety. The expression patterns of Mi05g09970 and Mi06g16680 in variety Tainong-1 also exhibited contrasting trends, with an initial increase followed by a decrease, and a decrease followed by an increase during fruit development and postharvest, respectively. However, the expression levels of these two genes remained relatively stable throughout the latter three periods of Renong-1. The expression levels of Mi20g08560 and Mi05g14240 exhibited a declining trend throughout the development and postharvest in both varieties, and were higher in Tainong-1 than in Renong-1 at most stages. The expression levels of Mi08g17750 and Mi04g18430 in Renong-1 were consistently higher than those in Tainong-1 across most stages, with a significant decrease observed during the postharvest stage, indicating that Mi08g17750 and Mi04g18430 might be involved in the inhibition of mango fruit maturation.



**Figure 9.** Expression of eight key MiMADS genes related to mango ripening in different developmental stages of two varieties. A, B, C, and D represent four different stages (young fruit stage, fruit enlargement stage, green mature stage, and full-ripe stage) of two varieties.

### 3.7. Protein Interaction Network Analysis of the MiMADS genes

Protein-protein interactions are crucial for the functioning of MADS-box proteins; therefore, evaluating the interaction networks is valuable for characterizing gene function mechanisms. In this study, Mi08g17750 and Mi04g18430 were selected to evaluate possible protein interaction using the STRING website to construct the protein interaction networks, *Arabidopsis thaliana* and *Solanum lycopersicum* were chosen as reference organisms, respectively. When *Arabidopsis thaliana* was used as a reference, both Mi04g18430 and Mi08g17750 showed highly homologous with AGL2, and may interact with several other MiMADS proteins including Mi03g15090, Mi20g08570, Mi18g07490, Mi18g11770, Mi19g11260, Mi06g16680, Mi04g18420, Mi05g08820, Mi05g14240, Mi09g07820, Mi18g00050, Mi18g00060, and Mi15g08080. When *Solanum lycopersicum* was taken as a reference, Mi04g18430 and Mi08g17750 exhibited highly homologous with MADS-RIN and TM29, respectively. Both genes may interact with Mi19g11260, Mi18g11770, Mi04g18420, Mi18g00050, Mi18g00060, Mi18g07490, Mi03g15090, Mi20g08570, Mi05g08820, Mi06g16680, Mi05g14240, Mi09g07820, and Mi15g08080. It is worth noting that the screened proteins interacting with Mi04g18430 and Mi08g17750 are consistent when *Arabidopsis thaliana* and *Solanum lycopersicum* were taken as reference respectively.



**Figure 10.** Functional interaction network of MiMADS proteins in mango based on their orthologs in *Arabidopsis thaliana* (A) and *Solanum lycopersicum* (B). Circle nodes in the network represent proteins, and straight lines connecting the nodes represent interaction relationships. Indigo blue and purple lines indicate known interactions; green, red, and navy blue balls indicate predicted interactions.

#### 4. Discussion

The MADS-box genes are ubiquitous in plants and play crucial roles in various processes including plant growth and development, stress response, signal transduction, and other processes. With the publication of high-quality genomes, numerous MADS-domain proteins have been extensively identified and characterized through comprehensive genome-wide identification and analysis of expression profiles, such as *Arabidopsis* (107 genes) [4], rice (75 genes) [65], tomato (131 genes) [66], potato (153 genes) [67], bread wheat (300 genes) [68], cucumber (43 genes) [69], *Brassica rapa* (167 genes) [70], radish (144 genes) [71], sesame (57 genes) [72], watermelon (39 genes) [73], moso bamboo (42 genes) [74], sorghum (65 genes) [75], apple (146 genes) [76], chrysanthemum (108 genes) [77], eggplant (*Solanum melongena*) (120 genes) [78], blueberry (249 genes) [79], chayote (*Sechium edule*) (70 genes) [80], poplar (105 genes) [81], pear (95 genes) [82], banana (96 genes) [83], grape (58 genes) [84], soybean (106 genes) [85], honeysuckle (*Lonicera japonica*) (48 genes) [86], *Zea mays* (211 genes) [87], and *Rhododendron ovatum* (77 genes) [88].

In this study, a total of 84 MiMADS genes were identified in the mango genome, comprising 22 type I genes and 62 type II genes based on their phylogenetic relationships. Comparing with previous studies, it was observed that the number of MADS-box genes varies across different species, but the number of members in type II exceeded that in type I, and this finding was consistent with the results obtained in the present study. Moreover, the number of MADS proteins in different subfamilies varies significantly among species. We found that while mango has a lower overall count of MADS proteins compared to *Arabidopsis thaliana*, there is a higher number of  $M\alpha$  MADS-box genes (16 genes) in mango than in *Arabidopsis thaliana* (11 genes) [4]. This difference may be attributed to gene segmental duplication during the evolutionary process of mango.

In this study, a total of 66 MiMADS proteins were predicted to localize in the nucleus, which is consistent with previous reports that transcription factor proteins mainly participate in regulating specific genes within the eukaryotic nucleus. Additionally, we identified 18 MiMADS proteins residing outside the nucleus, among which eight were found in chloroplasts and three were located in mitochondria. Furthermore, cytoplasmic localization was predicted for seven proteins. Notably, both chloroplasts and mitochondria contain DNA and RNA as well as possess essential machinery for gene transcription. It is plausible that certain MiMADS genes may play a regulatory role in photosynthesis and respiratory-related genes. Moreover, some transcription factor proteins exhibit nucleocytoplasmic shuttling through phosphorylation/dephosphorylation-mediated trafficking mechanisms [89], suggesting that MiMADS outside the nucleus might translocate into the nucleus

when required for gene regulation. Similar findings have been reported regarding MADS protein localization in blueberry [79].

The gene structure, domain, and motif analysis showed that there were differences in the MADS gene of Mango. Type II MiMADS genes exhibited a higher number of exons compared to type I genes, indicating their more intricate structural composition. Notably, the exon-intron patterns observed in both type I and type II genes remain conserved across diverse plant taxa, including Arabidopsis [4], rice [67], and grape [90], underscoring the high conservation of MADS-box TFs among plants. However, it is worth mentioning that several type I group genes also possess two or more intron structures, indicating these genes might have inequable functions. Different subfamilies exhibit variations in the number and types of domains and conserved motifs [91]. In our study, it was observed that some MIKCC genes and all MIKC\* genes lack K-box domains, whereas certain type I MiMADS genes contain K-box domains. The absence of K-box domains in the MIKC\* subfamily genes has been reported in several species, including foxtail millet [92], American beautyberry [93], and litchi [94].

Some studies have indicated that intron loss and insertion mutations are commonly observed during the evolutionary process of plant MADS-box genes [71]. MIKCC genes are believed to be the most ancient members of the MADS-box gene family, and type I genes likely evolving from MIKCC genes [5]. This suggests that MIKC\* genes may be a class of transition genes retained after the loss of the K-box domain during the evolution of MIKCC genes. Furthermore, certain type I MiMADS genes retain conserved K-box domains throughout their evolution. Additionally, it is worth noting that the MiMADS protein motifs within the same subfamily exhibit variations, indicating that intron loss or gain may serve as a pattern for MiMADS gene evolution and contribute significantly to the functional diversity within the MiMADS family.

The MADS-box genes are believed to originate from gene duplication events that initially occurred in the most recent common ancestor of extant eukaryote lineages [95]. Gene replication plays an important role in the amplification and evolution of transcription factors. In this study, 34 MiMADS genes (40.48%) were located in 14 clusters as tandem duplicated genes, and 47 pairs of MiMADS genes were found on segmental duplicated genome blocks. These findings highlight the significant contribution of both tandem duplication and segmental replication in facilitating the expansion of MADS-box genes within mango genomes. The same results were observed in *Cyclocarya paliurus* [96], *Rhododendron hainanense* Merr. [97], *Fagopyrum tataricum* [98], Arabidopsis [4] and potato [71]. However, only tandem duplications have been documented in cucumber [73]. These observations suggest that gene duplication plays diverse roles across different species. In addition, we observed that with the exception of seven pairs of segmental replication gene pairs with an undefined k value (NA), all other 40 pairs exhibited a k value less than 1, indicating a significant role of purifying selection in the segmental duplication events within the MiMADS gene family. These findings align with previous studies, protein neofunctionalization may not be the predominant status for the retained genes from recent WGD of mango [52].

Through collinearity analysis among different species, a total of 60 pairs of collinear MADS-box genes were identified between mango and Arabidopsis, while 89 pairs were found between mango and *Citrus sinensis* (sweet orange). The number of homologous events observed between mango and *Citrus sinensis* was significantly higher compared to that between mango and Arabidopsis, which is consistent with the relatively closer evolutionary relationship between mango and *Citrus sinensis* (sweet orange).

We identified a total of 287 cis-acting elements associated with phytohormones, which were found to be present in 76 MiMADS genes. Furthermore, we observed the presence of cis-acting elements related to stress responses, including anaerobic conditions, low temperatures, and wound signals in 77 MiMADS genes. These findings strongly suggest that the majority of MiMADS genes are under hormonal regulation and play crucial roles in the growth, development, and stress resistance mechanisms of mango fruit.

According to the RNA-seq data and phylogenetic tree analysis, we hypothesized that Mi20g08560, Mi03g15100, Mi05g09970, Mi13g10770, Mi04g18430, Mi08g17750 of the SEP1 subfamily

along with Mi05g14240 and Mi06g16680 of the AG subfamily were key MADS genes associated with mango ripening. Moreover, both Mi08g17750 and Mi04g18430 played a role in inhibiting mango fruit maturation, which aligns with previous findings [43]. The expression of four SEP genes in bananas has been shown to elevate ethylene content and influence banana maturation [99]. Similarly, the expression of SEP4-like genes essential for strawberry berry ripening [100].

## 5. Conclusions

In this study, 84 MADS-box gene family members were identified in the mango genome, and their physical and chemical properties, subcellular localization, chromosome position, gene structure, protein conserved domains, motifs, evolutionary connections, and cis-acting elements were analyzed. The expression profiles of MADS genes in different fruit development stages of two varieties were also analyzed. Our findings lay the foundation for a comprehensive functional characterization of the MADS-box gene family in mango, and provide candidate genes for studying the role of the MADS-box gene family in regulating fruit ripening. Additionally, the potential functions and regulatory mechanisms of Mi08g17750 and Mi04g18430 will be experimentally illuminated in our further research. This study will help us to comprehensively understand the characteristics of the MiMADS gene family and screen for MiMADS genes involved in mango fruit ripening.

**Supplementary Materials:** The following supporting information can be downloaded at the website of this paper posted on Preprints.org. Table S1: Protein information analysis of MADS-box gene family members in mango; Table S2: The Ka and Ks values for 47 duplicated gene pairs of MADS-box gene family members in mango; Table S3: Details of the cis-acting elements for MADS-box gene family members in mango.

**Author Contributions:** Methodology, B.Z. and H.W.; project administration, S.W., X.M. and J.F.; resources, S.W. and H.W.; supervision, S.W., H.W. and J.F.; validation, B.Z.; visualization, W.X. and K.X.; writing—original draft, B.Z.; writing—review and editing, L.S. and J.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Natural Science Foundation of Hainan Province (322QN378), Fund for Less Developed Regions of the National Natural Science Foundation of China (32360727), Key Research and Development Program of Hainan Province (ZDYF2022XDNY255), Seed Industry Revitalization Project of Guangdong Province (2022-NPY-00-030).

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank anonymous reviewers for their critical reading, stimulating discussions, and helpful comments and suggestions, which allowed us to improve the quality of this manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Riechmann J.L.; Heard J.; Martin G.; Reuber L.; Jiang C.; Keddie J.; Adam L.; Pineda O.; Ratcliffe O.J.; Samaha R.R.; Creelman R.; Pilgrim M.; Broun P.; Zhang J.Z.; Ghandehari D.; Sherman B.K.; Yu G. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 2000, 290(5499), 2105–2110. doi: 10.1126/science.290.5499.2105. PMID: 11118137.
2. Messenguy F.; Dubois E. Role of MADS box proteins and their cofactors in combinatorial control of gene expression and cell development. *Gene* 2003, 316, 1–21. doi: 10.1016/s0378-1119(03)00747-9. PMID: 14563547.
3. Alvarez-Buylla E.R.; Liljegren S.J.; Pelaz S.; Gold S.E.; Burgeff C.; Ditta G.S.; Vergara-Silva F.; Yanofsky M.F. MADS-box gene evolution beyond flowers: expression in pollen, endosperm, guard cells, roots and trichomes. *Plant J.* 2000, 24(4), 457–466. doi: 10.1046/j.1365-313x.2000.00891.x. PMID: 11115127.
4. Parenicová L.; de Folter S.; Kieffer M.; Horner D.S.; Favalli C.; Busscher J.; Cook H.E.; Ingram R.M.; Kater M.M.; Davies B.; Angenent G.C.; Colombo L. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world. *Plant Cell* 2003, 15(7), 1538–1551. doi: 10.1105/tpc.011544. PMID: 12837945; PMCID: PMC165399.

5. Henschel K.; Kofuji R.; Hasebe M.; Saedler H.; Münster T.; Theissen G. Two ancient classes of MIKC-type MADS-box genes are present in the moss *Physcomitrella patens*. *Mol Biol Evol.* 2002, 19(6), 801–814. doi: 10.1093/oxfordjournals.molbev.a004137. PMID: 12032236.
6. Kwantes M.; Liebsch D.; Verelst W. How MIKC\* MADS-box genes originated and evidence for their conserved function throughout the evolution of vascular plant gametophytes. *Mol Biol Evol.* 2012, 29(1), 293–302. doi: 10.1093/molbev/msr200. PMID: 21813465.
7. Falavigna V.; Severing E.; Lai X.; Estevan J.; Farrera I.; Hugouvieux V.; Revers L.F.; Zubieta C.; Coupland G.; Costes E.; Andrés F. Unraveling the role of MADS transcription factor complexes in apple tree dormancy. *New Phytol.* 2021, 232(5), 2071–2088. doi: 10.1111/nph.17710. PMID: 34480759; PMCID: PMC9292984.
8. Wang Z.; Wang F.; Hong Y.; Yao J.; Ren Z.; Shi H.; Zhu J.K. The flowering repressor SVP confers drought resistance in *Arabidopsis* by regulating abscisic acid catabolism. *Mol Plant.* 2018, 1184–1197. doi: 10.1016/j.molp.2018.06.009.
9. Chen R.; Ma J.; Luo D.; Hou X.; Ma F.; Zhang Y.; Meng Y.; Zhang H.; Guo W. CaMADS, a MADS-box transcription factor from pepper, plays an important role in the response to cold, salt, and osmotic stress. *Plant Sci.* 2019, 280, 164–174. DOI: 10.1016/j.plantsci.2018.11.020.
10. Ratcliffe O.J.; Kumimoto R.W.; Wong B.J.; Riechmann J.L. Analysis of the *Arabidopsis* MADS AFFECTING FLOWERING gene family: MAF2 prevents vernalization by short periods of cold. *Plant Cell* 2003, 15(5), 1159–1169. doi: 10.1105/tpc.009506. PMID: 12724541; PMCID: PMC153723.
11. Liao W.Y.; Lin L.F.; Lin M.D.; Hsieh S.C.; Li A.Y.; Tsay Y.S.; Chou M.L. Overexpression of *Lilium formosanum* MADS-box (LFMADS) causing floral defects while promoting flowering in *Arabidopsis thaliana*, whereas only affecting floral transition time in *Nicotiana tabacum*. *Int J Mol Sci.* 2018, 19(8), 2217. doi: 10.3390/ijms19082217. PMID: 30060634; PMCID: PMC6121541.
12. Benlloch R.; Roque E.; Ferrándiz C.; Cosson V.; Caballero T.; Penmetsa R.V.; Beltrán J.P.; Cañas L.A.; Ratet P.; Madueño F. Analysis of B function in legumes: PISTILLATA proteins do not require the PI motif for floral organ development in *Medicago truncatula*. *Plant J.* 2009, 60(1), 102–111. doi: 10.1111/j.1365-313X.2009.03939.x. PMID: 19500303.
13. Ferrario S.; Shchennikova A.V.; Franken J.; Immink R.G.; Angenent G.C. Control of floral meristem determinacy in *petunia* by MADS-box transcription factors. *Plant Physiol.* 2006, 140(3), 890–898. doi: 10.1104/pp.105.072660. PMID: 16428599; PMCID: PMC1400554.
14. Yu X.; Chen G.; Guo X.; Lu Y.; Zhang J.; Hu J.; Tian S.; Hu Z. Silencing SiAGL6, a tomato AGAMOUS-LIKE6 lineage gene, generates fused sepal and green petal. *Plant Cell Rep.* 2017, 36(6), 959–969. doi: 10.1007/s00299-017-2129-9. PMID: 28352968.
15. Li N.; Huang B.; Tang N.; Jian W.; Zou J.; Chen J.; Cao H.; Habib S.; Dong X.; Wei W.; Gao Y.; Li Z. The MADS-box gene SiMBP21 regulates sepal size mediated by ethylene and auxin in tomato. *Plant Cell Physiol.* 2017, 58(12), 2241–2256. doi: 10.1093/pcp/pcx158. PMID: 29069449.
16. Pelaz S.; Ditta G.S.; Baumann E.; Wisman E.; Yanofsky M.F. B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature* 2000, 405(6783), 200–203. doi: 10.1038/35012103. PMID: 10821278.
17. Pan. I.L.; Mcquinn R.; Giovannoni J.J.; Irish, V.F. Functional diversification of AGAMOUS lineage genes in regulating tomato flower and fruit development. *J. Exp. Bot.* 2010, 61(6), 1795–1806. doi: 10.1093/jxb/erq046. PMID: 20335407; PMCID: PMC2852668.
18. Giménez E.; Pineda B.; Capel J.; Antón M.T.; Atarés A.; Pérez-Martín F.; García-Sogo B.; Angosto T.; Moreno V.; Lozano R. Functional analysis of the Arlequin mutant corroborates the essential role of the Arlequin/TAGL1 gene during reproductive development of tomato. *PLoS One* 2010, 5(12), e14427. doi: 10.1371/journal.pone.0014427. PMID: 21203447; PMCID: PMC3009712.
19. Guo X.; Chen G.; Naeem M.; Yu X.; Tang B.; Li A.; Hu Z. The MADS-box gene SiMBP11 regulates plant architecture and affects reproductive development in tomato plants. *Plant Sci.* 2017, 258, 90–101. doi: 10.1016/j.plantsci.2017.02.005. PMID: 28330566.
20. Mizzotti C.; Mendes M.A.; Caporali E.; Schnittger A.; Kater M.M.; Battaglia R.; Colombo L. The MADS box genes SEEDSTICK and ARABIDOPSIS Bsister play a maternal role in fertilization and seed development. *Plant J.* 2012, 70(3), 409–420. doi: 10.1111/j.1365-313X.2011.04878.x. PMID: 22176531.
21. Malabarba J.; Buffon V.; Mariath J.E.A.; Gaeta M.L.; Dornelas M.C.; Margis-Pinheiro M.; Pasquali G.; Revers L.F. The MADS-box gene Agamous-like 11 is essential for seed morphogenesis in grapevine. *J Exp Bot.* 2017, 68(7), 1493–1506. doi: 10.1093/jxb/erx025. PMID: 28369525.

22. Chen C.; Begcy K.; Liu K.; Folsom J.J.; Wang Z.; Zhang C.; Walia H. Heat stress yields a unique MADS box transcription factor in determining seed size and thermal sensitivity. *Plant Physiol.* 2016, 171(1), 606–622. doi: 10.1104/pp.15.01992. PMID: 26936896; PMCID: PMC4854699.
23. Nesi N.; Debeaujon I.; Jond C.; Stewart A.J.; Jenkins G.I.; Caboche M.; Lepiniec L. The TRANSPARENT TESTA16 locus encodes the ARABIDOPSIS BSISTER MADS domain protein and is required for proper development and pigmentation of the seed coat. *Plant Cell* 2002, 14(10), 2463–2479. doi: 10.1105/tpc.004127. PMID: 12368498; PMCID: PMC151229.
24. Tapia-López R.; García-Ponce B.; Dubrovsky J.G.; Garay-Arroyo A.; Pérez-Ruiz R.V.; Kim S.H.; Acevedo F.; Pelaz S.; Alvarez-Buylla E.R. An AGAMOUS-related MADS-box gene, XAL1 (AGL12), regulates root meristem cell proliferation and flowering transition in Arabidopsis. *Plant Physiol.* 2008, 146(3), 1182–1192. doi: 10.1104/pp.107.108647. PMID: 18203871; PMCID: PMC2259045.
25. Li A.; Chen G.; Yu X.; Zhu Z.; Zhang L.; Zhou S.; Hu Z. The tomato MADS-box gene SiMBP9 negatively regulates lateral root formation and apical dominance by reducing auxin biosynthesis and transport. *Plant Cell Rep.* 2019, 38(8), 951–963. doi: 10.1007/s00299-019-02417-x. PMID: 31062133.
26. Ferrándiz C.; Fourquin C. Role of the FUL-SHP network in the evolution of fruit morphology and function. *J Exp Bot.* 2014, 65(16), 4505–4513. doi: 10.1093/jxb/ert479. PMID: 24482369.
27. Ferrándiz C.; Liljegren S.J.; Yanofsky M.F. Negative regulation of the SHATTERPROOF genes by FRUITFULL during Arabidopsis fruit development. *Science* 2000, 289(5478), 436–438. doi: 10.1126/science.289.5478.436. PMID: 10903201.
28. Huang B.; Routaboul J.M.; Liu M.; Deng W.; Maza E.; Mila I.; Hu G.; Zouine M.; Frasse P.; Vrebalov J.T.; Giovannoni J.J.; Li Z.; van der Rest B.; Bouzayen M. Overexpression of the class D MADS-box gene Si-AGL11 impacts fleshy tissue differentiation and structure in tomato fruits. *J Exp Bot.* 2017, 68(17), 4869–4884. doi: 10.1093/jxb/erx303. PMID: 28992179.
29. Kang I.H.; Steffen J.G.; Portereiko M.F.; Lloyd A.; Drews G.N. The AGL62 MADS domain protein regulates cellularization during endosperm development in Arabidopsis. *Plant Cell* 2008, 20(3), 635–647. doi: 10.1105/tpc.107.055137. PMID: 18334668; PMCID: PMC2329934.
30. Karlova R.; Chapman N.; David K.; Angenent G.C.; Seymour G.B.; de Maagd R.A. Transcriptional control of fleshy fruit development and ripening. *J Exp Bot.* 2014, 65(16), 4527–4541. doi: 10.1093/jxb/eru316. PMID: 25080453.
31. Causier B.; Kieffer M.; Davies B. Plant biology. MADS-box genes reach maturity. *Science* 2002, 296(5566), 275–276. doi: 10.1126/science.1071401. PMID: 11951024.
32. Fujisawa M.; Shima Y.; Higuchi N.; Nakano T.; Koyama Y.; Kasumi T.; Ito Y. Direct targets of the tomato-ripening regulator RIN identified by transcriptome and chromatin immunoprecipitation analyses. *Planta* 2012, 235(6), 1107–1122. doi: 10.1007/s00425-011-1561-2. PMID: 22160566.
33. Qin G.; Wang Y.; Cao B.; Wang W.; Tian S. Unraveling the regulatory network of the MADS box transcription factor RIN in fruit ripening. *Plant J.* 2012, 70(2), 243–255. doi: 10.1111/j.1365-313X.2011.04861.x. PMID: 22098335.
34. Martel C.; Vrebalov J.; Tafelmeyer P.; Giovannoni J.J. The tomato MADS-box transcription factor RIPENING INHIBITOR interacts with promoters involved in numerous ripening processes in a COLORLESS NONRIPENING-dependent manner. *Plant Physiol.* 2011, 157(3), 1568–1579. doi: 10.1104/pp.111.181107. PMID: 21941001; PMCID: PMC3252172.
35. Vrebalov J.; Pan I.L.; Arroyo A.J.; McQuinn R.; Chung M.; Poole M.; Rose J.; Seymour G.; Grandillo S.; Giovannoni J.; Irish V.F. Fleshy fruit expansion and ripening are regulated by the tomato SHATTERPROOF gene TAGL1. *Plant Cell* 2009, 21(10), 3041–3062. doi: 10.1105/tpc.109.066936. PMID: 19880793; PMCID: PMC2782289.
36. Dong T.; Hu Z.; Deng L.; Wang Y.; Zhu M.; Zhang J.; Chen G. A tomato MADS-box transcription factor, SiMADS1, acts as a negative regulator of fruit ripening. *Plant Physiol.* 2013, 163(2), 1026–1036. doi: 10.1104/pp.113.224436. PMID: 24006286; PMCID: PMC3793022.
37. Xie Q.; Hu Z.; Zhu Z.; Dong T.; Zhao Z.; Cui B.; Chen G. Overexpression of a novel MADS-box gene SiFYFL delays senescence, fruit ripening and abscission in tomato. *Sci Rep.* 2014, 4, 4367. doi: 10.1038/srep04367. PMID: 24621662; PMCID: PMC3952145.
38. Wang S.; Lu G.; Hou Z.; Luo Z.; Wang T.; Li H.; Zhang J.; Ye Z. Members of the tomato FRUITFULL MADS-box family regulate style abscission and fruit ripening. *J Exp Bot.* 2014, 65(12), 3005–3014. doi: 10.1093/jxb/eru137. PMID: 24723399; PMCID: PMC4071821.

39. Shima Y.; Kitagawa M.; Fujisawa M.; Nakano T.; Kato H.; Kimbara J.; Kasumi T.; Ito Y. Tomato FRUITFULL homologues act in fruit ripening via forming MADS-box transcription factor complexes with RIN. *Plant Mol Biol.* 2013, 82(4-5), 427–438. doi: 10.1007/s11103-013-0071-y. PMID: 23677393.
40. Yin W.; Hu Z.; Cui B.; Guo X.; Hu J.; Zhu Z.; Chen G. Suppression of the MADS-box gene SiMBP8 accelerates fruit ripening of tomato (*Solanum lycopersicum*). *Plant Physiol Biochem.* 2017, 118, 235–244. doi: 10.1016/j.plaphy.2017.06.019. PMID: 28649000.
41. Zhang J.; Hu Z.; Yao Q.; Guo X.; Nguyen V.; Li F.; Chen G. A tomato MADS-box protein, SiCMB1, regulates ethylene biosynthesis and carotenoid accumulation during fruit ripening. *Sci Rep.* 2018, 8(1), 3413. doi: 10.1038/s41598-018-21672-8. PMID: 29467500; PMCID: PMC5821886.
42. Seymour G.B.; Ryder C.D.; Cevik V.; Hammond J.P.; Popovich A.; King G.J.; Vrebalov J.; Giovannoni J.J.; Manning K. A SEPALLATA gene is involved in the development and ripening of strawberry (*Fragaria × ananassa* Duch.) fruit, a non-climacteric tissue. *J Exp Bot.* 2011, 62(3), 1179–1188. doi: 10.1093/jxb/erq360. PMID: 21115665; PMCID: PMC3022409.
43. Lu W.; Chen J.; Ren X.; Yuan J.; Han X.; Mao L. Ying T. Luo Z. One novel strawberry MADS-box transcription factor FaMADS1a acts as a negative regulator in fruit ripening. *Sci. Hortic.* 2018, 227, 124–131. doi: 10.1016/j.scienta.2017.09.042.
44. Liu J.; Xu B.; Hu L.; Li M.; Su W.; Wu J.; Yang J.; Jin Z. Involvement of a banana MADS-box transcription factor gene in ethylene-induced fruit ripening. *Plant Cell Rep.* 2009, 28(1), 103–111. doi: 10.1007/s00299-008-0613-y. PMID: 18820933.
45. Elitzur T.; Vrebalov J.; Giovannoni J.J.; Goldschmidt E.E.; Friedman H. The regulation of MADS-box gene expression during ripening of banana and their regulatory interaction with ethylene. *J Exp Bot.* 2010, 61(5), 1523–1535. doi: 10.1093/jxb/erq017. PMID: 20200120; PMCID: PMC2837265.
46. Elitzur T.; Yakir E.; Quansah L.; Fei Z.; Vrebalov J.; Khayat E.; Giovannoni J.J.; Friedman H. Banana MaMADS transcription factors are necessary for fruit ripening and molecular tools to promote shelf-life and food security. *Plant Physiol.* 2016, 171(1), 380–391. doi: 10.1104/pp.15.01866. PMID: 26956665; PMCID: PMC4854685.
47. Fu C.C.; Chen H.J.; Gao H.Y.; Wang S.L.; Han Y.C. Papaya CpMADS4 and CpNAC3 co-operatively regulate ethylene signal genes CpERF9 and CpEIL5 during fruit ripening. *Postharvest Biol. Technol.* 2021, 175, 111485. doi: 10.1016/j.postharvbio.2021.111485.
48. Lu S.; Zhang Y.; Zhu K.; Yang W.; Ye J.; Chai L.; Xu Q.; Deng X. The citrus transcription factor CsMADS6 modulates carotenoid metabolism by directly regulating carotenogenic genes. *Plant Physiol.* 2018, 176(4), 2657–2676. doi: 10.1104/pp.17.01830. PMID: 29463773; PMCID: PMC5884614.
49. Gomez-Lim M.A. Mango fruit ripening: physiology and molecular biology. *Acta Hort.* 1993, 341, 484–499. doi: 10.17660/ActaHortic.1993.341.52.
50. Wang P.; Luo Y.; Huang J.; Gao S.; Zhu G.; Dang Z.; Gai J.; Yang M.; Zhu M.; Zhang H.; Ye X.; Gao A.; Tan X.; Wang S.; Wu S.; Cahoon E.B.; Bai B.; Zhao Z.; Li Q.; Wei J.; Chen H.; Luo R.; Gong D.; Tang K.; Zhang B.; Ni Z.; Huang G.; Hu S.; Chen Y. The genome evolution and domestication of tropical fruit mango. *Genome Biol.* 2020, 21(1), 60. doi: 10.1186/s13059-020-01959-8. PMID: 32143734; PMCID: PMC7059373.
51. Eddy S.R. Accelerated profile HMM Searches. *PLoS Comput Biol.* 2011, 7(10), e1002195. doi: 10.1371/journal.pcbi.1002195. PMID: 22039361; PMCID: PMC3197634.
52. Chen C.; Chen H.; Zhang Y.; Thomas H.R.; Frank M.H.; He Y.; Xia R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol Plant.* 2020, 13(8), 1194–1202. doi: 10.1016/j.molp.2020.06.009. PMID: 32585190.
53. Horton P.; Park K.J.; Obayashi T.; Fujita N.; Harada H.; Adams-Collier C.J.; Nakai K. WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 2007, 35(Web Server issue), W585–587. doi: 10.1093/nar/gkm259. PMID: 17517783; PMCID: PMC1933216.
54. Letunic I.; Doerks T.; Bork P. SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res.* 2012, 40(Database issue), D302–305. doi: 10.1093/nar/gkr931. PMID: 22053084; PMCID: PMC3245027.
55. Lescot M.; Déhais P.; Thijs G.; Marchal K.; Moreau Y.; Van de Peer Y.; Rouzé P.; Rombauts S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 2002, 30(1), 325–327. doi: 10.1093/nar/30.1.325. PMID: 11752327; PMCID: PMC99092.

56. Li L.; Wu H.X.; Ma X.W.; Xu W.T.; Liang Q.Z.; Zhan R.L.; Wang S.B. Transcriptional mechanism of differential sugar accumulation in pulp of two contrasting mango (*Mangifera indica* L.) cultivars. *Genomics* 2020, 112(6), 4505–4515. doi: 10.1016/j.ygeno.2020.07.038. PMID: 32735916.
57. Pertea M.; Kim D.; Pertea G.M.; Leek J.T.; Salzberg S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, String Tie and Ballgown. *Nat Protoc.* 2016, 11(9), 1650–1667. doi: 10.1038/nprot.2016.095. PMID: 27560171; PMCID: PMC5032908.
58. Kofuji R.; Sumikawa N.; Yamasaki M.; Kondo K.; Ueda K.; Ito M.; Hasebe M. Evolution and divergence of the MADS-box gene family based on genome-wide expression analyses. *Mol Biol Evol.* 2003, 20(12), 1963–1977. doi: 10.1093/molbev/msg216. PMID: 12949148.
59. Lin H.; Zhu W.; Silva J.C.; Gu X.; Buell C.R. Intron gain and loss in segmentally duplicated genes in rice. *Genome Biol.* 2006, 7(5), R41. doi: 10.1186/gb-2006-7-5-r41. PMID: 16719932; PMCID: PMC1779517.
60. Davies B.; Egea-Cortines M.; de Andrade Silva E.; Saedler H.; Sommer H. Multiple interactions amongst floral homeotic MADS box proteins. *EMBO J.* 1996, 15(16), 4330–4343. PMID: 8861961; PMCID: PMC452158.
61. Lyons E.; Pedersen B.; Kane J.; Alam M.; Ming R.; Tang H.; Wang X.; Bowers J.; Paterson A.; Lisch D.; Freeling M. Finding and comparing syntenic regions among Arabidopsis and the outgroups papaya, poplar, and grape: CoGe with rosids. *Plant Physiol.* 2008, 148(4), 1772–1781. doi: 10.1104/pp.108.124867. PMID: 18952863; PMCID: PMC2593677.
62. Becker A.; Theissen G. The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol Phylogenet Evol.* 2003, 29(3), 464–489. doi: 10.1016/s1055-7903(03)00207-0. PMID: 14615187.
63. Irish V.F. The evolution of floral homeotic gene function. *Bioessays* 2003, 25(7), 637–646. doi: 10.1002/bies.10292. PMID: 12815719.
64. Schilling S.; Pan S.; Kennedy A.; Melzer R. MADS-box genes and crop domestication: the jack of all traits. *J Exp Bot.* 2018, 69(7), 1447–1469. doi: 10.1093/jxb/erx479. PMID: 29474735.
65. Arora R.; Agarwal P.; Ray S.; Singh A.K.; Singh V.P.; Tyagi A.K.; Kapoor S. MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. *BMC Genom.* 2007, 8, 242. doi: 10.1186/1471-2164-8-242. PMID: 17640358; PMCID: PMC1947970.
66. Wang Y.; Zhang J.; Hu Z.; Guo X.; Tian S.; Chen G. Genome-wide analysis of the MADS-box transcription factor family in *Solanum lycopersicum*. *Int J Mol Sci.* 2019, 20(12), 2961. doi: 10.3390/ijms20122961. PMID: 31216621; PMCID: PMC6627509.
67. Gao H.; Wang Z.; Li S.; Hou M.; Zhou Y.; Zhao Y.; Li G.; Zhao H.; Ma H. Genome-wide survey of potato MADS-box genes reveals that StMADS1 and StMADS13 are putative downstream targets of tuberigen StSP6A. *BMC Genom.* 2018, 19(1), 726. doi: 10.1186/s12864-018-5113-z. PMID: 30285611; PMCID: PMC6171223.
68. Raza Q.; Riaz A.; Atif R.M.; Hussain B.; Rana I.A.; Ali Z.; Budak H.; Alaraidh I.A. Genome-wide diversity of MADS-box genes in bread wheat is associated with its rapid global adaptability. *Front Genet.* 2022, 12, 818880. doi: 10.3389/fgene.2021.818880. PMID: 35111207; PMCID: PMC8801776.
69. Hu L.; Liu S. Genome-wide analysis of the MADS-box gene family in cucumber. *Genome* 2012, 55(3), 245–256. doi: 10.1139/g2012-009. PMID: 22376137.
70. Saha G.; Park J.I.; Jung H.J.; Ahmed N.U.; Kayum M.A.; Chung M.Y.; Hur Y.; Cho Y.G.; Watanabe M.; Nou I.S. Genome-wide identification and characterization of MADS-box family genes related to organ development and stress resistance in *Brassica rapa*. *BMC Genom.* 2015, 16(1), 178. doi: 10.1186/s12864-015-1349-z. PMID: 25881193; PMCID: PMC4422603.
71. Li C.; Wang Y.; Xu L.; Nie S.; Chen Y.; Liang D.; Sun X.; Karanja B.K.; Luo X.; Liu L. Genome-wide characterization of the MADS-box gene family in radish (*Raphanus sativus* L.) and assessment of its roles in flowering and floral organogenesis. *Front Plant Sci.* 2016, 7, 1390. doi: 10.3389/fpls.2016.01390. PMID: 27703461; PMCID: PMC5028395.
72. Wei X.; Wang L.; Yu J.; Zhang Y.; Li D.; Zhang X. Genome-wide identification and analysis of the MADS-box gene family in sesame. *Gene* 2015, 569(1), 66–76. doi: 10.1016/j.gene.2015.05.018. PMID: 25967387.
73. Wang P.; Wang S.; Chen Y.; Xu X.; Guang X.; Zhang Y. Genome-wide Analysis of the MADS-Box Gene Family in Watermelon. *Comput Biol Chem.* 2019, 80, 341–350. doi: 10.1016/j.compbiolchem.2019.04.013. PMID: 31082717.

74. Zhang Y.; Tang D.; Lin X.; Ding M.; Tong Z. Genome-wide identification of MADS-box family genes in moso bamboo (*Phyllostachys edulis*) and a functional analysis of PeMADS5 in flowering. *BMC Plant Biol.* 2018, 18(1), 176. doi: 10.1186/s12870-018-1394-2. PMID: 30176795; PMCID: PMC6122543.
75. Zhao Y.; Li X.; Chen W.; Peng X.; Cheng X.; Zhu S.; Cheng B. Whole-genome survey and characterization of MADS-box gene family in maize and sorghum. *Plant Cell Tissue Organ Cult.* 2011, 105(2), 159–173. doi: 10.1007/s11240-010-9848-8.
76. Tian Y.; Dong Q.; Ji Z.; Chi F.; Cong P.; Zhou Z. Genome-wide identification and analysis of the MADS-box gene family in apple. *Gene* 2015, 555(2), 277–290. doi: 10.1016/j.gene.2014.11.018. PMID: 25447908.
77. Won S.Y.; Jung J.A.; Kim J.S. Genome-wide analysis of the MADS-Box gene family in *Chrysanthemum*. *Comput Biol Chem.* 2021, 90, 107424. doi: 10.1016/j.compbiolchem.2020.107424. PMID: 33340990.
78. Chen Q.; Li J.; Yang F. Genome-wide analysis of the MADS-box transcription factor family in *Solanum melongena*. *Int J Mol Sci.* 2023, 24(1), 826. doi: 10.3390/ijms24010826. PMID: 36614267; PMCID: PMC9821028.
79. Wang X.; Huang Q.; Shen Z.; Baron G.C.; Li X.; Lu X.; Li Y.; Chen W.; Xu L.; Lv J.; Li W.; Zong Y.; Guo W. Genome-wide identification and analysis of the MADS-Box transcription factor genes in blueberry (*Vaccinium* spp.) and their expression pattern during fruit ripening. *Plants (Basel)* 2023, 12(7), 1424. doi: 10.3390/plants12071424. PMID: 37050050; PMCID: PMC10096547.
80. Cheng S.; Jia M.; Su L.; Liu X.; Chu Q.; He Z.; Zhou X.; Lu W.; Jiang C. Genome-wide identification of the MADS-box gene family during male and female flower development in chayote (*Sechium edule*). *Int J Mol Sci.* 2023, 24(7), 6114. doi: 10.3390/ijms24076114. PMID: 37047083; PMCID: PMC10094161.
81. Leseberg C.H.; Li A.; Kang H.; Duvall M.; Mao L. Genome-wide analysis of the MADS-box gene family in *Populus trichocarpa*. *Gene* 2006, 378, 84–94. doi: 10.1016/j.gene.2006.05.022. PMID: 16831523.
82. Wang R.; Ming M.; Li J.; Shi D.; Qiao X.; Li L.; Zhang S.; Wu J. Genome-wide identification of the MADS-box transcription factor family in pear (*Pyrus bretschneideri*) reveals evolution and functional divergence. *Peer J.* 2017, 5, e3776. doi: 10.7717/peerj.3776. PMID: 28924499; PMCID: PMC5598432.
83. Liu J.; Zhang J.; Zhang J.; Miao H.; Wang J.; Gao P.; Hu W.; Jia C.; Wang Z.; Xu B.; Jin Z. Genome-wide analysis of banana MADS-box family closely related to fruit development and ripening. *Sci Rep.* 2017, 7(1), 3467. doi: 10.1038/s41598-017-03897-1. PMID: 28615681; PMCID: PMC5471262.
84. Díaz-Riquelme J.; Lijavetzky D.; Martínez-Zapater J.M.; Carmona M.J. Genome-wide analysis of MIKCC-type MADS box genes in grapevine. *Plant Physiol.* 2009, 149(1), 354–369. doi: 10.1104/pp.108.131052. PMID: 18997115; PMCID: PMC2613716.
85. Shu Y.; Yu D.; Wang D.; Guo D.; Guo C. Genome-wide survey and expression analysis of the MADS-box gene family in soybean. *Mol Biol Rep.* 2013, 40(6), 3901–3911. doi: 10.1007/s11033-012-2438-6. PMID: 23559340.
86. Lin Y.; Qi X.; Wan Y.; Chen Z.; Fang H.; Liang C. Genome-wide analysis of the MADS-box gene family in *Lonicera japonica* and a proposed floral organ identity model. *BMC Genom.* 2023, 24(1), 447. doi: 10.1186/s12864-023-09509-9. PMID: 37553575; PMCID: PMC10408238.
87. Zhao D.; Chen Z.; Xu L.; Zhang L.; Zou Q. Genome-wide analysis of the MADS-Box gene family in maize: gene structure, evolution, and relationships. *Genes (Basel)* 2021, 12(12), 1956. doi: 10.3390/genes12121956. PMID: 34946905; PMCID: PMC8701013.
88. Sun F.; Fang H.; Wen X.; Zhang L. Phylogenetic and expression analysis of MADS-box gene family in *Rhododendron ovatum*. *Chinese Bulletin of Botany* 2023, 58(3), 404–416. doi: 10.11983/CBB22105.
89. Gramzow L.; Theißen G. Phylogenomics reveals surprising sets of essential and dispensable clades of MIKC(c)-group MADS-box genes in flowering plants. *J Exp Zool B Mol Dev Evol.* 2015, 324(4), 353–362. doi: 10.1002/jez.b.22598. PMID: 25678468.
90. Wang L.; Yin X.; Cheng C.; Wang H.; Guo R.; Xu X.; Zhao J.; Zheng Y.; Wang X. Evolutionary and expression analysis of a MADS-box gene superfamily involved in ovule development of seeded and seedless grapevines. *Mol Genet Genomics.* 2015, 290(3), 825–846. doi: 10.1007/s00438-014-0961-y. PMID: 25429734.
91. Wang B.; Hu W.; Fang Y.; Feng X.; Fang J.; Zou T.; Zheng S.; Ming R.; Zhang J. Comparative analysis of the MADS-Box genes revealed their potential functions for flower and fruit development in longan (*Dimocarpus longan*). *Front Plant Sci.* 2022, 12, 813798. doi: 10.3389/fpls.2021.813798. PMID: 35154209; PMCID: PMC8829350.
92. Zhao W.; Zhang L.L.; Xu Z.S.; Fu L.; Pang H.X.; Ma Y.Z.; Min D.H. Genome-Wide analysis of MADS-Box genes in foxtail millet (*Setaria italica* L.) and functional assessment of the role of SiMADS51 in the drought

- stress response. *Front Plant Sci.* 2021, 12, 659474. doi: 10.3389/fpls.2021.659474. PMID: 34262576; PMCID: PMC8273297.
93. Alhindi T.; Al-Abdallat A.M. Genome-wide identification and analysis of the MADS-box gene family in American beautyberry (*Callicarpa americana*). *Plants (Basel)* 2021, 10(9), 1805. doi: 10.3390/plants10091805. PMID: 34579338; PMCID: PMC8466759.
  94. Guan H.; Wang H.; Huang J.; Liu M.; Chen T.; Shan X.; Chen H.; Shen J. Genome-wide identification and expression analysis of MADS-Box family genes in litchi (*Litchi chinensis* Sonn.) and their involvement in floral sex determination. *Plants (Basel)* 2021, 10(10), 2142. doi: 10.3390/plants10102142. PMID: 34685951; PMCID: PMC8540616.
  95. Gramzow L.; Ritz M.S.; Theissen G. On the origin of MADS-domain transcription factors. *Trends Genet.* 2010, 26(4), 149–153. doi: 10.1016/j.tig.2010.01.004. PMID: 20219261.
  96. Qu Y.; Kong W.; Wang Q.; Fu X. Genome-wide identification MIKC-Type MADS-Box gene family and their roles during development of floral buds in wheel wingnut (*Cyclocarya paliurus*). *Int J Mol Sci.* 2021, 22(18), 10128. doi: 10.3390/ijms221810128. PMID: 34576289; PMCID: PMC8471257.
  97. Huo S.; Li Y.; Li R.; Chen R.; Xing H.; Wang J.; Zhao Y.; Song X. Genome-wide analysis of the MADS-box gene family in *Rhododendron hainanense* Merr. and expression analysis under heat and waterlogging stresses. *Ind Crops Prod.* 2021, 172, 114007. doi: 10.1016/j.indcrop.2021.114007.
  98. Liu M.; Fu Q.; Ma Z.; Sun W.; Huang L.; Wu Q.; Tang Z.; Bu T.; Li C.; Chen H. Genome-wide investigation of the MADS gene family and dehulling genes in tartary buckwheat (*Fagopyrum tataricum*). *Planta* 2019, 249(5), 1301–1318. doi: 10.1007/s00425-019-03089-3. PMID: 30617544.
  99. Liu J.; Liu M.; Wang J.; Zhang J.; Miao H.; Wang Z.; Jia C.; Zhang J.; Xu B.; Jin Z. Transcription factor MaMADS36 plays a central role in regulating banana fruit ripening. *J Exp Bot.* 2021, 72(20), 7078–7091. doi: 10.1093/jxb/erab341. PMID: 34282447.
  100. Wang Y.; Nie F.; Shahid M.Q.; Baloch F.S. Molecular footprints of selection effects and whole genome duplication (WGD) events in three blueberry species: detected by transcriptome dataset. *BMC Plant Biol.* 2020, 20(1), 250. doi: 10.1186/s12870-020-02461-w. PMID: 32493212; PMCID: PMC7268529.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.