horticulturae

MDPI

Article

Genome-Wide Classification and Evolutionary Analysis of the *KNOX* Gene Family in Plants

Zhi-Meng Gan [1], Xiao-Yan Ai [2], Chun-Gen Hu [1,*] and Jin-Zhi Zhang [1,*]

[1] National Key Laboratory for Germplasm Innovation & Utilization of Horticultural Crops, College of Horticulture and Forestry Science, Huazhong Agricultural University, Wuhan 430070, China; zhimenggan@hbaas.com

[2] Institute of Pomology and Tea, Hubei Academy of Agricultural Sciences, Wuhan 430070, China; aixiaoyan@hbaas.com

* Correspondence: chungen@mail.hzau.edu.cn (C.-G.H.); jinzhizhang@mail.hzau.edu.cn (J.-Z.Z.); Tel.: +86-27-6201-80231 (J.-Z.Z.); Fax: +86-27-8728-2010 (J.-Z.Z.)

Abstract: The *Knotted1-like homeobox* (*KNOX*) gene family plays a pivotal role in regulating meristem activity, organ differentiation, and cell meristematic identity. However, there has been a lack of large-scale, systematic, and comprehensive comparative analyses to explore their expression patterns and evolutionary mechanisms. In this study, a total of 1425 *KNOX* genes were identified across 118 plant species. The result showed that higher plants exhibited a significantly higher abundance of *KNOX* genes compared to lower plants. Phylogenetic analysis revealed that all *KNOX* genes can be divided into two classes (class I and II) and evolved independently after species differentiation. An analysis of gene duplication or loss showed that gene loss was more common than gene duplication in lower plants within the *KNOX* gene family. These findings suggest that gene loss in the *KNOX* gene family occurs after events such as whole-genome duplication (WGD) or whole-genome triplication (WGT). In addition, conserved motif analysis was also conducted to uncover the evolutionary trajectories of *KNOX* genes. We found that three motifs (M1, M2, and M4) were present in nearly all *KNOX* genes, while four novel motifs (M7–M10) were lost in lower plants but present in higher plants. Moreover, the loss of certain motifs in the *KNOX* genes was also observed in higher plants, indicating sequence divergence in *KNOX* genes throughout evolution. To understand the expression patterns of *KNOX* genes, a gene expression pattern analysis was performed in *A. thaliana* and *O. sativa*. The results showed that class I *KNOX* genes exhibit conserved high expression in stems, suggesting their potential similar biological roles across different plant species and the conservation of their functions during evolution. Additionally, we analyzed the *KNOX* genes in the Citrus genus and closely related species, and we found that the number of *KNOX* genes evolved at a slower rate in these species, indicating a relatively conservative nature. In conclusion, this study provides valuable resources for the investigation of *KNOX* gene family evolution and function in plants.

Keywords: *Knotted1-like homeobox* family; evolutionary analysis; gene duplication and loss; conserved motif analysis; expression analysis

Citation: Gan, Z.-M.; Ai, X.-Y.; Hu, C.-G.; Zhang, J.-Z. Genome-Wide Classification and Evolutionary Analysis of the *KNOX* Gene Family in Plants. *Horticulturae* **2023**, *9*, 1174. https://doi.org/10.3390/horticulturae9111174

Academic Editor: Radu E. Sestras

Received: 20 September 2023
Revised: 23 October 2023
Accepted: 25 October 2023
Published: 27 October 2023

1. Introduction

The *KNOX* gene family is one of the plant-specific homeobox transcription factors, belonging to the TALE (three amino acid loop extension) subclass, and it plays a vital role in leaf morphogenesis and shoot apical meristem (SAM) development [1–3]. Previous reports have shown that *KNOX* genes are very conserved and mainly contain four conserved domains (KNAT1, KNAT2, ELK, and Homeobox KN domain) [4,5]. The members of the *KNOX* gene family are divided into two classes based on their similarity and expression pattern, class I or class II [6]. In Arabidopsis, class I *KNOX* genes such as *SHOOT MERISTEMLESS* (*STM*), *BREVI-PEDICELLUS* (*BP* or *KNAT1*), *KNAT2*, and *KNAT6* exhibit specific expression in meristems. These genes play a significant role in regulating plant organ

development and shaping plant morphology. Among the class II *KNOX* genes are *KNAT3*, *KNAT4*, *KNAT5*, and *KNAT7*, which are expressed in the tissues of all plants. Currently there are few studies on these genes and their functions in plants are not yet clear.

The first *KNOX* gene (*KN1* or *ZmKNAT1*) was cloned and characterized in *Zea mays* [7,8]. Until now, the *KNOX* gene has been characterized in most species, such as 8 members in *A. thaliana* [6], 23 members in *Brassica juncea* [9], 20 members in *Brassica napus* [10], 9 members in *Brassica oleracea* [11], 12 members in *Brassica rapa* [9], 8 members in *Cardamine hirsute* [12], 8 members in *Citrus clementina* [13], 11 members in *Cucumis sativus* [14], 8 members in *Fragaria vesca* [15], 28 members in *Glycine max* [16], 8 members in *Lactuca sativa* [17], 17 members in *Malus domestica* [18], 10 members in *Medicago truncatula* [19], 7 members in *Prunus persica* [20], 7 members in *Solanum lycopersicum* [21], 8 members in *Solanum tuberosum* [22], 9 members in *Triticum aestivum* [23], and 9 members in *Vitis vinifera* [24]. These studies will provide a rich resource for further analysis of the function and evolution of the *KNOX* family gene in plants.

Currently, the functions of *KNOX* genes in model plants have been widely reported, such as the maintenance of apical meristems [25], leaf morphogenesis [26], stem elongation [27,28], and floral organ development [29,30]. In *Arabidopsis*, *STM* is expressed in the shoot meristem and interacts with CLAVATA3 (CLV3) to jointly maintain stem cell homeostasis [27,28]. Additionally, *STM* can prevent cell differentiation of in the meristems by maintaining the expression of *WUSCHEL* (*WUS*) [31], indicating that *STM* is essential for stem cell maintenance in the SAM. In terms of leaf morphogenesis, the abnormal expression of *KNOX* genes causes maize leaves to shrink and become "knotted" [7]. In plants with compound leaves, such as tomatoes, the class I *KNOX* genes *Tkn1* and *Tkn2/LeT6* are expressed in leaf primordia. When *tkn1* and *tkn2/let6* are mutated, it results in accelerated leaf differentiation and reduced leaf complexity [12,32]. In both *M. truncatula* and *Cardamine hirsuta*, KNOXI proteins are also necessary and sufficient for the formation of leaflets [33]. The ectopic expression of class I *KNOX* members can similarly enhance leaf complexity, indicating that class I *KNOX* genes play a role in regulating leaf development. For instance, *PagKNAT2/6b* downregulates the expression of *GA20ox1* in poplar, inhibiting cell elongation and expansion, leading to a dwarf phenotype [34]. Similarly, similar phenotypes have been observed in *A. thaliana* [28], *Oryza sativa* [26], and *Z. mays* [35], indicating that class I KNOX genes play an important role in regulating plant cell elongation and expansion. Class I *KNOX* family genes also play a crucial role in floral morphogenesis. The ectopic expression of *STM* in *Arabidopsis* promotes the transformation of homologous ovules into carpels, with a reduction in the number of petals and other floral organs in *stm* mutants [36,37]. Furthermore, *STM* may interact with *APETALA1* (*AP1*) during the regulation of the development of floral organs [38]. *KNAT1* inhibits *KNAT2/KNAT6* to ensure the normal upward orientation of the pedicel, while the ectopic expression of *KNAT2* and *KNAT6* in *knat1* null mutants results in an abnormal inflorescence structure [29].

So far, there have been few comparative and evolutionary analysis studies across multiple different species. There are related transcription factor databases, such as the Plant Transcription Factor Database (PlantTFDB, http://planttfdb.gao-lab.org/, accessed on 12 July 2023) [39], which can query and analyze *KNOX* family genes. However, limited genomes, low quality, and the inability to perform targeted updates affect the comprehensiveness of identification. With the development of sequencing technology, an increasing number of genomes have been updated or assembled with higher quality, making it possible to identify the whole genome of *KNOX* genes in more species. Here, *KNOX* genes from 118 plants were comprehensively identified. Subsequently, we further conducted comparative analysis, gene duplication and loss analysis, conserved motif analysis, collinearity analysis, replication type identification, expression pattern analysis, and pan-genome analysis on *KNOX* genes.

2. Materials and Methods

2.1. Data Collection

In this study, a total of 118 genomes were selected, including 2 Rhodophyta, 8 Chlorophyta, 1 basal plant, and 107 land plants (2 basal angiosperms, 26 monocots, 74 eudicots and 5 non-angiosperms) (Table S1). It is noteworthy that 60 of these land plants are horticultural plants, including 23 fruit trees (*Actinidia chinensis*, *Amborellatrichopoda*, *Ananas comosus*, *Arachis ipaensis*, *Citrullus lanatus*, *C. clementina*, *Citrus maxima*, *Citrus sinensis*, *Cucumis melo*, *Diospyros lotus*, *Ficus carica*, *F. vesca*, *M. domestica*, *Mangifera indica*, *Musa acuminata*, *Musa acuminata subsp*, *Pistacia vera*, *Prunus avium*, *Prunus dulcis*, *Prunus persica*, *Punica granatum*, *Vitis riparia*, *V. vinifera*), 24 vegetables (*Aquilegia coerulea*, *Arachis duranensis*, *Asparagus officinalis*, *Beta vulgaris*, *B. juncea*, *B. napus*, *B. oleracea*, *B. rapa*, *Capsicum annuum*, *Carica papaya*, *Cicer arietinum*, *Cucumis sativus*, *Daucus carota*, *G. max*, *Juglans regia*, *L. sativa*, *Manihot esculenta*, *Olea europaea*, *Phaseolus vulgaris*, *Pisum sativum*, *S. dulcamara*, *S. lycopersicum*, *Solanum pennellii*, *S. tuberosum*), and 13 other horticultural plants (*Elaeis guineensis*, *Eucalyptus grandis*, *Gossypium arboreum*, *Ipomoea triloba*, *Kalanchoe laxiflora*, *Lotus japonicus*, *M. truncatula*, *Nymphaea colorata*, *Picea abies*, *Prunus mume*, *Rosa chinensis*, *Salix purpurea*, *Salvia miltiorrhiza*).

The genome data and annotation information for *A. thaliana* were downloaded from the Arabidopsis Information Resource (http://www.arabidopsis.org/, accessed on 5 August 2023), data for *O. sativa* were downloaded from the Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/, accessed on 5 August 2023), data for 67 species were downloaded from the EnsemblPlants database (https://ftp.ebi.ac.uk/ensemblgenomes/pub/release-57/plants/, accessed on 5 August 2023), data for 30 species were downloaded from the NCBI Refseq database (https://ftp.ncbi.nlm.nih.gov/genomes/refseq/plant/, accessed on 5 August 2023), and the other 19 species were downloaded from the Phytozome database (https://phytozome-next.jgi.doe.gov/, accessed on 5 August 2023). The 17 pan-genome sequences of Citrinae species were downloaded from the CPBD database (http://citrus.hzau.edu.cn/, accessed on 10 August 2023) [40].

The transcriptome data of *A. thaliana* and *O. sativa* were downloaded from the Sequence Read Archive (SRA, https://www.ncbi.nlm.nih.gov/sra/, accessed on 12 August 2023) The dataset for *A. thaliana* consisted of 94 root samples, 24 stem samples, 315 leaf samples, 133 flower samples, 36 seed samples, 60 embryo samples, and 370 seedling samples. Furthermore, we integrated 281 samples that underwent hormone treatments, with 38 samples treated with IAA, 113 samples with ABA, 21 samples with GA, and 109 samples with JA. The dataset for *O. sativa* encompassed 487 samples obtained from various tissues of rice, along with 48 samples treated with different hormone conditions. Specifically, the rice dataset comprised 125 root samples, 7 stem samples, 242 leaf samples, 8 flower samples, 56 seed samples, 7 embryo samples, and 42 seedling samples.

2.2. Identification of KNOX Gene

The *KNOX* genes were identified by using the KNOX protein sequences of *A. thaliana* [6] and *O. sativa* [41] as the query (BLASTP, -perc_identit y > 50 and e-value < 1×10^{-10}) and the hidden Markov model file of the KNOX1 domain (PF03790), KNOX2 domain (PF03791), ELK domain (PF03789), and Homeobox KN domain (PF05920) as probes (HMMER3.0). Furthermore, the CDD databases (https://www.ncbi.nlm.nih.gov/cdd/, accessed on 15 August 2023), Pfam databases (http://pfam.xfam.org/, accessed on 15 August 2023) and SMART (http://smart.embl.de/, accessed on 15 August 2023) databases were used to ensure the result accuracy.

2.3. Multiple Sequence Alignment and Phylogenetic Analyses

The phylogenetic tree was made according to the relationship between the species according NCBI taxonomy and plotted by using the R package, ggtree [42]. The phylogenetic tree was constructed by using Mafft (v7.471) [43] for the protein sequences and using IQ-tree (v2.1.11) [44] (the LG and JTT model was used, and bootstrap replications

were set at 1000) for the maximum likelihood tree construction and plotted using the R package, ggtree.

2.4. Gene Duplication and Loss Inference

Gene duplication and loss were identified using the phylogenetic species of trees and genes using the java software, Notung (v2.9) [45]. All of this information on the phylogenetic trees was illustrated using the R package, ggtree.

2.5. Collinearity Analysis and Replication Type Identification of KNOX Gene

The collinearity of the *KNOX* gene among seven representative plants was detected using MCScanX software [46]. Firstly, the protein sequences of these species were aligned using the blastp program with an e-value of 10^{-5}. Then, the collinear blocks were detected using MCScanX with the default parameters. Finally, the gene duplication types (WGD/segmental, tandem, singleton, dispersed, proximal duplications, and singleton) were identified using a duplicate gene classifier program from the MCScanX software. A significance analysis of the duplication type for the *KNOX* genes compared with the whole-genome genes was conducted using the χ^2 test ($p < 0.01$).

2.6. Expression Pattern Analysis of the KNOX Gene

We conducted a search in databases such as the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/, accessed on 11 August 2023), the Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra/, accessed on 11 August 2023), the European Nucleotide Archive (https://www.ebi.ac.uk, accessed on 11 August 2023), and the DNA Data Bank (https://www.ddbj.nig.ac.jp/ddbj/, accessed on 11 August 2023) of Japan to retrieve samples of different tissues and hormone treatments in *Arabidopsis* and rice. Subsequently, we downloaded the raw data from all libraries and performed the analysis using a standardized pipeline (fastp–STAR–featureCounts). Ultimately, we obtained normalized FPKM (fragments per kilobase of transcript per million mapped reads) values in Araport11 for each library. Afterwards, the data were further processed and analyzed using Python. Grouping and statistical analysis were performed to gain insights into the data. Additionally, visualizations such as box plots and heatmaps were generated.

2.7. Nonsynonymous (Ka) and Synonymous (Ks) Substitution Rate Calculation Analysis

The calculation of non-synonymous (Ka) and synonymous (Ks) substitution rates for the *KNOX* gene pairs was performed using a model averaging approach and the Ka/Ks Calculator software [47]. This software enabled the estimation of Ka and Ks substitution rates, which were subsequently used to calculate the Ka/Ks ratio. Python was then utilized for grouping the data and generating visualizations.

3. Results

3.1. Identifying the KNOX Gene in 118 Plants

A total of 1425 *KNOX* genes were identified in 118 plants (Figure 1; Table S1). Among these species there are 2 Rhodophyta, 8 Chlorophyta, 1 basal plant and 107 land plants, which are further divided into 26 monocots, 74 eudicots, 2 basal angiosperms, and 5 non-angiosperms (*Amborella trichopoda*, *Nymphaea colorata*, *Marchantia polymorpha*, *Physcomitrium patens*, *Picea abies*, *Selaginella moellendorffii*, *Sphagnum fallax*). Among these land plants, 60 are horticultural plants, 23 are fruit trees, 24 are vegetables, and 13 are other horticultural plants (Figure 1 and Table S1).

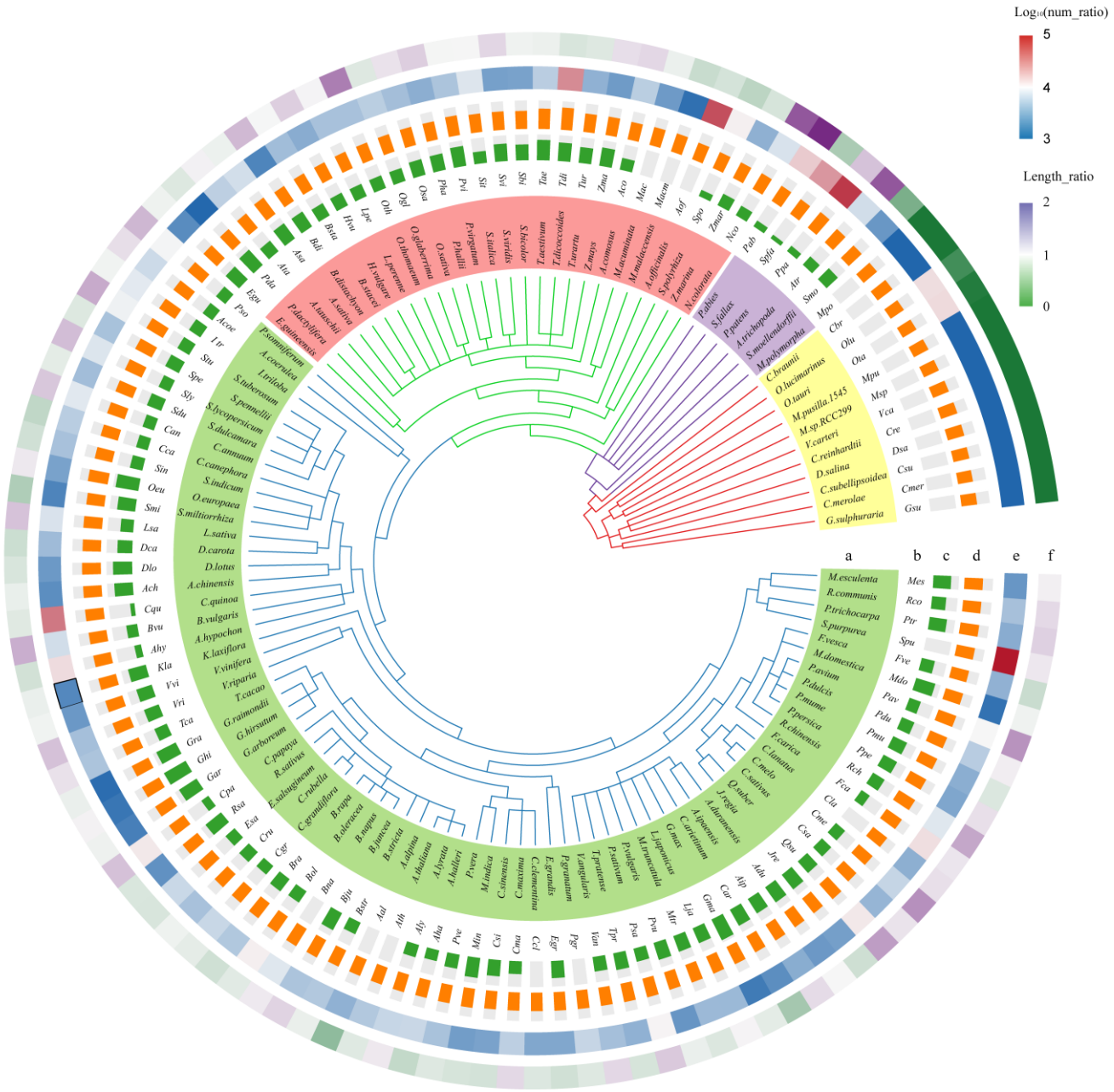

Figure 1. Comparative analysis of *KNOX* genes in 118 plants. (a) The abbreviated Latin name of each species. (b) The base classification of each species. (c) The Log_{10} (number of all genes) in each species. (d) The Log_2 (number of *KNOX* genes) in each species. (e) The Log_{10} (number ratio of *KNOX* genes) in each species. (f) The length ratio of *KNOX* genes in each species.

Statistical analysis of the identified *KNOX* genes found that the average number of *KNOX* genes was 13.3 and most species (81, 68.64%) had >8 KNOX genes compared to *A. thaliana* (8) (Figures 1 and S1a). Subsequently, we further calculated the length ratio of the *KNOX* gene (average length of *KNOX* genes/average length of all genes); the results showed that the length ratio of the *KNOX* gene was 55 (55.61%), between 0.8 and 1.2, which is greater than >1.2 (51, 46.22%) (Figures 1 and S1b).

3.2. Comparative Analysis of the KNOX Gene in 118 Plants

Investigating the evolution of plant gene families is of great significance for exploring the process of plant evolution and morphological development [48,49]. The statistical

analysis found that the number of *KNOX* genes in higher plants is significantly higher than in lower plants (Figures 1 and 2a). Among the top 10 species with a higher proportion of *KNOX* genes, all species belong to higher plants, including seven eudicots and three monocots (Figure 2b and Table S1). Monocot species include *P. dactylifera* (date palm), *M. acuminata* (banana), and *B. distachyon* (brachypodium). Interestingly, the three species with the highest proportion of *KNOX* genes belong to the Malvaceae family and include *G. raimondii*, *G. hirsutum*, and *G. arboretum* (Figure 2b and Table S1). These results indicate that the leaf morphology and development of the Malvaceae family may exhibit higher diversity and complexity.

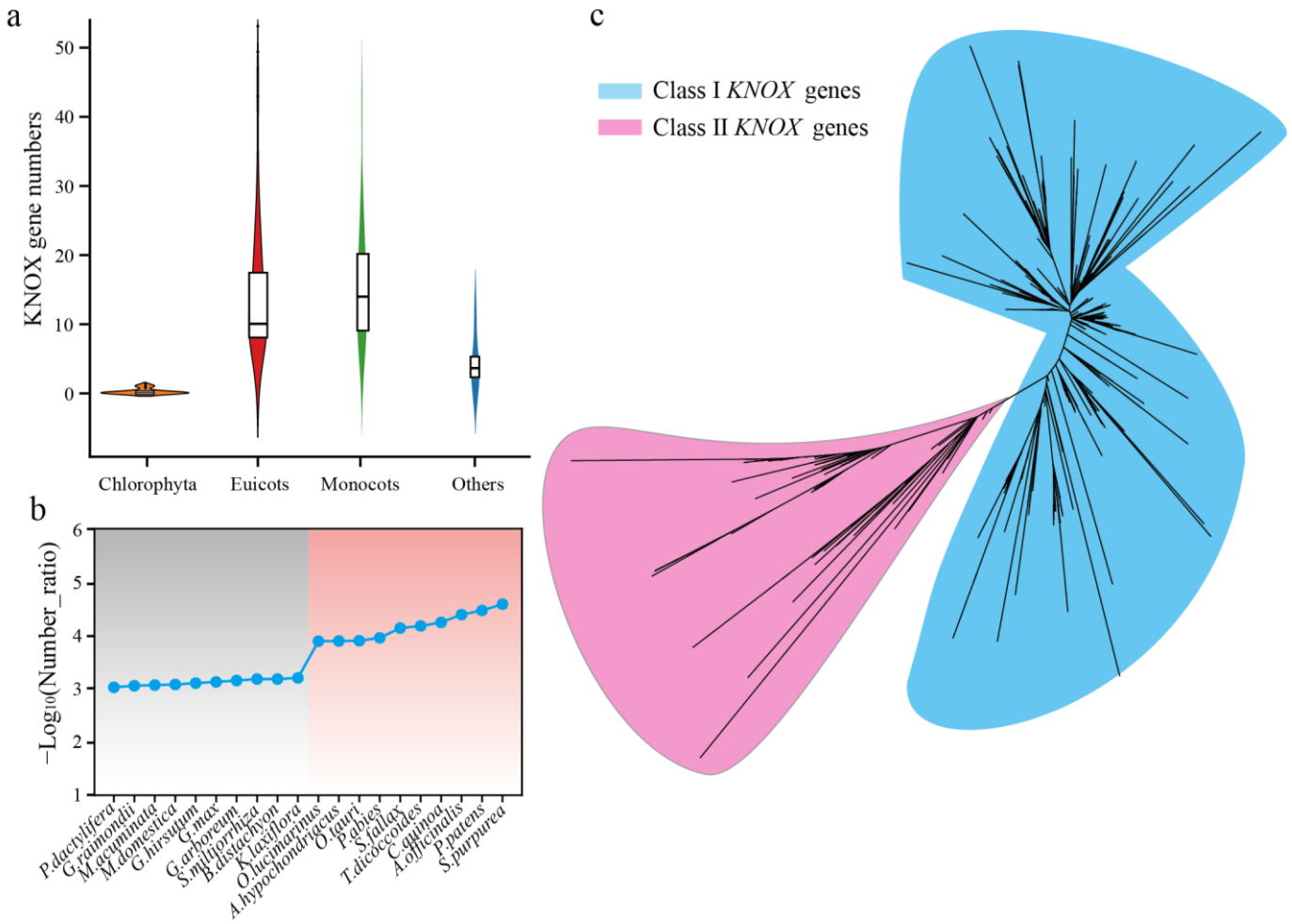

Figure 2. Phylogenetic analysis of *KNOX* genes in plants. (**a**) Boxplot of *KNOX* genes for four categories of plants (Rhodophyta and Chlorophyta, Eudicots, Monocots, and other plants). (**b**) Number ratio of *KNOX* genes in representative plants. The gray gradient indicates the 10 plants with the highest ratio of *KNOX* genes, and the orange gradient indicates the 10 plants with the lowest ratio of *KNOX* genes. The *Y*-axis represents the -log10 ratio of the *KNOX* gene. (**c**) Phylogenetic tree using 1425 *KNOX* proteins from 118 plants. Cyan and magenta represent class I and class II *KNOX* genes, respectively.

Most of the lower plants did not have any detected *KNOX* genes, and only one *KNOX* gene was detected in two Ostreococcus species, including *O. lucimarinus* and *O. tauri* (Table S1). Among the eight other species with a lower proportion of *KNOX* genes, were four eudicots, one monocot, and three other higher plants (Figure 2b). The only monocotyledon was *T. dicoccoides* (emmer), an allopolyploid plant with a total of 295,286 genes (Table S1). This result implies that the *KNOX* gene did not increase linearly with

genome doubling, it is speculated that there may be some negative regulatory mechanism in *T. dicoccoides* that inhibits the replication of the *KNOX* gene or promotes the deletion of gene family members, which inhibits the expansion of the *KNOX* gene family. Additionally, among the three other higher plants, *P. abies* (Norway spruce), *S. fallax* (Peat moss), and *P. patens* (a moss) have the lowest proportion of *KNOX* gene family members. In *P. patens*, only three *KNOX* genes were detected out of the total 86,669 genes in the whole genome, accounting for only 0.0035% of all genes (Figure 1 and Table S1). These results suggest that these ancient plant species may utilize alternative approaches to regulate the leaf morphology and development instead of relying on *KNOX* genes.

To investigate the evolution of the *KNOX* gene in plants, we constructed an interspecies phylogenetic tree using 1425 *KNOX* proteins from 118 plant species (Figure 2c). The results show that all KNOX proteins were grouped into two clades (class I and class II), and, corresponding to previous reports [5], the number of class I KNOX proteins was significantly higher than the number in class II. This shows that the evolution process of *KNOX* genes is relatively conservative. Moreover, most branches contain *KNOX* genes from different plant classifications. This result indicates that the KNOX proteins of each branch evolved independently after species divergence. Also, *KNOX* genes are clearly separated in some clades into monocot and eudicot species. These results suggest that the *KNOX* genes may generate sequence variation during evolution.

3.3. Gene Duplication and Loss Analysis of the KNOX Gene

Gene duplication or loss is prevalent in plant evolution and plays an important role in WGD (whole-genome duplication) or WGT (whole-genome triploidy) events [50]. To elucidate the evolution of the *KNOX* gene, we selected 17 representative plants for gene duplication and loss analysis, including two Chlorophyta (*O. lucimarinus*, *O. tauri*), two Bryophyta (*S. fallax*, *P. patens*), one Lycopodiophyta (*S. moellendorffii*), one Gymnospermae (*P. abies*), one basal Angiospermae (*A. trichopoda*), five eudicots (*A. thaliana*, *C. clementina*, *F. vesca*, *M. domestica*, and *V. vinifera*), and five monocots (*M. acuminata*, *O. sativa*, *P. dactylifera*, *Z. mays*, and *Z. marina*) (Figure 3a). To obtain the gene duplication and loss of *KNOX* genes during the evolutionary process, we used the Notung software to reconstruct the species tree and gene tree of these 17 representative plants and 20 typical dicotyledonous plants.

In the most recent common ancestor (MRCA) of these 17 representative plants, 29 genes underwent duplication without any gene loss (Figure 3a). However, in the common ancestors of two green algae (*O. lucimarinus* and *O. tauri*) and two bryophyte phyla (*S. fallax* and *P. patens*), gene loss occurred, but no gene duplication was found, and they lost 24 and 25 genes, respectively. In the common ancestors of mosses and other higher plants, 28 genes were duplicated and only one gene was lost. Although most WGD and WGT events occur in most higher plants, the gene loss in the *KNOX* gene family was more prevalent than duplications in *S. moellendorffii* (Lycopodiophyta), *P. abies* (Gymnospermae) and *A. trichopoda* (basal Angiospermae) (Figure 3a). These results indicated that the loss of the *KNOX* gene family may be involved in the processes of WGD or WGT during the plant evolution process.

In addition, we further analyzed the duplication and loss of *KNOX* genes in 20 typical dicotyledonous plants (Figure 3b). In the MRCA of these 20 plants, 141 genes underwent duplication without any gene loss, and in 16 species, the losses of the *KNOX* genes were more prevalent than duplications. This result further suggests that *KNOX* gene loss occurs after WGD or WGT events in most eudicots.

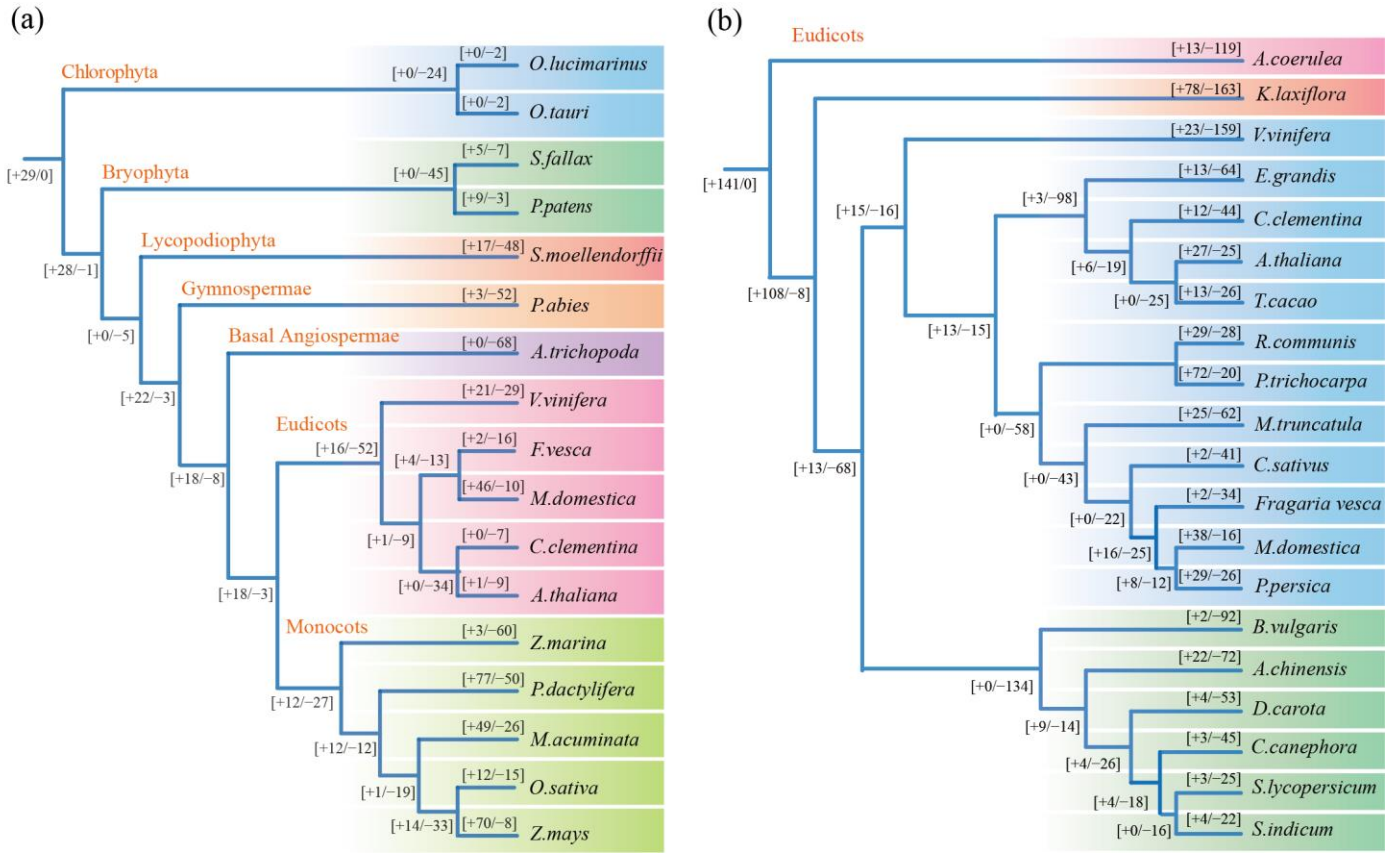

Figure 3. Gene duplication and loss analysis of *KNOX* gene. (**a**). Gene duplication and loss analysis of the *KNOX* gene in 17 representative plants. (**b**). Gene duplication and loss analysis of the *KNOX* gene in 20 represent eudicots. "+" and "−", respectively represent gene duplication and loss, respectively, while the numerical values represent the number of genes.

3.4. Conserved Motifs Analysis of the KNOX Gene

Conserved motif analysis can reveal the conserved patterns of a gene family [51]. Here, we explored the conserved motif of the *KNOX* gene in seven representative plants, ranging from *O. tauri* (lower plants) to the higher plants, including the Bryophyta species (*P. patens*), the Lycopodiophyta species (*S. moellendorffii*), the Gymnospermae species (*P. abies*), the basal angiosperm species (*A. trichopoda*), the monocot species (*Z. mays* and *O. sativa*), and the eudicot species (*C. sinensis* and *A. thaliana*) (Figure 4a and Table S1).

In total, 66 *KNOX* genes were identified from the whole genomes of nine species. The largest number of *KNOX* genes was detected in *Z. mays* (13), followed by *S. moellendorffii* (12), and *O. sativa* (11) (Figure 4a and Table S1). However, there was only one *KNOX* gene in the lower plant (*O. tauri*). To explore the phylogenetic relationship and classification of *KNOX* genes, a maximum likelihood tree was constructed using protein sequences. The results revealed that all of the *KNOX* genes could be divided into two groups, which we defined as class I and class II (Figure 4a).

Furthermore, we detected the motifs (motifs (M) 1–10) of the *KNOX* genes using the MEME program (Figure 4a and Table S2). A total of 479 motifs were detected with an average of 7.3 motifs per gene, and most genes (58, 85.29%) had >4 motifs. M4 was the longest, followed by M1 and M2. Three motifs (M1, M2, and M4) were present in almost all *KNOX* genes. However, some of these motifs were also absent. Furthermore, the Pfam analysis showed that the M1 motif was also partial in *S. moellendorffii* and *A. trichopoda* (Figure 4a). In addition, we also found that the M5 motif was present in almost all class I genes, while it was absent in almost all of the class II group. Similarly, M6 and M10 were only detected in the *KNOX* genes of class II and class I. These motifs might be associated

with the functional specificity of a different class of *KNOX* genes. These findings suggest that the motifs within the same class exhibited a high degree of similarity, aligning with the phylogenetic relationship among these genes.

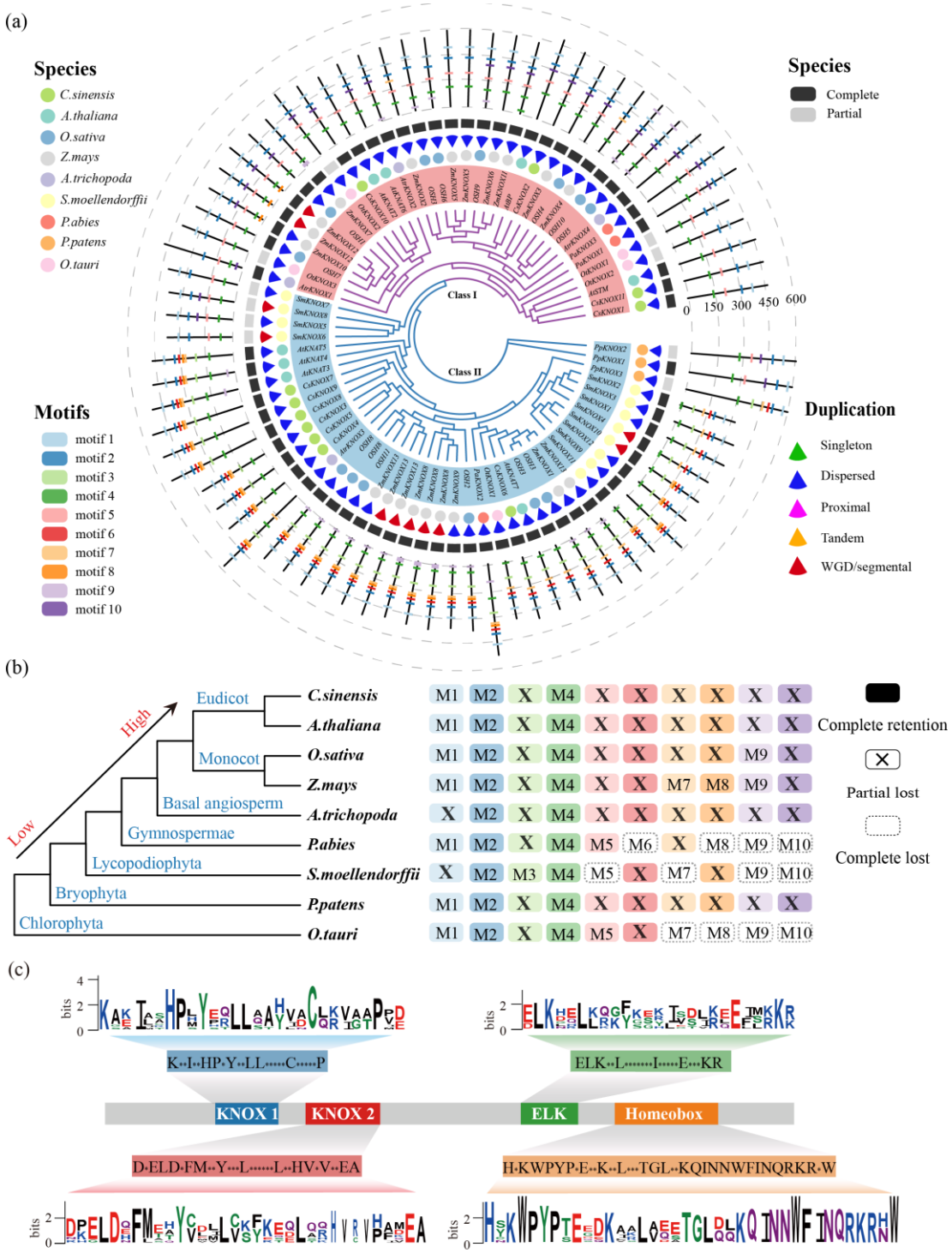

Figure 4. Conserved motif analysis of the *KNOX* gene. (**a**) Conserved motif analysis and duplication-type analysis of the *KNOX* genes in nine representative plants. (**b**) The motif evolutionary trajectories of the *KNOX* genes in nine representative plants. Blank and dashed boxes indicate complete loss or nonexistence of the motif. White with an X indicates loss of the motif in a certain gene. Black indicates the motif is present in all genes. (**c**). WebLogo analysis of the four domains (KNOX1, KNOX2, ELK, and Homeobox domains).

Moreover, the motif evolutionary trajectories of the *KNOX* gene were obtained to investigate the evolutionary trajectories of the domains of the *KNOX* genes (Figure 4b and Table S2). In the lower plant (*O. tauri*), only six motifs (M1, M2, M3, M4, M5, and M6) were detected, while four motifs (M7, M8, M9, and M10) were completely lost. In the bryophyte (*P. patens*), all 10 motifs were present, but M3, M5, M6, M7, M8, M9, and M10 were partially lost in some *KNOX* proteins. M5, M7, M9, and M10 were completely lost in Lycopodiophyta (*S. moellendorffii*) and M6, M8, M9 and M10 were completely lost in Gymnospermae (*P.abies*), while some motifs were partially lost in some genes. In two Monocots and two eudicots, M1, M2, and M4 were completely retained in all *KNOX* proteins. Further WebLogo analysis of all of the motif sequence revealed amino acid variability (Figure 4c). This result indicated that the occurrence of sequence divergence in the *KNOX* gene throughout the evolutionary course of the species.

3.5. Collinearity Analysis and Replication Type Identification of KNOX Gene

Gene duplication patterns may reflect gene family expansion mechanisms [52]. MCScanX was employed to analyze the entire genome as well as the *KNOX* genes, identifying five types of gene duplication, including singleton, dispersed, proximal, tandem, and WGD/segmental duplications (Figure 5 and Table S3). In all selected species, there were no *KNOX* genes classified under the proximal and tandem duplication types. As for singletons, only *O. sativa* and *Z. mays* possessed one singleton, while the remaining species did not have any. The majority of the *KNOX* genes belonged to the dispersed duplication type.

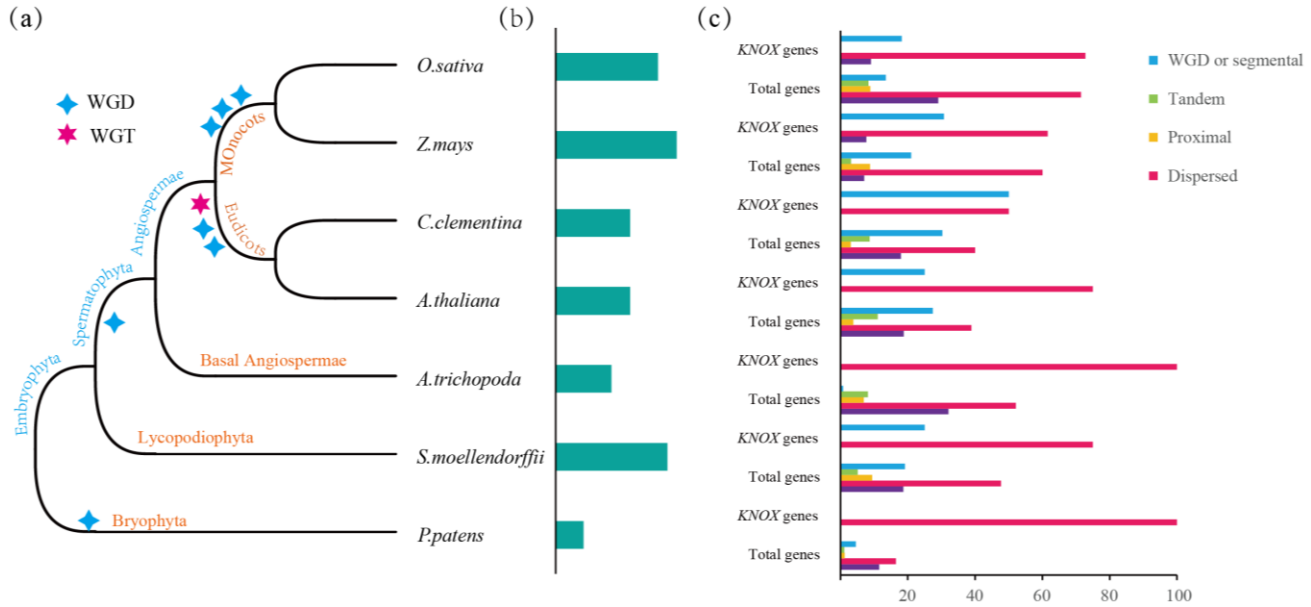

Figure 5. Number of *KNOX* genes and duplication types for each representative species. (**a**) The phylogenetic tree of the seven species. (**b**) The number of *KNOX* genes in the seven species. (**c**) The proportions of different replication types and genes belonging to different replication types in all *KNOX* genes in seven species. The number of blue quadrilaterals represents the occurrences of whole-genome duplication (WGD) events during the evolutionary process of the species, while the number of purple-red stars represents the occurrences of whole genome triplication (WGT) events during the evolutionary process of the species.

In *A. thaliana*, *C. clementina*, *O. sativa*, *S. moellendorffii*, and *Z. mays*, several KNOX genes were identified as belonging to the WGD/segmental duplication type (Figure 5c and Table S3). In *C. clementina*, the proportion of *KNOX* genes belonging to the WGD/segmental duplication type was 50.00% (Table S3). This proportion was significantly higher than the average proportion of all genome genes belonging to the WGD/segmental duplication

type, which was 30.28% ($p < 0.01$). Therefore, the WGD/segmental duplication type may play a crucial role in the expansion of the *KNOX* gene family in *C. clementina*.

3.6. Expression Pattern Analysis of the KNOX Gene in Model Plants

To investigate the expression pattern of the *KNOX* genes, a comprehensive analysis was conducted using a dataset that comprised 1032 samples obtained from diverse tissues of *A. thaliana*, as well as samples treated with various hormones (Tables S4 and S6). Through a visual expression analysis of the FPKM values, it was found that all class I *KNOX* genes (*AtSTM*, *AtKNAT1*, *AtKNAT2*, and *AtKNAT6*) are highly expressed in the stem while class II *KNOX* genes are not, but they have a small amount of expression in different tissues, This indicates that class I *KNOX* genes may be maintaining meristem activity, and play a key role in the process of organ differentiation (Figures 6a–d and S2). In addition, a further cluster analysis of its response to hormones also found that the expression of *AtSTM*, *AtBP1*, *AtKNAT2*, and *AtKNAT6* belonging to the class I *KNOX* gene in SAM under IAA treatment conditions was significantly higher than that in other tissues (Figure 6e and Table S6), further indicating that class I *KNOX* genes may cooperate with IAA to regulate the development of SAM.

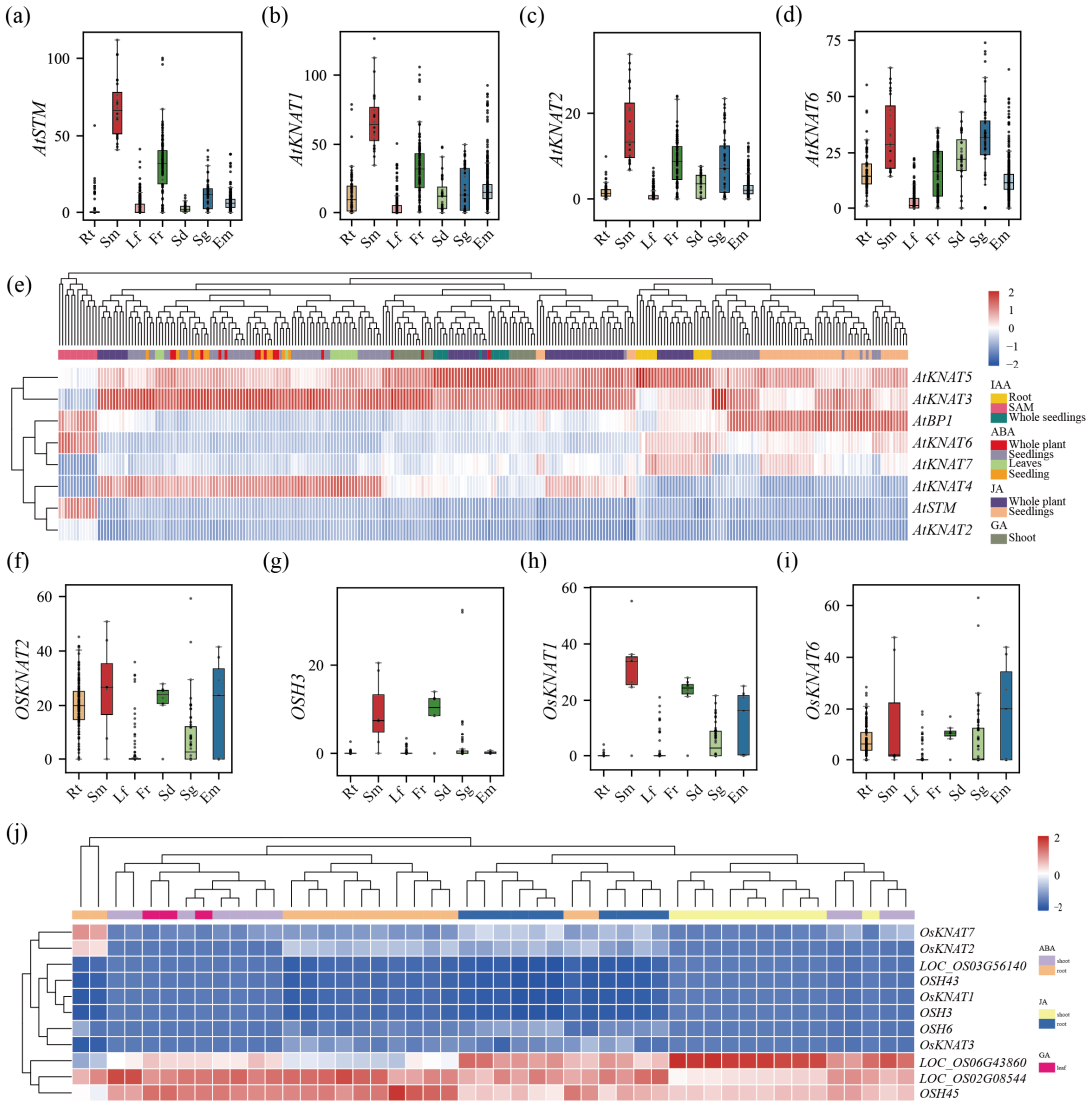

Figure 6. Analysis of expression patterns of *KNOX* genes in different tissues and under different hormone treatments in *A. thaliana* and *O. sativa*. (**a**–**d**) Box plots of expression patterns of class I

KNOX gene in seven different tissue species of *A. thaliana*. The *Y*-axis represents the FPKM of the genes. Rt: root; Sm: stem; Lf: leaf; Fr: flower; Sd: seed; Sg: seeding; Em: embryo. (**e**) Heatmap of expression pattern of *KNOX* gene in different tissues in *A. thaliana* under different hormone treatments. (**f**–**i**) Box plots of expression patterns of class I *KNOX* gene in seven different tissue species of *O. sativa*. The *Y*-axis represents the FPKM of the gene. Rt: root; Sm: stem; Lf: leaf; Sd: seed; Sg: seeding; Em: embryo. (**j**) Heat map of expression patterns of *KNOX* gene in different tissues in *O. sativa* under different hormone treatments.

To assess the differences between monocotyledonous and dicotyledonous plants, we also collected transcriptomic data for *O. sativa* as a representative monocot plant (Tables S5 and S7). The results revealed that the class I *KNOX* gene exhibited not only high expression levels in the stems but also prominent expression in the root, seed, and embryo tissues (Figure 6f–i). In contrast, among the class II *KNOX* genes, except for *OSH43* which showed significant up-regulation in seeds, the remaining genes demonstrated low levels of expression across different tissues (Figure S3). Therefore, these findings suggest the conservation of the developmental and functional roles of class I *KNOX* genes in these tissues across monocotyledonous and dicotyledonous plants. These results imply that these genes may play similar biological roles in different plant types, and their functions have been conserved during evolution. This conservation may be related to the importance and functionality of these tissues.

3.7. Pan-Genome Analysis of the KNOX Gene Family in 17 Citrinae

To understand the evolutionary relationships of the *KNOX* gene among different species in one genus, we performed *KNOX* family analysis in the Citrus genus and its closely related species. We selected a total of 17 species, including *Murraya paniculata*, *Atalantia buxfoliata*, *Citrus trifoliata*, *Citrus mangshanensis*, *Citrus ichangensis*, *Citrus linwuensis*, *Fortunella hindsii*, *Citrus media*, *Citrus australasica*, *Citrus hongheensis*, *Citrus grandis* (L.) Osbeck.cv. "Zipiyou", *Citrus grandis* (L.) Osbeck. cv. "Majiayou", *Citrus sinensis*, *Citrus reticulata*, *Citrus clementina*, and *Citrus reticulata* Blanco cv. "Ponkan". We performed gene family identification analysis and identified a total of 164 *KNOX* genes (Figure 7a and Table S8).

Among them, *M. paniculata*, a close relative of the citrus genus, has nine *KNOX* family genes, which is the same as *A. thaliana*, *F. hindsii*, *A. buxfoliata*, *C. mangshanensis*, *C. media*, and *C. ichangensis*, which are primitive or wild species of the Citrus genus with little difference from modern cultivated species. All of them have 10 *KNOX* genes, more than *M. paniculata*, indicating gene duplication events in the *KNOX* gene family. Only *C. linwuensis* has 12 *KNOX* genes, which may be related to its larger genome size. *C. trifoliata* and *C. grandis* have eight and seven *KNOX* genes, respectively, suggesting slower evolution in these species (Figure 7a and Table S8).

In order to further explore the evolutionary relationship of *KNOX* genes in the Rutaceae species, we performed multiple sequence alignment between the 164 identified *KNOX* proteins and 9 *A. thaliana* KNOX proteins and used FastTree to construct a phylogenetic tree (Figure 7b). All KNOX proteins were also divided into two categories, class I and class II. Among them, the class II genes of most citrus species are relatively conservative, with only one or missing orthologous genes, and class I genes basically have gene duplication, indicating that class I *KNOX* genes have evolved more actively in citrus species.

To explore the origin of duplicated *KNOX* genes, the MCScanX analysis package was used to examine the duplication types of each *KNOX* gene, including WGD/segmented duplications, tandem duplications, proximal duplications, and dispersed duplications (Table S9). Among them, citrus mainly includes two types, WGD/segmented duplication and dispersed duplication, accounting for 22.22–80% and 20–77% of the total *KNOX* genes in the species, respectively, indicating that these two repeat types played a key role in the

expansion of the *KNOX* gene family. Collinearity analysis also yielded the same results (Figures 5a and 7c).

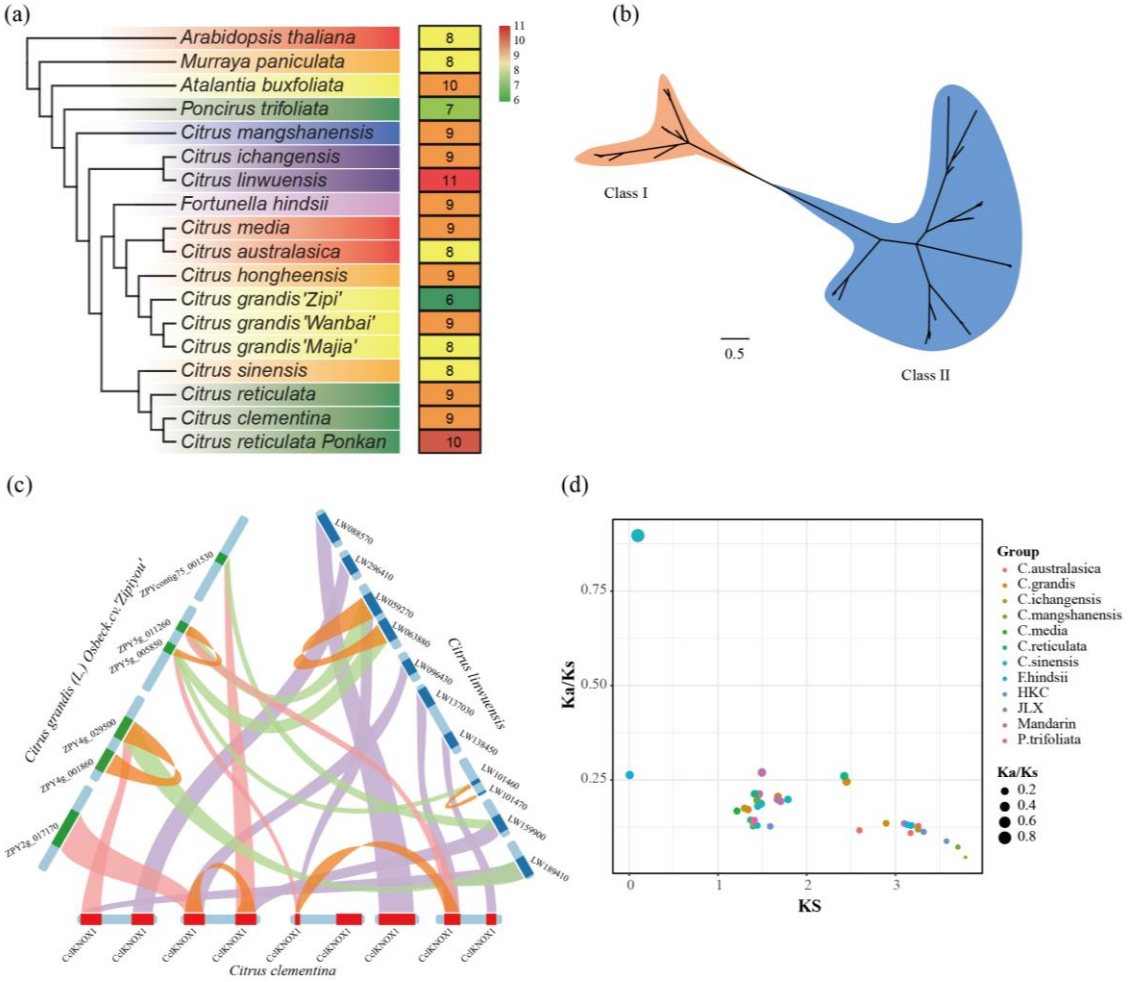

Figure 7. Pan-genome analysis of the *KNOX* genes in 17 Citrinae. (**a**) The phylogenetic tree and number of *KNOX* genes in the Citrus genus and its closely related species. (**b**) The phylogenetic tree of all of the KNOX proteins in the Citrus genus and its closely related species. (**c**) Collinear analysis of *KNOX* genes in *C. clementine*, *C. grandis* (L.) Osbeck cv. "Zipiyou", and *C. linwuensis*. (**d**) Ka/Ks analysis scatter plot of different citrus *KNOX* family duplicate gene pairs, the *X*-axis is the Ks value, and the *Y*-axis is the corresponding Ka/Ks ratio.

The Ka/Ks ratio reflects the selection pressure on duplicated genes. To explore the selective pressure of *KNOX* family genes in duplications, we calculated the Ka/Ks ratio of WGD/segmented duplication in the *KNOX* genes and plotted the Ks value and Ka/Ks ratio scatter plot of all repeated *KNOX* gene pairs (Figure 7d and Table S10). Among the 17 citrus species, the Ka/Ks ratios of all gene pairs ranged from 0.045 to 0.878, indicating that all repetitive genes in citrus were subject to purification selection, indicating that citrus *KNOX* genes are relatively conserved. In summary, the *KNOX* genes show a relatively conservative nature in Rutaceae plants and have undergone limited expansion.

4. Discussion

With the rapid development and advancements in sequencing technology, an increasing number of plant genomes have been completely sequenced. This provides valuable resources for studying the functional aspects of plant gene families and has become a prominent area of research in understanding the evolution of plant gene families [35,53]. For instance, Cenci et al. (2022) utilized nine flowering plant genomes to unravel the important

roles and evolutionary history of the GELP (GDSL-Type Esterase/Lipase) gene family in plant development [54]. Additionally, Qiao et al. (2019) developed a workflow called DupGen_finder, which identified different types of gene duplication events in 141 sequenced plant genomes, shedding light on the impact of gene and genome duplications on biodiversity evolution [55]. Furthermore, the evolutionary history of the same gene family across all species within a particular plant family can also be effectively analyzed [56]. In this study, we utilized a total of 118 plant genomes, all of which underwent comprehensive whole genome sequencing. Notably, some of these genomes were recently updated or represented higher-quality assemblies, such as the T2T genomes of *A. thaliana* [57], *O. sativa* [26], and *Z. mays* [58]. Among the selected species, there were 2 Rhodophyta, 8 Chlorophyta, 1 basal plant, and 107 land plants. The land plants were further classified into 26 monocots, 74 eudicots, 2 basal angiosperms, and 5 non-angiosperms. Compared to the PlantTFDB [39], the selection of species in this study offered greater comprehensiveness and representativeness. Additionally, the genomes utilized in this study exhibited higher quality. These factors collectively provide essential references for obtaining a thorough understanding of the evolution and functional changes within the *KNOX* gene family.

Our analysis found that the number of *KNOX* genes in higher plants is significantly higher than that in lower plants. This may be due to the fact that many WGD or WGT events have occurred in plants during the evolutionary process [59]. Especially in higher plants, which adapt to different environments, leaf morphogenesis occurs with higher diversity and complexity [49,60]. This finding is consistent with the discovery by Yu et al. (2022) that the proportion of heat shock transcription factors (*Hsf*) family genes is higher in the monocotyledonous plant banana (*M. acuminata*). This may be attributed to the fact that banana is a tropical fruit and requires a higher proportion of heat-tolerant genes to adapt to high-temperature environments. These results indicate a synergistic relationship between changes in the number of gene families and environmental adaptations [61]. Interestingly, three species with a higher proportion of *KNOX* genes belong to the Malvaceae family, further demonstrating that WGD or WGT events can affect the number of *KNOX* genes present, which it is consistent with previous reports that WGD and WGT events affect the number of *Hs* family genes in Brassicaceae species [61–64]. In gene duplication and loss analysis, the number of *KNOX* gene losses in most lycophytes, gymnosperms and angiosperms is greater than the number of gene duplications, indicating that frequent *KNOX* gene loss events occurring in higher plants occurred after WGD or WGT events.

Conserved motif analysis can reveal conserved patterns in gene families [43]. We found that three motifs (M1, M2, and M4) were present in nearly all *KNOX* genes and showed that the *KNOX* genes are structurally conserved. When we performed replication-type identification on seven plants, we found that the dispersed and WGD/segmental duplication type is the main type in the *KNOX* gene family. This is consistent with previous reports. In eudicots, segmental duplication is caused by whole genome duplication events [65,66]. Similar situations have been reported in other gene families, such as the MADS-box family gene [67], NAC transcription factors [68] and TCP transcription factors [69].

A. thaliana and *O. sativa* are typical model plants for dicotyledonous and monocotyledonous plants, respectively, and are often used for comparative analysis to reveal the conservation or differences in gene function and evolution [70]. In this study, we found that all class I *KNOX* genes are highly expressed in the stem, while class II *KNOX* genes are not, but have a small amount of expression in different tissues in *A. thaliana*, Consistent with extensive previous reports, class I *KNOX* members are mainly expressed in meristems and play a role in transcriptional regulation [3]. Among them, *STM* is highly expressed in the SAM, which maintains embryonic development [71]. *KNAT1* is recorded throughout the stem and inflorescence, but not in the leaves [72]. *KNAT2* is highly expressed in the SAM base and inflorescence tissue, but is low in the leaves, affecting leaf and flower development [73]. *KNAT6* is expressed on the lateral root edge and SAM and plays a role in lateral root initiation, meristem activity and organ formation [74,75]. There are few studies on class II *KNOX* genes, as their expression patterns are inconsistent with class I *KNOX* genes,

and some of them are functionally redundant [76]. In *O. sativa*, similar results were also shown, indicating that the expression patterns of *KNOX* genes are conserved. When treated with hormones, it was found that the class I *KNOX* gene in SAM under IAA treatment conditions was significantly higher than that in other tissues, indicating that class I *KNOX* genes may cooperate with IAA to regulate the development of SAM. Related findings have also been reported in barley [77], tomato [78], and maize [79].

Citrus is the largest fruit tree family in the world, ranking first in terms of production and planting area among fruit crops [80]. There are abundant types of leaf morphology, indicating that there may be differences in the development of leaf morphogenesis [81]. Previous studies have shown that *CiKN1* and *CiKN6* are differentially transcribed in the initial stages of leaf formation in trifoliate orange and lemon, and both of them can be directly involved in leaf development regulation through the *miR164a-CUC2* pathway, indicating that *KNOX* family genes are related to leaf morphogenesis [13]. Recently, the pan-genomes of 23 citrus accessions were reported [40], and *KNOX* gene family analysis was performed on 17 of them. The study found that there are differences in the number of *KNOX* genes between different species, reflecting the expansion of the *KNOX* gene in citrus, mainly including WGD/segmented duplication and dispersed duplication.

5. Conclusions

To gain a deeper understanding of the expression patterns and evolutionary mechanisms of *KNOX* family genes, we conducted an extensive, systematic, and comprehensive comparative analysis of *KNOX* family genes across 118 plant species. In this study, a total of 1425 *KNOX* genes were identified from these 118 plant species. Phylogenetic analysis classified them into two classes (class I and class II), and it was observed that they underwent independent evolution following species differentiation. Furthermore, it was noted that in non-angiosperm plants, gene loss was more prevalent than gene duplication within the *KNOX* gene family. Additionally, a conserved motif analysis revealed that three motifs (M1, M2, and M4) were present in nearly all *KNOX* genes, while four novel motifs (M7–M10) were lost in non-angiosperm plants, suggesting sequence divergence in *KNOX* genes during their evolution. Gene expression pattern analysis in Arabidopsis and rice showed that class I *KNOX* genes exhibited consistently high expression in stems, implying their potentially similar biological roles across different plant species and the conservation of their functions during evolution. Notably, this study marked the first pan-genome analysis of *KNOX* family genes in 17 citrus species. This analysis revealed that the number of *KNOX* family genes evolved at a slower rate in these citrus species, indicating a relatively conservative nature. These findings may offer new insights in the understanding of the evolutionary trajectory of *KNOX* genes and their contributions to leaf morphogenesis.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/horticulturae9111174/s1, Supporting information Figures S1–S3: Figure S1. (a) The statistics of *KNOX* gene number and percentage for three categories in plants (0, 1–8, ≥8). (b) The statistics of *KNOX* gene length ratio (average length of *KNOX* genes/average length of all genes), and their percentage for four categories (length ratio: 0, ≤0.8, 0.8–1.0, ≥1.2) in plants. Figure S2. (a–d) Box plots of expression patterns of *AtKNAT3*, *AtKNAT4*, *AtKNAT5*, and *AtKNAT7* in seven different tissue species of *A. thaliana*. The *Y*-axis represents the FPKM of the genes. Rt: root; Sm: stem; Lf: leaf; Fr: flower; Sd: seed; Sg: seeding; Em: embryo. Figure S3. (a–g) Box plots of expression patterns of other class II *KNOX* genes in seven different tissue species of *O. sativa*. The *Y*-axis represents the FPKM of the gene. Rt: root; Sm: stem; Lf: leaf; Sd: seed; Sg: seeding; Em: embryo. Supporting information Tables S1–S10: Table S1. Information of identified *KNOX* gene in 118 plants. Table S2. Conserved motifs and domains analysis of identified *KNOX* gene in seven representative plants. Table S3. Percentage of each duplication type for *KNOX* gene and all genes in seven representative species. Table S4. Fpkm value of *KNOX* gene from diverse tissues in *A. thaliana*. Table S5. Fpkm value of *KNOX* gene from diverse tissues in *O.sativa*. Table S6. Fpkm value of *KNOX* gene underwent hormone treatments in *A. thaliana*. Table S7. Fpkm value of *KNOX* gene underwent hormone treatments in *O.sativa*. Table S8. Identifying the *KNOX* gene in 17 Citrinae.

Table S9. Percentage of each duplication type for *KNOX* gene in 17 Citrinae. Table S10. Percentage of each duplication type for *KNOX* gene and all genes in seven representative species.

Author Contributions: Z.-M.G. collected and analyzed the data. C.-G.H. and J.-Z.Z. designed the study and the experiments. Z.-M.G. wrote the manuscript. X.-Y.A., C.-G.H. and J.-Z.Z. reviewed the manuscript. All authors discussed the data obtained. All authors have read and agreed to the published version of the manuscript.

Funding: This research was financially supported by the National Major Research and Development Plan (No. 2022YFD1200503) and the National Natural Science Foundation of China (Nos. 319702356, 32072521, and 31872045).

Conflicts of Interest: The authors have no conflict of interest to declare.

References

1. Efroni, I.; Eshed, Y.; Lifschitz, E. Morphogenesis of simple and compound leaves: A critical review. *Plant Cell* **2010**, *22*, 1019–1032. [CrossRef]
2. Tsuda, K.; Hake, S. Diverse functions of *KNOX* transcription factors in the diploid body plan of plants. *Curr. Opin. Plant Biol.* **2015**, *27*, 91–96. [CrossRef]
3. Frangedakis, E.; Saint, D.; Moody, L.A.; Rabbinowitsch, E.; Langdale, J.A. Nonreciprocal complementation of *KNOX* gene function in land plants. *New Phytol.* **2017**, *216*, 591–604. [CrossRef]
4. Qin, W.; Yin, Q.; Chen, J.; Zhao, X.; Yue, F.; He, J.; Yang, L.; Liu, L.; Zeng, Q.; Lu, F. The class II *KNOX* transcription factors *KNAT3* and *KNAT7* synergistically regulate monolignol biosynthesis in *Arabidopsis*. *J. Exp. Bot.* **2020**, *71*, 5469–5483. [CrossRef]
5. Gao, J.; Yang, X.; Zhao, W.; Lang, T.; Samuelsson, T. Evolution, diversification, and expression of KNOX proteins in plants. *Front. Plant Sci.* **2015**, *6*, 882. [CrossRef]
6. Serikawa, K.A.; Martinez-Laborda, A.; Zambryski, P. Three *knotted1-like homeobox* genes in *Arabidopsis*. *Plant Mol. Biol.* **1996**, *32*, 673–683. [CrossRef]
7. Vollbrecht, E.; Veit, B.; Sinha, N.; Hake, S. The developmental gene *Knotted-1* is a member of a maize homeobox gene family. *Nature* **1991**, *350*, 241–243. [CrossRef]
8. Kerstetter, R.; Vollbrecht, E.; Lowe, B.; Veit, B.; Yamaguchi, J.; Hake, S. Sequence analysis and expression patterns divide the maize *knotted1-like* homeobox genes into two classes. *Plant Cell* **1994**, *6*, 1877–1887.
9. Wechter, W.P.; Farnham, M.W.; Smith, J.P.; Keinath, A.P. Identification of resistance to peppery leaf spot among *Brassica juncea* and *Brassica rapa* plant introductions. *HortScience* **2007**, *42*, 1140–1143. [CrossRef]
10. Bhat, S.; Sarla, N. IIdentification and overcoming barriers between *Brassica rapa* (L. *) and B. nigra (L.) Koch crosses for the resynthesis of B. juncea (L.) Czern. Genet. Resour. Crop Evol.* **2004**, *51*, 455–469.
11. Cheng, F.; Sun, R.; Hou, X.; Zheng, H.; Zhang, F.; Zhang, Y.; Liu, B.; Liang, J.; Zhuang, M.; Liu, Y. Subgenome parallel selection is associated with morphotype diversification and convergent crop domestication in *Brassica rapa* and *Brassica oleracea*. *Nat. Genet.* **2016**, *48*, 1218–1224. [CrossRef] [PubMed]
12. Hay, A.; Tsiantis, M. The genetic basis for differences in leaf form between *Arabidopsis thaliana* and its wild relative *Cardamine hirsuta*. *Nat. Genet.* **2006**, *38*, 942–947. [CrossRef]
13. Zeng, R.F.; Fu, L.M.; Deng, L.; Liu, M.F.; Gan, Z.M.; Zhou, H.; Hu, S.F.; Hu, C.G.; Zhang, J.Z. *CiKN1* and *CiKN6* are involved in leaf development in citrus by regulating *CimiR164*. *Plant J.* **2022**, *110*, 828–848. [CrossRef] [PubMed]
14. Yang, Z.; Song, M.; Cheng, F.; Zhang, M.; Davoudi, M.; Chen, J.; Lou, Q. A SNP Mutation in Homeodomain-DDT (HD-DDT) Transcription Factor Results in Multiple Trichomes (mt) in Cucumber (*Cucumis sativus* L.). *Genes* **2021**, *12*, 1478. [CrossRef] [PubMed]
15. Kabak, G.; Şehsuvars, S.; Turgut, S.; Gökdemir, Ş. Genome-wide analysis of *Fragaria vesca* three-amino-acid-loop-extension (TALE) genes. *Biotechnol. Stud.* **2021**, *30*, 79–85. [CrossRef]
16. Wang, L.; Yang, X.; Gao, Y.; Yang, S. Genome-Wide Identification and Characterization of TALE Superfamily Genes in Soybean (*Glycine max* L.). *Int. J. Mol. Sci.* **2021**, *22*, 4117. [CrossRef]
17. Giannino, D.; Frugis, G.; Mele, G.; Nicolodi, C.; Chiappetta, A.; Bitonti, M.; Innocenti, A.; Mariotti, D. The *Knotted-like* homeobox genes (*Knox*) and cytokinins are involved in the differentiation and morphogenesis of lettuce leaves. In Proceedings of the Atti XVII Convegno Nazionale Della Societa'Italiana di Chimica Agraria, Portoferraio, Italy, 29 September–1 October 1999; pp. 181–189.
18. Jia, P.; Zhang, C.; Xing, L.; Li, Y.; Shah, K.; Zuo, X.; Zhang, D.; An, N.; Han, M.; Ren, X. Genome-Wide Identification of the *MdKNOX* Gene Family and Characterization of Its Transcriptional Regulation in *Malus domestica*. *Front. Plant Sci.* **2020**, *11*, 128. [CrossRef]
19. Di Giacomo, E.; Sestili, F.; Iannelli, M.A.; Testone, G.; Mariotti, D.; Frugis, G. Characterization of *KNOX* genes in *Medicago truncatula*. *Plant Mol. Biol.* **2008**, *67*, 135–150. [CrossRef]

20. Testone, G.; Condello, E.; Verde, I.; Nicolodi, C.; Caboni, E.; Dettori, M.T.; Vendramin, E.; Bruno, L.; Bitonti, M.B.; Mele, G. The peach (*Prunus persica* L. Batsch) genome harbours 10 *KNOX* genes, which are differentially expressed in stem development, and the class 1 *KNOPE1* regulates elongation and lignification during primary growth. *J. Exp. Bot.* **2012**, *63*, 5417–5435. [CrossRef]
21. Wang, J.; Zhao, P.; Cheng, B.; Zhang, Y.; Shen, Y.; Wang, X.; Zhang, Q.; Lou, Q.; Zhang, S.; Wang, B. Identification of TALE transcription factor family and expression patterns related to fruit chloroplast development in tomato (*Solanum lycopersicum* L.). *Int. J. Mol. Sci.* **2022**, *23*, 4507. [CrossRef]
22. Bằng, C.P. Identification, classification and expression analysis of *KNOX* gene family in potato (*Solanum tuberosum* L.) by using bioinformatic methods. *J. Sci. Technol.* **2017**, *3*, 84–87.
23. Song, N.; Liang, H.; An, Y.; Bai, S.; Ma, F.; Zhang, Z.; Li, H.; Zhou, Y.; Guo, G.; Song, C. Identification and Bioinformatics Analysis of *KNOX* Gene Family in Wheat (*Triticum aestivum* L.). *Mol. Plant Breed.* **2021**, *12*, 19. [CrossRef]
24. Yao, J.; Zhang, S.; Wu, N.; Li, X.; Ahmad, B.; Wu, J.; Guo, R.; Wang, X. KNOX transcription factor *VvHB63* affects grape seed development by interacting with protein *VvHB06*. *Plant Sci.* **2023**, *330*, 111665. [CrossRef] [PubMed]
25. Scofield, S.; Murray, J.A. *KNOX* gene function in plant stem cell niches. *Plant Mol. Biol.* **2006**, *60*, 929–946. [CrossRef]
26. Song, J.-M.; Xie, W.-Z.; Wang, S.; Guo, Y.-X.; Koo, D.-H.; Kudrna, D.; Gong, C.; Huang, Y.; Feng, J.-W.; Zhang, W. Two gap-free reference genomes and a global view of the centromere architecture in rice. *Mol. Plant* **2021**, *14*, 1757–1767. [CrossRef]
27. Yang, T.; Jiao, Y.; Wang, Y. Stem cell basis of shoot branching. *Plant Cell Physiol.* **2023**, *64*, 291–296. [CrossRef]
28. Su, Y.H.; Zhou, C.; Li, Y.J.; Yu, Y.; Tang, L.P.; Zhang, W.J.; Yao, W.J.; Huang, R.; Laux, T.; Zhang, X.S. Integration of pluripotency pathways regulates stem cell maintenance in the *Arabidopsis* shoot meristem. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 22561–22571. [CrossRef]
29. Zhao, M.; Yang, S.; Chen, C.-Y.; Li, C.; Shan, W.; Lu, W.; Cui, Y.; Liu, X.; Wu, K. *Arabidopsis BREVIPEDICELLUS* interacts with the *SWI2/SNF2* chromatin remodeling ATPase BRAHMA to regulate *KNAT2* and *KNAT6* expression in control of inflorescence architecture. *PLoS Genet.* **2015**, *11*, e1005125. [CrossRef]
30. Tabata, R.; Ikezaki, M.; Fujibe, T.; Aida, M.; Tian, C.-E.; Ueno, Y.; Yamamoto, K.T.; Machida, Y.; Nakamura, K.; Ishiguro, S. *Arabidopsis* auxin response factor6 and 8 regulate jasmonic acid biosynthesis and floral organ development via repression of class 1 *KNOX* genes. *Plant Cell Physiol.* **2010**, *51*, 164–175. [CrossRef]
31. Scofield, S.; Dewitte, W.; Murray, J.A. STM sustains stem cell function in the *Arabidopsis* shoot apical meristem and controls *KNOX* gene expression independently of the transcriptional repressor AS1. *Plant Signal. Behav.* **2014**, *9*, e28934. [CrossRef]
32. .Shani, E.; Burko, Y.; Ben-Yaakov, L.; Berger, Y.; Amsellem, Z.; Goldshmidt, A.; Ori, N. Stage-specific regulation of *Solanum lycopersicum* leaf maturation by class 1 KNOTTED1-LIKE HOMEOBOX proteins. *Plant Cell* **2009**, *21*, 3078–3092. [CrossRef] [PubMed]
33. Zhou, C.; Han, L.; Li, G.; Chai, M.; Fu, C.; Cheng, X.; Wen, J.; Tang, Y.; Wang, Z.-Y. *STM/BP-like KNOXI* is uncoupled from ARP in the regulation of compound leaf development in *Medicago truncatula*. *Plant Cell* **2014**, *26*, 1464–1479. [CrossRef] [PubMed]
34. Song, X.; Zhao, Y.; Wang, J.; Lu, M.-Z. 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.* **2021**, *72*, 5625–5637. [CrossRef]
35. Sufyan, M.; Daraz, U.; Hyder, S.; Zulfiqar, U.; Iqbal, R.; Eldin, S.M.; Rafiq, F.; Mahmood, N.; Shahzad, K.; Uzair, M. An overview of genome engineering in plants, including its scope, technologies, progress and grand challenges. *Funct. Integr. Genom.* **2023**, *23*, 119. [CrossRef]
36. Pautot, V.; Dockx, J.; Hamant, O.; Kronenberger, J.; Grandjean, O.; Jublot, D.; Traas, J. *KNAT2*: Evidence for a link between knotted-like genes and carpel development. *Plant Cell* **2001**, *13*, 1719–1734. [CrossRef]
37. Scofield, S.; Dewitte, W.; Murray, J.A. The *KNOX* gene *SHOOT MERISTEMLESS* is required for the development of reproductive meristematic tissues in *Arabidopsis*. *Plant J.* **2007**, *50*, 767–781. [CrossRef]
38. Roth, O.; Alvarez, J.P.; Levy, M.; Bowman, J.L.; Ori, N.; Shani, E. The *KNOXI* transcription factor *SHOOT MERISTEMLESS* regulates floral fate in Arabidopsis. *Plant Cell* **2018**, *30*, 1309–1321. [CrossRef]
39. Jin, J.; Tian, F.; Yang, D.-C.; Meng, Y.-Q.; Kong, L.; Luo, J.; Gao, G. PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* **2016**, *45*, gkw982. [CrossRef]
40. Liu, H.; Wang, X.; Liu, S.; Huang, Y.; Guo, Y.-X.; Xie, W.-Z.; Liu, H.; Qamar, M.T.U.; Xu, Q.; Chen, L.-L. Citrus Pan-Genome to Breeding Database (CPBD): A comprehensive genome database for citrus breeding. *Mol. Plant* **2022**, *15*, 1503–1505. [CrossRef]
41. Sato, Y.; Fukuda, Y.; Hirano, H.Y. Mutations that cause amino acid substitutions at the invariant positions in homeodomain of OSH3 KNOX protein suggest artificial selection during rice domestication. *Genes Genet. Syst.* **2001**, *76*, 381–392. [CrossRef]
42. Yu, G.; Lam, T.T.-Y.; Zhu, H.; Guan, Y. Two methods for mapping and visualizing associated data on phylogeny using ggtree. *Mol. Biol. Evol.* **2018**, *35*, 3041–3043. [CrossRef] [PubMed]
43. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef] [PubMed]
44. Nguyen, L.; Schmidt, H.A.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [CrossRef] [PubMed]
45. Chen, K.; Durand, D.; Farach, M. NOTUNG: A program for dating gene duplications and optimizing gene family trees. *J. Comput. Biol.* **2000**, *7*, 429–447. [CrossRef]
46. Wang, Y.; Tang, H.; DeBarry, J.; Tan, X.; Li, J.; Wang, X.; Lee, T.; Jin, H.; Marler, B.; Guo, H. MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **2012**, *40*, e49. [CrossRef]

47. Wang, D.; Zhang, Y.; Zhang, Z.; Zhu, J.; Yu, J. KaKs_Calculator 2.0: A toolkit incorporating gamma-series methods and sliding window strategies. *Genom. Proteom. Bioinform.* **2010**, *8*, 77–80. [CrossRef]
48. Nam, J.; dePamphilis, C.W.; Ma, H.; Nei, M. Antiquity and evolution of the MADS-box gene family controlling flower development in plants. *Mol. Biol. Evol.* **2003**, *20*, 1435–1447. [CrossRef]
49. De La Torre, A.R.; Piot, A.; Liu, B.; Wilhite, B.; Weiss, M.; Porth, I. Functional and morphological evolution in gymnosperms: A portrait of implicated gene families. *Evol. Appl.* **2020**, *13*, 210–227. [CrossRef]
50. Song, X.; Wang, J.; Ma, X.; Li, Y.; Lei, T.; Wang, L.; Ge, W.; Guo, D.; Wang, Z.; Li, C.; et al. Origination, expansion, evolutionary trajectory, and expression bias of AP2/ERF superfamily in Brassica napus. *Front. Plant Sci.* **2016**, *7*, 1186. [CrossRef]
51. Zhang, H.; Jin, J.; Jin, L.; Li, Z.; Xu, G.; Wang, R.; Zhang, J.; Zhai, N.; Chen, Q.; Liu, P.; et al. Identification and analysis of the chloride channel gene family members in tobacco (*Nicotiana tabacum*). *Gene* **2018**, *676*, 56–64. [CrossRef]
52. Sheridan, P.O.; Raguideau, S.; Quince, C.; Holden, J.; Zhang, L.; Williams, T.A.; Gubry, C. Gene duplication drives genome expansion in a major lineage of Thaumarchaeota. *Nat. Commun.* **2020**, *11*, 5494. [CrossRef] [PubMed]
53. Cheng, Q.-Q.; Ouyang, Y.; Tang, Z.-Y.; Lao, C.-C.; Zhang, Y.-Y.; Cheng, C.-S.; Zhou, H. Review on the development and applications of medicinal plant genomes. *Front. Plant Sci.* **2021**, *12*, 791219. [CrossRef] [PubMed]
54. Hou, X.; Wang, D.; Cheng, Z.; Wang, Y.; Jiao, Y. A near-complete assembly of an *Arabidopsis thaliana* genome. *Mol. Plant* **2022**, *15*, 1247–1250. [CrossRef] [PubMed]
55. Cenci, A.; Concepción-Hernández, M.; Guignon, V.; Angenon, G.; Rouard, M. Genome-wide classification and phylogenetic analyses of the GDSL-type esterase/lipase (GELP) family in flowering plants. *Int. J. Mol. Sci.* **2022**, *23*, 12114. [CrossRef]
56. Qiao, X.; Li, Q.; Yin, H.; Qi, K.; Li, L.; Wang, R.; Zhang, S.; Paterson, A.H. Gene duplication and evolution in recurring polyploidization–diploidization cycles in plants. *Genome Biol.* **2019**, *20*, 38. [CrossRef]
57. Wu, Y.; Li, D.; Hu, Y.; Li, H.; Ramstein, G.P.; Zhou, S.; Zhang, X.; Bao, Z.; Zhang, Y.; Song, B.; et al. Phylogenomic discovery of deleterious mutations facilitates hybrid potato breeding. *Cell* **2023**, *186*, 2313–2328. [CrossRef]
58. Chen, J.; Wang, Z.; Tan, K.; Huang, W.; Shi, J.; Li, T.; Hu, J.; Wang, K.; Wang, C.; Xin, B. A complete telomere-to-telomere assembly of the maize genome. *Nat. Genet.* **2023**, *55*, 1221–1231. [CrossRef]
59. Zhang, J.; Tian, Y.; Yan, L.; Zhang, G.; Wang, X.; Zeng, Y.; Zhang, J.; Ma, X.; Tan, Y.; Long, N. Genome of plant maca (*Lepidium meyenii*) illuminates genomic basis for high-altitude adaptation in the central Andes. *Mol. Plant* **2016**, *9*, 1066–1077. [CrossRef]
60. Del Pozo, J.C.; Ramirez-Parra, E. Whole genome duplications in plants: An overview from *Arabidopsis*. *J. Exp. Bot.* **2015**, *66*, 6991–7003. [CrossRef]
61. Yu, T.; Bai, Y.; Liu, Z.; Wang, Z.; Yang, Q.; Wu, T.; Feng, S.; Zhang, Y.; Shen, S.; Li, Q. Large-scale analyses of heat shock transcription factors and database construction based on whole-genome genes in horticultural and representative plants. *Hortic. Res.* **2022**, *9*, uhac035. [CrossRef]
62. Wang, X.; Wang, H.; Wang, J.; Sun, R.; Wu, J.; Liu, S.; Bai, Y.; Mun, J.-H.; Bancroft, I. The genome of the mesopolyploid crop species Brassica rapa. *Nat. Genet.* **2011**, *43*, 1035–1039. [CrossRef] [PubMed]
63. Yang, J.; Liu, D.; Wang, X.; Ji, C.; Cheng, F.; Liu, B.; Hu, Z.; Chen, S.; Pental, D.; Ju, Y. The genome sequence of allopolyploid Brassica juncea and analysis of differential homoeolog gene expression influencing selection. *Nat. Genet.* **2016**, *48*, 1225–1232. [CrossRef] [PubMed]
64. Zhang, L.; Xu, Y.; Zhang, X.; Ma, X.; Zhang, L.; Liao, Z.; Zhang, Q.; Wan, X.; Cheng, Y.; Zhang, J. The genome of kenaf (*Hibiscus cannabinus* L.) provides insights into bast fibre and leaf shape biogenesis. *Plant Biotechnol. J.* **2020**, *18*, 1796–1809. [CrossRef]
65. Schmutz, J.; Cannon, S.B.; Schlueter, J.; Ma, J.; Mitros, T.; Nelson, W.; Hyten, D.L.; Song, Q.; Thelen, J.J.; Cheng, J. Genome sequence of the palaeopolyploid soybean. *Nature* **2010**, *463*, 178–183. [CrossRef]
66. Tuskan, G.A.; Difazio, S.; Jansson, S.; Bohlmann, J.; Grigoriev, I.; Hellsten, U.; Putnam, N.; Ralph, S.; Rombauts, S.; Salamov, A.; et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **2006**, *313*, 1596–1604.
67. Theissen, G.; Becker, A.; Di Rosa, A.; Kanno, A.; Kim, J.T.; Münster, T.; Winter, K.-U.; Saedler, H. A short history of MADS-box genes in plants. *Plant Mol. Biol.* **2000**, *42*, 115–149. [CrossRef]
68. Hu, R.; Qi, G.; Kong, Y.; Kong, D.; Gao, Q.; Zhou, G. Comprehensive Analysis of NAC Domain Transcription Factor Gene Family in *Populus trichocarpa*. *BMC Plant Biol.* **2010**, *10*, 145. [CrossRef]
69. Wang, J.L.; Wang, H.W.; Cao, Y.N.; Kan, S.L.; Liu, Y.Y. Comprehensive evolutionary analysis of the TCP gene family: Further insights for its origin, expansion, and diversification. *Front. Plant Sci.* **2022**, *13*, 994567. [CrossRef]
70. Rouhier, N.; Couturier, J.; Jacquot, J.P. Genome-wide analysis of plant glutaredoxin systems. *J. Exp. Bot.* **2006**, *57*, 1685–1696. [CrossRef]
71. Long, J.A.; Barton, M.K. The development of apical embryonic pattern in *Arabidopsis*. *Development* **1998**, *125*, 3027–3035. [CrossRef]
72. Li, G.; Yang, J.; Chen, Y.; Zhao, X.; Chen, Y.; Kimura, S.; Hu, S.; Hou, H. *SHOOT MERISTEMLESS* participates in the heterophylly of Hygrophila difformis (Acanthaceae). *Plant Physiol.* **2022**, *190*, 1777–1791. [CrossRef]
73. Ragni, L.; Belles-Boix, E.; Gunl, M.; Pautot, V. Interaction of *KNAT6* and *KNAT2* with *BREVIPEDICELLUS* and *PENNYWISE* in *Arabidopsis* inflorescences. *Plant Cell* **2008**, *20*, 888–900. [CrossRef]
74. Butenko, M.A.; Shi, C.-L.; Aalen, R.B. *KNAT1*, *KNAT2* and *KNAT6* act downstream in the *IDA-HAE/HSL2* signaling pathway to regulate floral organ abscission. *Plant Signal. Behav.* **2012**, *7*, 135–138. [CrossRef] [PubMed]
75. Belles-Boix, E.; Hamant, O.; Witiak, S.M.; Morin, H.; Traas, J.; Pautot, V. *KNAT6*: An Arabidopsis homeobox gene involved in meristem activity and organ separation. *Plant Cell* **2006**, *18*, 1900–1907. [CrossRef]

76. Reiser, L.; Sánchez-Baracaldo, P.; Hake, S. Knots in the family tree: Evolutionary relationships and functions of knox homeobox genes. *Plant Mol. Evol.* **2000**, 151–166. [CrossRef]
77. Richardson, A.; Rebocho, A.B.; Coen, E. Ectopic *KNOX* expression affects plant development by altering tissue cell polarity and identity. *Plant Cell* **2016**, *28*, 2079–2096. [CrossRef]
78. Yan, F.; Deng, W.; Pang, X.; Gao, Y.; Chan, H.; Zhang, Q.; Hu, N.; Chen, J.; Li, Z. Overexpression of the *KNOX* gene *Tkn4* affects pollen development and confers sensitivity to gibberellin and auxin in tomato. *Plant Sci.* **2019**, *281*, 61–71. [CrossRef] [PubMed]
79. Bolduc, N.; Yilmaz, A.; Mejia, M.K.; Morohashi, K.; O'Connor, D.; Grotewold, E.; Hake, S. Unraveling the *KNOTTED1* regulatory network in maize meristems. *Genes Dev.* **2012**, *26*, 1685–1690. [CrossRef] [PubMed]
80. Wu, G.A.; Terol, J.; Ibanez, V.; López-García, A.; Pérez-Román, E.; Borredá, C.; Domingo, C.; Tadeo, F.R.; Carbonell-Caballero, J.; Alonso, R.; et al. Genomics of the origin and evolution of Citrus. *Nature* **2018**, *554*, 311–316. [CrossRef]
81. Da, X.; Yu, K.; Shen, S.; Zhang, Y.; Wu, J.; Yi, H. Identification of differentially expressed genes in a spontaneous altered leaf shape mutant of the navel orange [*Citrus sinensis* (L.) Osbeck]. *Plant Physiol. Biochem.* **2012**, *56*, 97–103. [CrossRef]