

Article

Not peer-reviewed version

Genome-Wide Identification and Characterization of the RWP-RK Proteins in *Zanthoxylum armatum*

[Xianzhe Zheng](#), Yanling Duan, [Huifang Zheng](#), Hao Tang, [Liumeng Zheng](#), [Xiaobo Yu](#)*

Posted Date: 11 April 2024

doi: 10.20944/preprints202404.0818.v1

Keywords: *Zanthoxylum armatum*; RWP-RK; apomixis; gene expression



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 and Characterization of the RWP-RK proteins in *Zanthoxylum armatum*

Xianzhe Zheng, Yanling Duan, Huifang Zheng, Hao Tang, Liumeng Zheng and Xiaobo Yu *

Southwest Research Center for Cross Breeding of Special Economic Plants, School of Life Science, Leshan Normal University, 614000 Sichuan, China

* Correspondence: yuxiaobosky@163.com

Abstract: *Zanthoxylum armatum* is an forest tree with apomixis characteristic, and RWP-RKs are plant-specific transcription factors known to regulate nitrogen response and embryonic development. However, the genome-wide analysis of RWP-RK family genes in *Z. armatum* is unclear. In this study, 36 *ZaRWP-RK* transcription factors were identified in the genome of *Z. armatum*, among which 15 genes were belonged to the RKD subfamily and 21 belonged to the NLP subfamily. Duplication events of *ZaRWP-RK* genes were mainly segmental duplication, and synteny analysis revealed close phylogenetic relationship between *Z. armatum* and *Arabidopsis*. The analysis of cis-elements indicated that *ZaRWP-RK* genes may be involved in the regulation of embryonic development of *Z. armatum* by responding to plant hormones such as abscisic acid, auxin, and gibberellin. Real-time PCR showed that the expression levels of most *ZaRWP-RK* genes were significantly increased from flowers to young fruits. Protein-protein interaction network analysis further revealed the potential roles of the *ZaRWP-RK* proteins in apomixis. Collectively, this study provide a theoretical basis for future investigations into the *ZaRWP-RK* genes and their regulatory mechanisms in the apomixis process of *Z. armatum*.

Keywords: *Zanthoxylum armatum*; RWP-RK; apomixis; gene expression

1. Introduction

Gene transcription is activated or repressed by transcription factors, which bind to specific cis-elements of the promoter [1]. The RWP-RK protein family, with a conserved RWP-RK domain, is a class of plant-specific transcription factors that is expressed in green algae, mosses, and vascular plants [2]. It is divided into two subfamilies: NLPs (NIN-like proteins) and RKDs (RWP-RK domain proteins) [3]. Unlike RKDs, all members of the NLP subfamily possess the PB1 (Phox and Bem 1) domain at their C-termini [4].

Two subfamilies of RWP-RK have been to participate in the modulation of various processes in plants. Members of the RKD subfamily are primarily involved in embryonic development [5]. For example, *AtRKD4* deficiency may lead to inhibition of zygote elongation and disrupt early cell division patterns [6]. Overexpression of *AtRKD1* and *AtRKD2* was found to induce the formation of egg-like structures [7]. Members of the NLP subfamily are also linked to the nitrogen response [8]. The *AtNLP7* receptor directly binds nitrate in intracellular environments [9]. Loss of *NLP2* might trigger a decreased in nitrogen fixation and nitrogen content in plants [10]. Moreover, evidence suggests that RWP-RK proteins modulate plant responses to abiotic stress, such as heat stress [11].

Zanthoxylum armatum DC., also known as 'Tengjiao' or 'Qinghuajiao', belongs to the rutaceae family and is widely distributed in southwest China [12,13]. The fruit of *Z. armatum* can be utilized to prepare food spice, medicine, and oil, highlighting its significant economic value [14]. In addition, *Z. armatum*, like citrus, exhibit apomixis traits [15]. *CitRWP*, a gene that encodes a protein with a RWP-PK domain, was associated with apomixis in citrus [16]. To date, analysis of the RWP-RK family has only been conducted for a few species such as soybean, tea, and *Brassica napus* [5,17,18]. However, and no family analysis of RWP-RK proteins in *Z. armatum* has been reported. Recently, the genome of the *Z. armatum* was published allowing researchers to analyze various gene families [15,19].

In this study, 36 *ZaRWP-RK* transcription factors were identified from the *Z. armatum* genome. Chromosomal localization, phylogenetic relationships, physicochemical properties, gene structure, and cis-element of the *ZaRWP-RK* transcription factors were explored. Furthermore, the expression levels of *ZaRWP-RK* in different tissues and protein-protein interaction predictions were conducted which showed that some *ZaRWP-RK* transcription factors may regulate apomixis process. Therefore, this study provides a theoretical basis for further investigations on RWP-RK genes and their regulatory mechanisms in the apomixis process in *Z. armatum*.

2. Materials and Methods

2.1. Identification and Physicochemical Properties of *ZaRWP-RK* Genes

The data of *Z. armatum* genome was obtained from figshare (https://figshare.com/articles/dataset/Genome_Data_of_Chinese_pepper/20217635) as described by Hu et al. [19]. A hidden markov model (HMM) file of the RWP-RK domain (PF02042) was downloaded from the Pfam database (<http://pfam-legacy.xfam.org/>), and the candidate *ZaRWP-RK* proteins were identified using HMMER 3.0 software. All candidate *ZaRWP-RK* proteins were examined by NCBI CDD (<https://www.ncbi.nlm.nih.gov/cdd/>) and Pfam (<http://pfam.xfam.org/>). The number of amino acids, molecular weight (Mw), and isoelectric point value (pI) of *ZaRWP-RK* proteins were assessed and screened using the ExPasy tools (<http://web.expasy.org/protparam/>). The subcellular localization of *ZaRWP-RK* proteins was predicted using WoLF PSORT (<https://wolfpsort.hgc.jp/>).

2.2. Chromosomal Location and Gene Duplication Events

The chromosomal location of *ZaRWP-RK* genes was identified on the *Z. armatum* genome and visualized with TBtools-II software [20]. The gene duplication events of *ZaRWP-RK* genes were analyzed using Multiple Collinearity Scan toolkit (MCScanX) [21]. The synteny analysis of the *ZaRWP-RK* genes between *Z. armatum* and other species (*Arabidopsis thaliana*, *Oryza sativa*, and *Citrus grandis*) was visualized using TBtools-II software [20].

2.3. Phylogenetic analysis of *ZaRWP-RK* genes

Alignment of RWP-RK protein sequences from *Z. armatum*, *Arabidopsis thaliana*, *Oryza sativa*, and *Citrus grandis* was achieved using the MAFFT multiple-alignment software [22]. Phylogenetic trees were constructed using the Maximum likelihood (ML) method in raxmlGUI 2.0 with 1000 bootstrap replicates [23]. They were also visualized with the online tool EvolView (<https://www.evolgenius.info/evolview-v2/>).

2.4. Conserved Structural Domains, Conserved Motifs, and Cis-element Analyses of *ZaRWP-RK* genes

Conserved structural domains of *ZaRWP-RK* proteins were identified using the online tool CD-Search (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) [24]. Subsequently, conserved motifs were determined using the online program MEME (<https://meme-suite.org/meme/tools/meme>). Promoter regions 2000 bp upstream of the start codon (ATG) were subjected to cis-elements analyses using PlantCARE (<https://bioinformatics.psb.ugent.be/webtools/plantcare/html/>). Next, the conserved structural domains, conserved motifs, and cis-element analyses of *ZaRWP-RK* genes were visualized using TBtools-II software [20].

2.5. Plant Materials, RNA Extraction and qRT-PCR Analysis

Different tissues (female flowers, young fruits, stems and leaves) of *Z. armatum* were collected from Juzi town in Leshan of Sichuan Province, China (29.32°N, 103.43°E). Total RNA samples were extracted using the FastPure® Plant Total RNA Isolation Kit (Polysaccharides and Polyphenolics-rich) (Vazyme, Nanjing, China) according to the manufacturer's instructions. 1 µg of total RNA was

then used to synthesize cDNA using ABScript III RT Master Mix for qPCR with gDNA Remover (ABclonal, RK20428). The cDNA was subjected to qRT-PCR on a QX96M quantitative PCR system (JLM, Chengdu, China). The reaction protocol was as follows: one cycle at 95°C for 30 s; 40 cycles at 95°C for 10 s, 60°C for 30 s, and melting curve generated automatically. All data presented was from at least three repeats. The results were analyzed using IBM SPSS Statistics 27 software. The primers for qPCR were designed using primer premier 5.0 (see Supplementary Table S2).

2.6. Protein-Protein Interaction Network Prediction of ZaRWP-RK Proteins

The ZaRWP-RK proteins from *Citrus sinensis* that were homologous to the ZaRWP-RK proteins in *Z. armatum* were submitted to the STRING 12 (<https://cn.string-db.org/>) to construct a protein-protein interaction network.

3. Results

3.1. Genome-Wide Identification and Physicochemical Properties of ZaRWP-RK Genes

Based on the RWP-RK domain (PF02042) in the Pfam database, 36 *ZaRWP-RK* genes were identified from the *Z. armatum* genome. The *ZaRWP-RK* genes were then named from *ZaRWP-RK1* to *ZaRWP-RK36*, and two *ZaRWP-RKs* (*ZaRWP-RK35* and *ZaRWP-RK36*) were not annotated to chromosomes based on the order of their chromosomal locations (Figure 1). Details of the physical and chemical properties of the *ZaRWP-RK* genes are presented in Table S1. The coding sequences (CDS) of *ZaRWP-RK* genes ranged from 438 bp (*ZaRWP-RK12*) to 3024 bp (*ZaRWP-RK29*). Similarly, *ZaRWP-RK* genes encode amino acids ranging from 145 aa (*ZaRWP-RK12*) to 1007 aa (*ZaRWP-RK29*), with molecular weights between 17.16 and 110.13 kDa. The isoelectric point (pI) distribution of *ZaRWP-RK* proteins varied from 4.56 (*ZaRWP-RK17*) to 9.57 (*ZaRWP-RK12*). Most of the *ZaRWP-RK* members were predicted to localize in the nucleus (32), and a few in the cytoskeleton (1), plasma membrane (1), cytosol (1), and chloroplast (1).

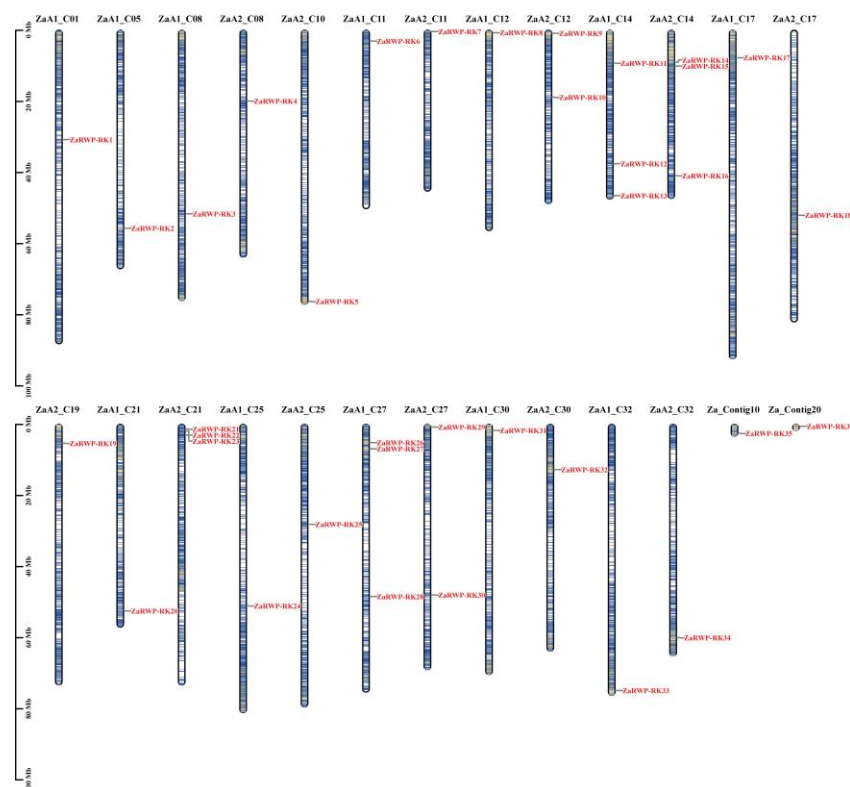


Figure 1. Chromosomal locations of the *ZaRWP-RK* genes. The chromosomal information of *ZaRWP-RK* genes was available in the database of *Z. armatum* genome. ZaA1 and ZaA2 represent the allelic chromosomes of autotetraploid. *ZaRWP-RK* genes are numbered in the order of chromosomes.

Next, we analyzed the duplication events of *ZaRWP-RK* genes (Figure 2). Ten pairs of segmental duplications were identified, while no tandem duplicated genes were found, indicating that segmental duplication likely had a substantial impact on the evolutionary trajectory of the *ZaRWP-RK* family. To gain further insights into the evolutionary relationships among the 36 *ZaRWP-RK* genes between *Z. armatum* and other plants, three collinearity maps were constructed using the *RWP-RK* members of *Z. armatum*, *Arabidopsis thaliana*, *Oryza sativa*, and *Citrus grandis* (Figure 3). The number of orthologous gene pairs between *Z. armatum* and *A. thaliana*, *Oryza sativa*, and *Citrus grandis* was 48, 11, and 35, respectively. Among the three species, *ZaRWP-RK* genes are most closely related to *AtRWP-RK* genes.

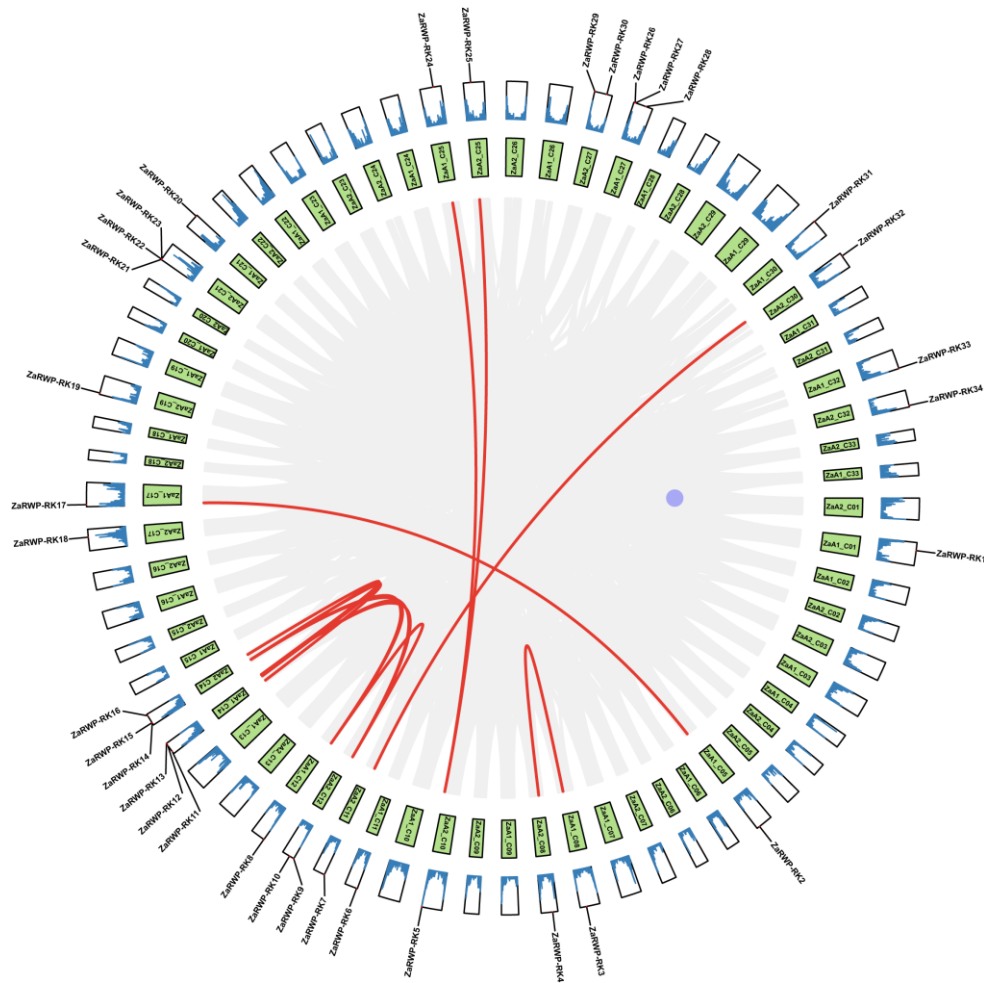


Figure 2. Collinearity analysis of the *ZaRWP-RK* genes. Red lines represent segmental duplication events in *ZaRWP-RK* family. Green boxes indicate the chromosomes of *Z. armatum*. Gene density is shown by the blue lines in the outermost box.

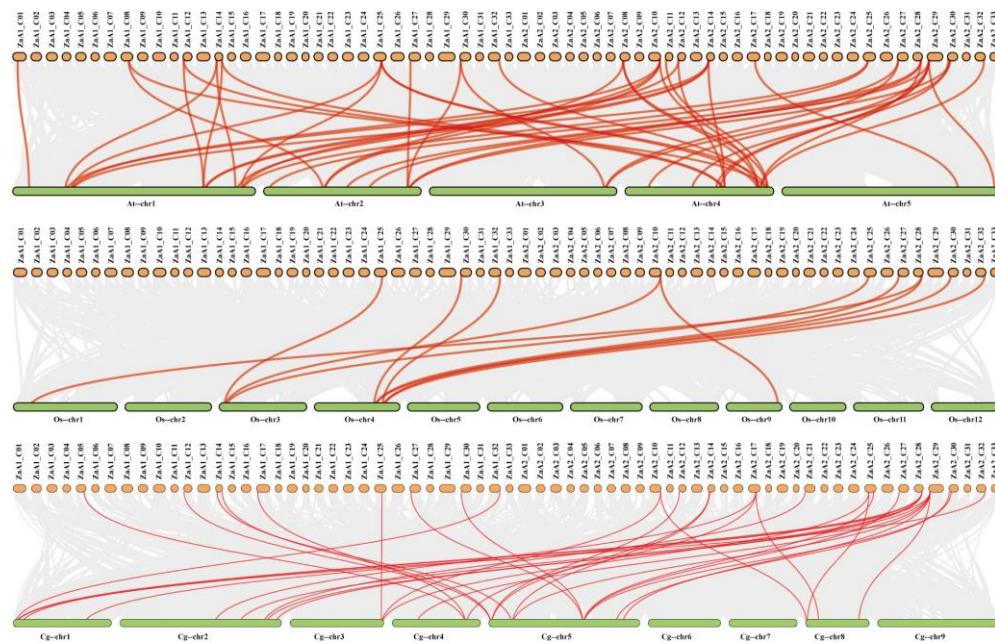


Figure 3. Synteny analysis of *ZaRWP-RK* genes between *Z. armatum*, *Arabidopsis thaliana*, *Oryza sativa*, and *Citrus grandis*. The gray lines in the backdrop indicate the collinear connections between *Z. armatum* and three other species, while the red lines signify the collinear pairs of *ZaRWP-RK* genes.

3.2. Phylogenetic Relationship, Conserved Structural Domains and Motifs Analysis of *ZaRWP-RK* proteins

To explore the phylogeny of the RWP-RK proteins, a phylogenetic tree was constructed based on the amino acid sequences from 36 *Z. armatum* *ZaRWP-RK* proteins, 14 *Arabidopsis* *AtRWP-RK* proteins, 13 rice *OsRWP-RK* proteins, and 8 pummelo *CgRWP-RK* proteins using raxmlGUI 2.0 (Figure 4). The *ZaRWP-RK* proteins were clustered into two subfamilies (NLP and RKD) using the results of *A. thaliana* as a reference [25]. Among them, NLP subfamily contained 21 members (*ZaNLP1* - *ZaNLP21*), and RKD subfamily contained 15 members (*ZaRKD1* - *ZaRKD15*) (Figure 4, Table S1).

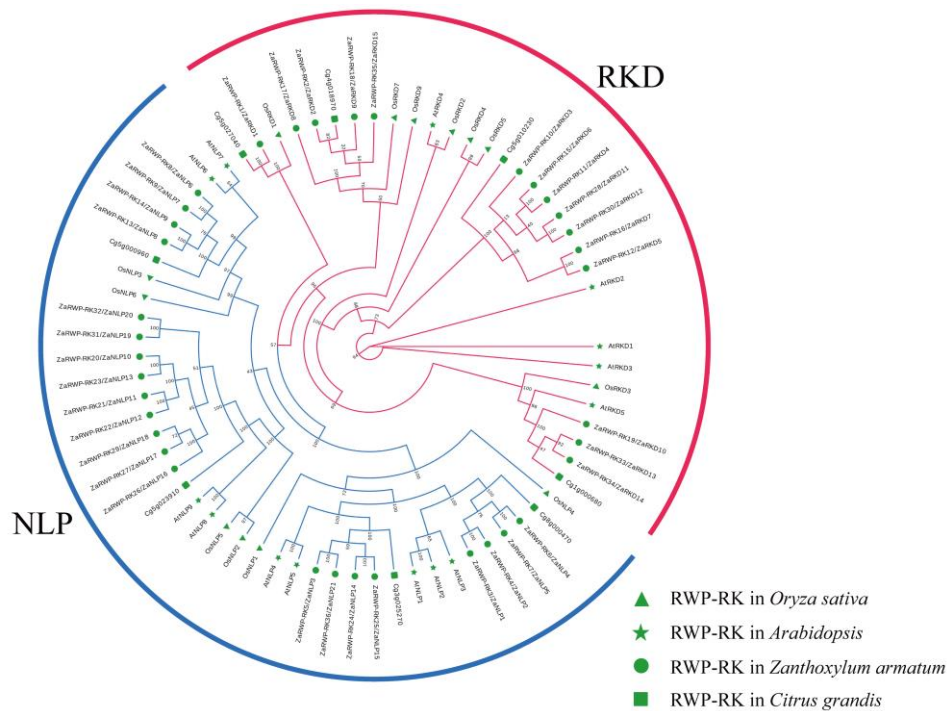


Figure 4. Phylogenetic relationships between *ZaRWP-RK* proteins and other plant RWP-RK proteins. The phylogenetic trees was constructed using the Maximum likelihood (ML) method in raxmlGUI 2.0

3.3. Cis-Element Analysis of *ZaRWP-RK* Promoters

To investigate the regulatory mechanisms of *ZaRWP-RK* genes, promoter regions 2000 bp upstream of the start codon (ATG) were subjected to cis-elements analyses (Figure 6, Table S2). Based on the results of cis-acting regulatory element analysis, 21 types of hormone-responsive elements were identified and classified into three categories, including abiotic stress, response to phytohormone, and plant growth and development. Among these elements, most of them were phytohormone-responsive elements. Research has demonstrated a close association of abscisic acid (ABA), auxin, and gibberellin with embryonic development. In terms of hormone response, ABRE (58) elements were the most prevalent and were identified as ABA response elements. Additionally, numerous auxin-responsive elements (14 TGA-element and 4 AuxRR-core) and gibberellin-responsive elements (12 P-box, 8 GARE-motif, and 7 TATC-box) were found in some *ZaRWP-RK* genes. Moreover, among the plant growth and development-related elements, unnamed_1 elements were the most abundant (42, 60k protein binding site) followed by CCAAT box (16, MYBHv1 binding site). Interestingly, some cis-elements were specific to certain *ZaRWP-RK* genes, such as MSA-like (*ZaRKD2*), DRE (*ZaNLP10* and *ZaRKD10*), HD-Zip 3 (*ZaNLP13*) and Box III (*ZaRKD1*, *ZaRKD11*, and *ZaRKD12*). These results suggested that the *ZaRWP-RK* genes might be regulated by hormones and some transcription factors, thereby modulating embryonic development.

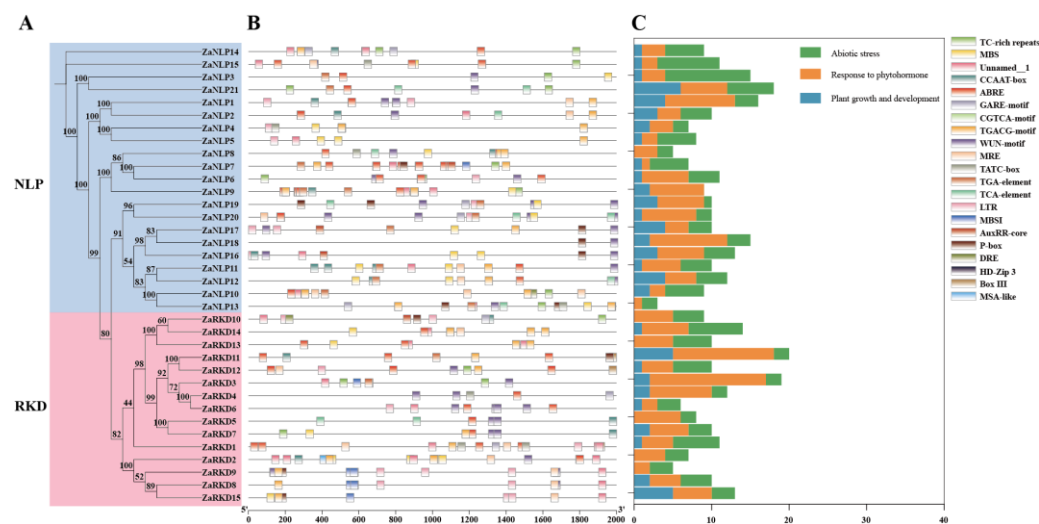


Figure 6. Analysis of cis-elements in the promoter of the *ZaRWP-RK* genes. (A) Phylogenetic tree of 36 *ZaRWP-RK* proteins. (B) Visualized location of cis-elements in the promoter of the *ZaRWP-RK* genes. Each of the cis-elements is distinguished by different colors. (C) The number of cis-elements for each *ZaRWP-RK* gene in three categories.

3.4. Tissue-Specific Expression of *ZaRWP-RK* Genes

To investigate the tissue-specific patterns of *ZaRWP-RK* genes, their expression profiles were explored in four different tissues of *Z. armatum* (female flower, young fruit, stem, and leaf) using qRT-PCR (Figure 7). It was observed that most *ZaRWP-RK* genes were highly expressed in young fruits, suggesting that these genes participate in the regulation of apomixis. Among them, *ZaNLP6*, *ZaNLP10*, *ZaNLP18*, and *ZaRKD3* were significantly upregulated in the young fruit, whereas *ZaRKD5* was highly expressed in female flower and young fruit, indicating their potential role in apomixis. In contrast, the *ZaNLP7* expression was significantly lower in the female flower and young fruit than in the stems and leaf, suggesting it might regulate apomixis in *Z. armatum* plants and perform negative regulatory function.

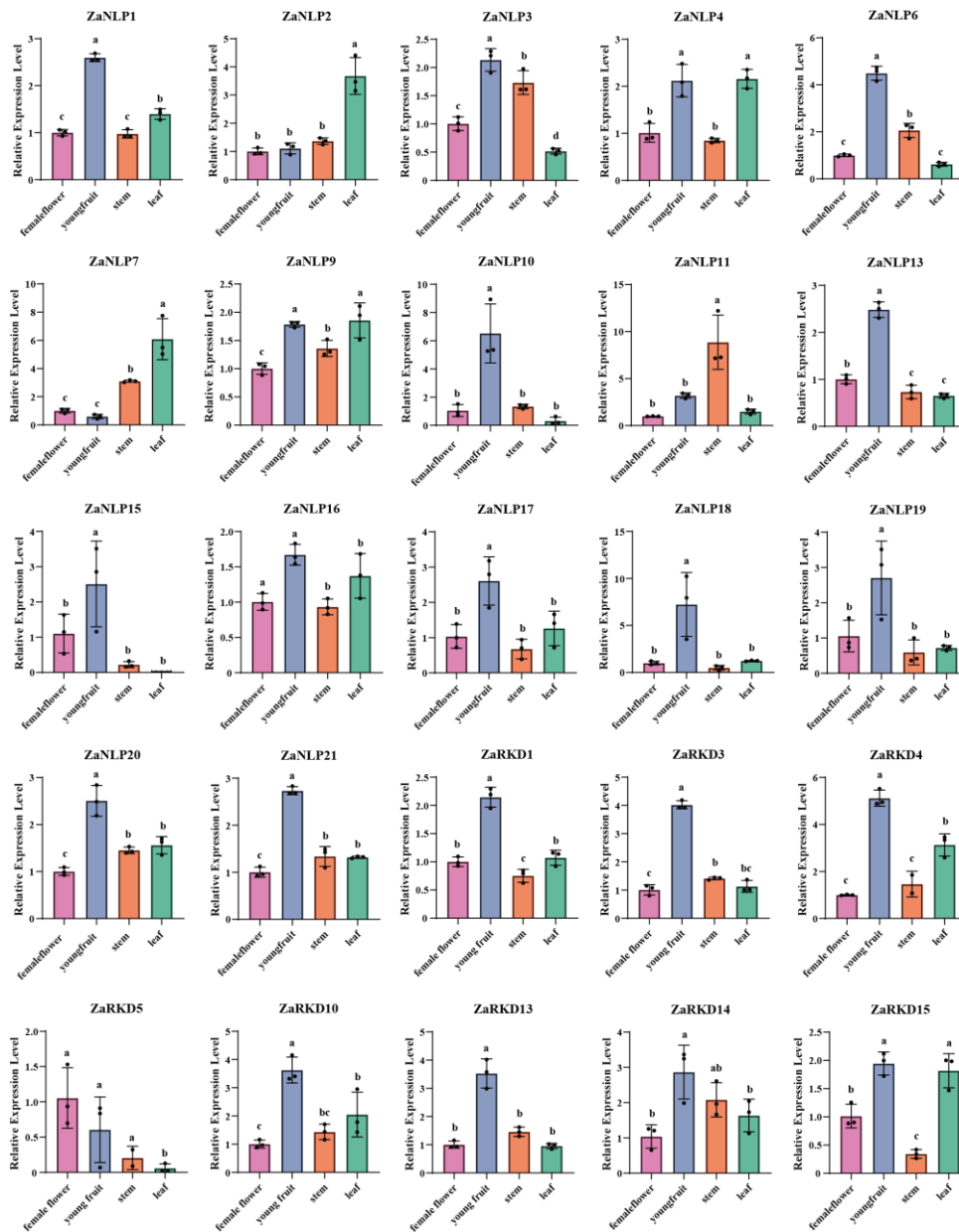


Figure 7. Expression patterns of ZaRWP-RK Genes in different tissues of *Z. armatum*. Data are presented as the mean \pm SD of three independent experiments. Different letters indicate significant differences between different tissues ($P < 0.05$), as determined by Student's t test.

3.5. Correlation Analysis of Apomixis-Related ZaRWP-RK Genes

Previous research has demonstrated that citrus apomixis may be caused by the expression of a gene known as *CitRWP* (Cg4g018970), which encodes a protein with an RWP-PK domain [16]. Similarly, *OsRKD3* can regulate somatic embryogenesis in black rice [26]. According to the results of phylogenetic tree analysis, four genes (*ZaRKD2*, *ZaRKD8*, *ZaRKD9*, and *ZaRKD15*) in *Z. armatum* were most closely related to *CitRWP*, and three genes (*ZaRKD10*, *ZaRKD13*, and *ZaRKD14*) were most closely related to *OsRKD3* (Figure 4). *Zanthoxylum* is a close relative of *citrus* in the Rutaceae family [27]. To further investigate the functions of these genes, the STRING database was employed to construct a protein-protein interaction network of these ZaRWP-RK proteins based on their CsRWP-RK homologs (Figure 8). Cs_ont_4g023450 (homolog of *ZaRKD2*, *ZaRKD10* and *ZaRKD14*) and Cs_ont_5g040080 (homolog of *ZaRKD8* and *ZaRKD9*) were directly connected to ATG5, ATG12, AP2A1, AP2S1, CHC, FIS1. Moreover, Cs_ont_5g040080 (homolog of *ZaRKD8* and *ZaRKD9*) directly interacted with Alba3.

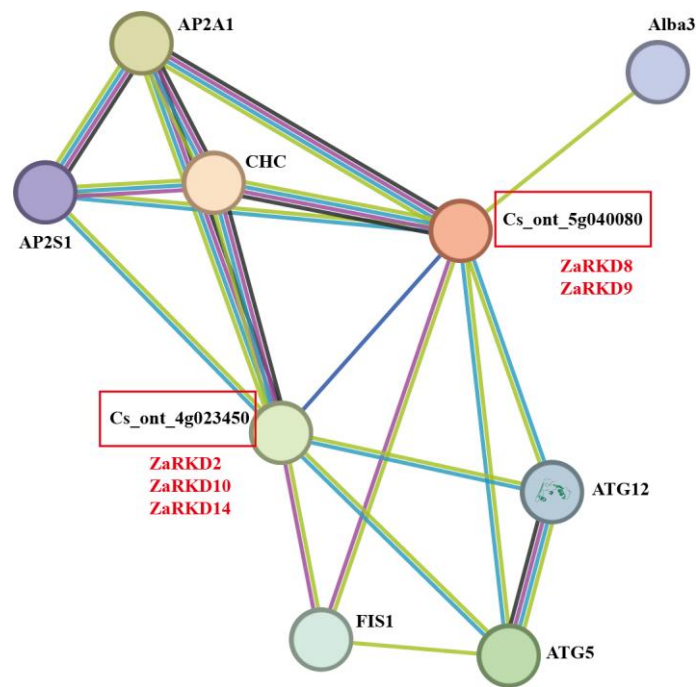


Figure 8. Protein-protein interaction network of RWP-RK homologs in *Citrus sinensis* visualized by STRING.

4. Discussion

The RWP-RK transcription factor is widely expressed in many plants and regulates various physiological processes such as nitrogen response, gametophyte development, and abiotic stress regulation [28]. In many plants, including *Arabidopsis*, rice, elephant grass and soybean, members of the RWP-RK transcription factor family have been reported [5,25,29,30]. However, no study has analyzed this family in *Z. armatum*. In this study, we performed the genome-wide identification of the RWP-RK family members in *Z. armatum* and analyzed their expression during reproductive processes, based on the recently published genome information of *Z. armatum* [19].

Using the conserved motif (RWPXRK) of RWP-RK transcription factor, 36 members of the RWP-RK family in Chili pepper genome were identified, which surpassed that reported in other plants, for example, 14 in *Arabidopsis*, 13 in rice, and 8 in pummelo (Figures 1 and 4). *Z. armatum*, having undergone a whole-genome duplication, has higher number of chromosomes ($2n = 4x = 132$) [15,19]. This indicates that the *ZaRWP-RK* genes have extensively expanded throughout the evolutionary process, probably due to the diversity of their functions. Gene duplication is a significant driving force for the diversification of gene functions and species evolution [31]. In eukaryotes, there are two main types of duplication events: tandem and segmental [32]. Previous studies have shown that the NLP subfamily in *Arabidopsis* evolved through segmental duplication rather than tandem amplification [25]. Similarly, we found that the *ZaRWP-RK* family in *Z. armatum* evolved primarily through segmental duplication (Figure 2). Moreover, we observed that the number of orthologous gene pairs between *Z. armatum* and dicots (*Arabidopsis* and pummelo) exceeded that between *Z. armatum* and monocots (rice), indicating that after the divergence of monocots and dicots, the *ZaRWP-RK* transcription factors have undergone extensive evolution and duplication (Figure 3).

The RWP-RK transcription factor family is generally divided into two subfamilies: the RKD subfamily and the NLP subfamily [25]. The available studies indicate that in *Arabidopsis*, the RKD subfamily has 5 members and the NLP subfamily has 7 members [29,33]. In rice, there are 7 members in the RKD subfamily and 6 members in the NLP subfamily [25]. In this study, we analyzed 36 *ZaRWP-RK* transcription factors from the *Z. armatum* genome. Through a phylogenetic analysis, we classified 15 of these members into the RKD subfamily and 21 members into the NLP subfamily (Figure 4, Table S1). Notably, the transcriptional activity, DNA-binding activity, and protein-protein interactions of transcription factors are often controlled by their motifs [34]. Among the 10 conserved

motifs identified in ZaRWP-RK family, motifs 1 and 7 were expressed in all ZaRWP-RK members, indicating that they might constitute the most conserved part of ZaRWP-RK (Figure 5B). Motifs 2-6 and 8-10 were only identified in members of the NLP subfamily, suggesting that these motifs might play an important role in unique biological functions of the NLP subfamily. Analysis of the conserved domains of ZaRWP-RK showed that the PB1 domain was present in all NLP subfamily members, which is the main distinguishing point between RKD and NLP subfamily (Figure 5C). Interestingly, our analyses showed that ZaNLP2 contained a conserved GAF_2 domain. A recent study reported that nitrate-triggered ROS signaling and the detection of nitrate deficiency are dependent upon the GAF domain of AtNLP7 [35]. Therefore, we speculated that ZaNLP2 might be involved in the regulation of nitrogen response. In addition, the PoIC superfamily domain was detected in ZaRKD10. PoIC is a major DNA polymerase involved in genome duplication, suggesting that ZaRKD10 might participate in rapid cell division stage, such as embryonic development [36].

The activation of gene expression is regulated by Cis-elements present in the promoter region [37]. Therefore, the physiological functions of cis-elements on the promoter can be used to predict the biological function of the genes [38]. In ZaRWP-RK family, 21 types of cis-elements were identified, categorized into abiotic stress (6), response to phytohormone (9), and plant growth and development (6) (Figure 6, Table S2). Most of them were hormone response elements, with the ABRE (ABA response element) accounting for the largest proportion. ABA are key hormones that control plant growth, embryogenesis, seed physiology reactions, and stress tolerance [39]. Pretreatment with ABA in embryo maturation medium increased cassava somatic embryo conversion into plants [40]. In addition, a significant number of auxin response elements and gibberellin response elements were identified. Auxin and gibberellin have been shown to regulate plant embryo development [41,42]. Furthermore, numerous elements related to plant growth and development were identified in the promoters of ZaRWP-RK members, suggesting that ZaRWP-RK transcription factors participate in embryogenesis. Consistent with this view, numerous studies have shown that ZaRWP-RK modulates plant reproduction processes [7]. Among the five RKD members in *Arabidopsis*, *AtRKD1* and *AtRKD4* are highly expressed during early embryo development, and the *AtRKD4* mutant leads to abnormal embryo development [6,29]. The MpRKD controls the development of gametophytes and contributes to the formation of gemma cups [43]. Therefore, our results indicate that ZaRWP-RK transcription factors may participate in plant reproductive processes by responding to and regulating hormones such as ABA, auxin, and gibberellin.

Apomixis, found in over 300 genera and 40 families throughout the plant kingdom, displays a wide taxonomic distribution [44]. In this mechanism, the fertilization step is skipped, allowing cells from nucellus or integument tissues to develop into somatic embryos. Previous studies have highlighted the crucial role of RWP-RK transcription factors in both somatic embryogenesis and apomixis processes [15,26,45]. Therefore, investigating the expression pattern of the *ZaRWP-RK* transcription factors during the apomictic reproduction process in *Z. armatum* is of great importance. In this study, we found that the expression levels of most *ZaRWP-RK* genes increase from flowers to young fruits significantly (Figure 7). Among them, the expression levels of *ZaNLP6*, *ZaNLP10*, *ZaNLP18*, and *ZaRKD3* and *ZaRKD5* show the most significant differences, indicating their potential functions in apomixis. In contrast, *ZaNLP7* expression was significantly lower in the female flower and young fruit than in the stems and leaf, suggesting it might negatively regulate apomixis.

Autophagy proteins are required for multiple functions during embryogenesis [46]. Our predictions based on the protein-protein interaction network suggested that Cs_ont_4g023450 (homolog of ZaRKD2, ZaRKD10 and ZaRKD14) and Cs_ont_5g040080 (homolog of ZaRKD8 and ZaRKD9) may interact with ATG5 and ATG12, suggesting that ZaRKDs may be involved in autophagy-regulated embryonic development (Figure 8). In addition, our results showed that Cs_ont_4g023450 and Cs_ont_5g040080 may interact with AP2A1, AP2S1, and CHC1 (Figure 8). The adaptor protein complex 2 (AP-2) proteins participate in various biological processes, such as plant growth and development, and stresses response [47]. AP2 σ and AP1/2 β modulate embryogenesis and plant growth in *Arabidopsis* [48,49]. AP-2 has been shown to interact with CHC and regulate flower organ development. Therefore, we speculate that ZaRKDs may interact with AP2A1, AP2S1,

and CHC1, thereby influencing the reproductive process in *Z. armatum*. Moreover, we also found that Cs_ont_4g023450 and Cs_ont_5g040080 interacted with MIS1 and Alba3. The mitochondrial fission complex has been implicated in *Arabidopsis* heat tolerance [50]. ALBA proteins modulate various stress responses [51]. Recent studies also have shown that overexpression of one RWP-RK gene in pearl millet can significantly enhance plant heat tolerance [11]. Therefore, it can be hypothesized that ZaRKDs may regulate heat tolerance in *Z. armatum*.

5. Conclusions

In this study, we identified 36 ZaRWP-RK transcription factors in the genome of *Z. armatum*. Most ZaRWP-RK genes were distributed on 24 chromosomes except ZaRWP-RK35 and ZaRWP-RK36. The ZaRWP-RK genes were into two subfamilies: NLP and RKD, through phylogenetic analysis. The conserved structural domain and motif analysis revealed the functional similarity and specificity of these ZaRWP-RK genes. Meanwhile, cis-element analysis showed that the ZaRWP-RK genes had multiple functions in plant reproductive processes. Analysis of gene expression in different tissues and protein-protein interaction network further highlighted the potential regulatory role of ZaRWP-RK genes in the apomixis of *Z. armatum*. In summary, through genome-wide identification and characterization of the RWP-RK transcription factor in *Z. armatum*, our results not only improve our understanding of the ZaRWP-RK family but also offer theoretical support for further studies into the regulatory mechanism of apomixis in *Z. armatum*.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org. Table S1: The characteristic and subcellular localization of ZaRWP-RK genes; Table S2: Primers for qRT-PCR.

Author Contributions: Conceptualization, X.Z. and X.Y.; software, L.Z.; validation, X.Z., Y.D. and H.Z.; formal analysis, H.T.; investigation, X.Y.; resources, X.Z.; data curation, X.Z.; writing—original draft preparation, X.Z.; writing—review and editing, X.Y.; visualization, X.Z. and D.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Natural Science Foundation of Sichuan Province (2022NSFSC0089), Key Research Projects of Science and Technology Bureau of Leshan Town (22ZDYJ0082), Scientific research project of Leshan Normal University (RC2023005), and Scientific research and cultivation project of Leshan Normal University (KYPY2024-0011).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The main data supporting the results of this article are included within the article (and its additional files).

Acknowledgments: I thank my wife Haohua Lu for useful comments, suggestions and patience.

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

References

1. Zhou, T., Zhu, X., Ye, Z., Wang, Y. F., Yao, C., Xu, N., Zhou, M., Ma, J., Qin, Y., Shen, Y., Tang, Y., Yin, Z., Xu, H., Zhang, Y., Zang, X., Ding, H., Yang, W., Guo, Y., Harley, J. B., Namjou, B., Kaufman, K.M., Kottyan, L.C., Weirauch, M.T., Hou, G., Shen, N. Lupus enhancer risk variant causes dysregulation of IRF8 through cooperative lncRNA and DNA methylation machinery. *Nat Commun* **2022**, *13*, 1855.
2. Sakuraba, Y., Zhuo, M., Yanagisawa, S. RWP-RK domain-containing transcription factors in the Viridiplantae: biology and phylogenetic relationships. *J Exp Bot* **2022**, *73*, 4323-4337.
3. Yin, M., Zhang, Z., Xuan, M., Feng, H., Ye, W., Zheng, X., Wang, Y. Conserved subgroups of the plant-specific RWP-RK transcription factor family are present in oomycete pathogens. *Front Microbiol* **2020**, *11*, 1724.
4. Mu, X., Luo, J. Evolutionary analyses of NIN-like proteins in plants and their roles in nitrate signaling. *Cell Mol Life Sci* **2019**, *76*, 3753-3764.
5. Amin, N., Ahmad, N., Khalifa, M.A.S., Du, Y., Mandozai, A., Khattak, A.N., Piwu, W. Identification and molecular characterization of RWP-RK transcription factors in soybean. *Genes (Basel)* **2023**, *14*, 369.

6. Waki, T., Hiki, T., Watanabe, R., Hashimoto, T., Nakajima, K. The Arabidopsis RWP-RK protein RKD4 triggers gene expression and pattern formation in early embryogenesis. *Curr Biol* **2011**, 21, 1277-1281.
7. Tedeschi, F., Rizzo, P., Rutten, T., Altschmied, L., Bäumllein, H. RWP-RK domain-containing transcription factors control cell differentiation during female gametophyte development in Arabidopsis. *New Phytol* **2017**, 213, 1909-1924.
8. Wu, Z., Liu, H., Huang, W., Yi, L., Qin, E., Yang, T., Wang, J., Qin, R. Genome-wide identification, characterization, and regulation of RWP-RK gene family in the nitrogen-fixing clade. *Plants (Basel)* **2020**, 9, 1178.
9. Liu, K.H., Liu, M., Lin, Z., Wang, Z.F., Chen, B., Liu, C., Guo, A., Konishi, M., Yanagisawa, S., Wagner, G., Sheen, J. NIN-like protein 7 transcription factor is a plant nitrate sensor. *Science* **2022**, 377, 1419-1425.
10. Jiang, S., Jardinaud, M. F., Gao, J., Pecrix, Y., Wen, J., Mysore, K., Xu, P., Sanchez-Canizares, C., Ruan, Y., Li, Q., Zhu, M., Li, F., Wang, E., Poole, P. S., Gamas, P., Murray, J. D. NIN-like protein transcription factors regulate leghemoglobin genes in legume nodules. *Science* **2021**, 374, 625-628.
11. Yan, H., Sun, M., Zhang, Z., Jin, Y., Zhang, A., Lin, C., Wu, B., He, M., Xu, B., Wang, J., Qin, P., Mendieta, J.P., Nie, G., Wang, J., Jones, C.S., Feng, G., Srivastava, R.K., Zhang, X., Bombarely, A., Luo, D., Jin, L., Peng, Y., Wang, X., Ji, Y., Tian, S., Huang, L. Pangenomic analysis identifies structural variation associated with heat tolerance in pearl millet. *Nat Genet* **2023**, 55, 507-518.
12. Liu, J., Wan, J., Zhang, Y., Hou, X., Shen, G., Li, S., Luo, Q., Li, Q., Zhou, M., Liu, X., Wen, C., Zhu, X., Zhang, Z. The establishment of comprehensive quality evaluation model for flavor characteristics of green Sichuan pepper (*Zanthoxylum armatum* DC.) in Southwest China. *Food Chem X* **2023**, 18, 100721.
13. Tang, N., Cao, Z., Wu, P., Zhang, X., Lou, J., Liu, Y., Wang, Q., Hu, Y., Si, S., Sun, X., Chen, Z. Genome-wide identification, interaction of the MADS-box proteins in *Zanthoxylum armatum* and functional characterization of ZaMADS80 in floral development. *Front Plant Sci* **2022**, 13, 1038828.
14. Huang, Y., Chen, J., Li, J., Li, Y., Zeng, X. Genome-wide identification and analysis of the growth-regulating factor family in *Zanthoxylum armatum* DC and functional analysis of ZaGRF6 in leaf size and longevity regulation. *Int J Mol Sci* **2022**, 23, 9043.
15. Wang, M., Tong, S., Ma, T., Xi, Z., Liu, J. Chromosome-level genome assembly of Sichuan pepper provides insights into apomixis, drought tolerance, and alkaloid biosynthesis. *Mol Ecol Resour* **2021**, 21, 2533-2545.
16. Wang, X., Xu, Y., Zhang, S., Cao, L., Huang, Y., Cheng, J., Wu, G., Tian, S., Chen, C., Liu, Y., Yu, H., Yang, X., Lan, H., Wang, N., Wang, L., Xu, J., Jiang, X., Xie, Z., Tan, M., Larkin, R. M., Chen, L.L., Ma, B.G., Ruan, Y., Deng, X., Xu, Q. Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. *Nat Genet* **2017**, 49, 765-772.
17. Li, D., Jin, Y., Lu, Q.H., Ren, N., Wang, Y.Q., Li, Q.S. Genome-wide identification and expression analysis of NIN-like protein (NLP) genes: Exploring their potential roles in nitrate response in tea plant (*Camellia sinensis*). *Plant Physiol Biochem* **2024**, 207, 108340.
18. Liu, M., Chang, W., Fan, Y., Sun, W., Qu, C., Zhang, K., Liu, L., Xu, X., Tang, Z., Li, J., Lu, K. Genome-wide identification and characterization of NODULE-INCEPTION-Like protein (NLP) family genes in *Brassica napus*. *Int J Mol Sci* **2018**, 19, 2270.
19. Hu, L., Xu, Z., Fan, R., Wang, G., Wang, F., Qin, X., Yan, L., Ji, X., Meng, M., Sim, S., Chen, W., Hao, C., Wang, Q., Zhu, H., Zhu, S., Xu, P., Zhao, H., Lindsey, K., Daniell, H., Wendel, J. F., Jin, S. The complex genome and adaptive evolution of polyploid Chinese pepper (*Zanthoxylum armatum* and *Zanthoxylum bungeanum*). *Plant Biotechnol J* **2023**, 21, 78-96.
20. Chen, C., Wu, Y., Li, J., Wang, X., Zeng, Z., Xu, J., Liu, Y., Feng, J., Chen, H., He, Y., Xia, R. TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. *Mol Plant* **2023**, 16, 1733-1742.
21. Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., Lee, T. H., Jin, H., Marler, B., Guo, H., Kissinger, J. C., Paterson, A. H. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res* **2012**, 40, e49.
22. Katoh, K., Rozewicki, J., Yamada, K.D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* **2019**, 20, 1160-1166.
23. Edler, D., Klein, J., Antonelli, A., Silvestro, D. RaxmlGUI 2.0: A graphical interface and toolkit for phylogenetic analyses using RAXML. *Methods Ecol Evol* **2021**, 12, 373-377.

24. Wang, J., Chitsaz, F., Derbyshire, M. K., Gonzales, N. R., Gwadz, M., Lu, S., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C. J., Marchler-Bauer, A. The conserved domain database in 2023. *Nucleic Acids Res* **2023**, 51, 384-388.
25. Schauser, L., Wieloch, W., Stougaard, J. Evolution of NIN-like proteins in *Arabidopsis*, rice, and *Lotus japonicus*. *J Mol Evol* **2005**, 60, 229-237.
26. Purwestri, Y. A., Lee, Y. S., Meehan, C., Mose, W., Susanto, F. A., Wijayanti, P., Fauzia, A. N., Nuringtyas, T. R., Hussain, N., Putra, H. L., Gutierrez-Marcos, J. RWP-RK Domain 3 (OsRKD3) induces somatic embryogenesis in black rice. *BMC Plant Biol* **2023**, 23, 202.
27. Feng, S., Liu, Z., Cheng, J., Li, Z., Tian, L., Liu, M., Yang, T., Liu, Y., Liu, Y., Dai, H., Yang, Z., Zhang, Q., Wang, G., Zhang, J., Jiang, H., Wei, A. Zanthoxylum-specific whole genome duplication and recent activity of transposable elements in the highly repetitive paleotetraploid *Z. bungeanum* genome. *Hortic Res* **2021**, 8, 205.
28. Chardin, C., Girin, T., Roudier, F., Meyer, C., Krapp, A. The plant RWP-RK transcription factors: key regulators of nitrogen responses and of gametophyte development. *J Exp Bot* **2014**, 65, 5577-5587.
29. Koszegi, D., Johnston, A. J., Rutten, T., Czihal, A., Altschmied, L., Kumlehn, J., Wüst, S. E., Kirioukhova, O., Gheyselinck, J., Grossniklaus, U., Bäumllein, H. Members of the RKD transcription factor family induce an egg cell-like gene expression program. *Plant J* **2011**, 67, 280-291.
30. Jin, Y., Luo, J., Yang, Y., Jia, J., Sun, M., Wang, X., Khan, I., Huang, D., Huang, L. The evolution and expansion of RWP-RK gene family improve the heat adaptability of elephant grass (*Pennisetum purpureum* Schum.). *BMC Genomics* **2023**, 24, 510.
31. Bowers, J.E., Chapman, B.A., Rong, J., Paterson, A.H. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **2003**, 422, 433-438.
32. Cannon, S.B., Mitra, A., Baumgarten, A., Young, N.D., May, G. The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol* **2004**, 4, 10.
33. Konishi, M., Yanagisawa, S. Arabidopsis NIN-like transcription factors have a central role in nitrate signalling. *Nat Commun* **2013**, 4, 1617.
34. Hu, H., Ma, L., Chen, X., Fei, X., He, B., Luo, Y., Liu, Y., Wei, A. Genome-wide identification of the NAC gene family in *Zanthoxylum bungeanum* and their transcriptional responses to drought stress. *Int J Mol Sci* **2022**, 23, 4769.
35. Wu, J., Song, Y., Zhang, Z. S., Wang, J. X., Zhang, X., Zang, J. Y., Bai, M. Y., Yu, L. H., Xiang, C. B. GAF domain is essential for nitrate-dependent AtNLP7 function. *BMC Plant Biol* **2022**, 22, 366.
36. Evans, R. J., Davies, D. R., Bullard, J. M., Christensen, J., Green, L. S., Guiles, J. W., Pata, J. D., Ribble, W. K., Janjic, N., Jarvis, T. C. Structure of PolC reveals unique DNA binding and fidelity determinants. *Proc Natl Acad Sci USA* **2008**, 105, 20695-20700.
37. Siepel, A., Arbiza, L. Cis-regulatory elements and human evolution. *Curr Opin Genet Dev* **2014**, 29, 81-89.
38. Wu, R., Duan, L., Pruneda-Paz, J. L., Oh, D. H., Pound, M., Kay, S., Dinneny, J. R. The 6xABRE synthetic promoter enables the spatiotemporal analysis of ABA-mediated transcriptional regulation. *Plant Physiol* **2018**, 177, 1650-1665.
39. Ali, F., Qanmber, G., Li, F., Wang, Z. Updated role of ABA in seed maturation, dormancy, and germination. *J Adv Res* **2021**, 35, 199-214.
40. Danso, K., Elgba, W. Optimisation of somatic embryogenesis in cassava. In *Biotechnologies for Plant Mutation Breeding*, eds.; Jankowicz-Cieslak, J., Tai, T. H., Kumlehn, J., Bradley, J. T. Eds.; Springer Nature: Switzerland, 2017; 73-90.
41. Li, M., Wrobel-Marek, J., Heidmann, I., Horstman, A., Chen, B., Reis, R., Angenent, G. C., Boutilier, K. Auxin biosynthesis maintains embryo identity and growth during BABY BOOM-induced somatic embryogenesis. *Plant Physiol* **2022**, 188, 1095-1110.
42. Park, J., Oh, D. H., Dassanayake, M., Nguyen, K. T., Ogas, J., Choi, G., Sun, T. P. Gibberellin signaling requires chromatin remodeler PICKLE to promote vegetative growth and phase transitions. *Plant Physiol* **2017**, 173, 1463-1474.
43. Rövekamp, M., Bowman, J.L., Grossniklaus, U. Marchantia MpRKD regulates the gametophyte-sporophyte transition by keeping egg cells quiescent in the absence of fertilization. *Curr Biol* **2016**, 26, 1782-1789.
44. Wang, Y., Underwood, C.J. Apomixis. *Curr Biol* **2023**, 33, R293-R295.

45. Shimada, T., Endo, T., Fujii, H., Nakano, M., Sugiyama, A., Daido, G., Ohta, S., Yoshioka, T., Omura, M. MITE insertion-dependent expression of CitRKD1 with a RWP-RK domain regulates somatic embryogenesis in citrus nucellar tissues. *BMC Plant Biol* **2018**, *18*, 166.
46. Zhou, X., Zhao, P., Sun, M.X. Autophagy in sexual plant reproduction: new insights. *J Exp Bot* **2021**, *72*, 7658-7667.
47. Yang, Z., Jin, H., Chen, J., Li, C., Wang, J., Luo, J., Wang, Z. Identification and analysis of the AP2 subfamily transcription factors in the pecan (*Carya illinoensis*). *Int J Mol Sci* **2021**, *22*, 13568.
48. Fan, L., Hao, H., Xue, Y., Zhang, L., Song, K., Ding, Z., Botella, M. A., Wang, H., Lin, J. Dynamic analysis of Arabidopsis AP2 σ subunit reveals a key role in clathrin-mediated endocytosis and plant development. *Development* **2013**, *140*, 3826-3837.
49. Liu, C., Li, Z., Tian, D., Xu, M., Pan, J., Wu, H., Wang, C., Otegui, M. S. AP1/2 β -mediated exocytosis of tapetum-specific transporters is required for pollen development in *Arabidopsis thaliana*. *Plant Cell* **2022**, *34*, 3961-3982.
50. Tsukimoto, R., Isono, K., Kajino, T., Iuchi, S., Shinozawa, A., Yotsui, I., Sakata, Y., Taji, T. Mitochondrial fission complex is required for long-term heat tolerance of *Arabidopsis*. *Plant Cell Physiol* **2022**, *63*, 296-304.
51. Náprstková, A., Malínská, K., Záveská Drábková, L., Billey, E., Náprstková, D., Sýkorová, E., Bousquet-Antonelli, C., Honys, D. Characterization of ALBA family expression and localization in *Arabidopsis thaliana* generative organs. *Int J Mol Sci* **2021**, *22*, 1652.

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.