BMC Plant Biology

RESEARCH Open Access

Genome-wide identification analysis of the 4-Coumarate: CoA ligase (*4CL*) gene family expression profiles in *Juglans regia* and its wild relatives *J. Mandshurica* resistance and salt stress

Jiayu Ma[1†], Dongjun Zuo[3†], Xuedong Zhang[1], Haochen Li[1], Hang Ye[1], Nijing Zhang[1], Mengdi Li[1], Meng Dang[1], Fangdong Geng[1], Huijuan Zhou[2*] and Peng Zhao[1*]

Abstract

Persian walnut (*Juglans regia*) and Manchurian walnut (*Juglans mandshurica*) belong to Juglandaceae, which are vulnerable, temperate deciduous perennial trees with high economical, ecological, and industrial values. 4-Coumarate: CoA ligase (4CL) plays an essential function in plant development, growth, and stress. Walnut production is challenged by diverse stresses, such as salinity, drought, and diseases. However, the characteristics and expression levels of *4CL* gene family in *Juglans* species resistance and under salt stress are unknown. Here, we identified 36 *Jr4CL* genes and 31 *Jm4CL* genes, respectively. Based on phylogenetic relationship analysis, all *4CL* genes were divided into three branches. WGD was the major duplication mode for *4CL*s in two *Juglans* species. The phylogenic and collinearity analyses showed that the *4CL*s were relatively conserved during evolution, but the gene structures varied widely. 4CLs promoter region contained multiply *cis*-acting elements related to phytohormones and stress responses. We found that *Jr4CLs* may be participated in the regulation of resistance to anthracnose. The expression level and some physiological of *4CLs* were changed significantly after salt treatment. According to qRT-PCR results, positive regulation was found to be the main mode of regulation of *4CL* genes after salt stress. Overall, *J. mandshurica* outperformed *J. regia*. Therefore, *J. mandshurica* can be used as a walnut rootstock to improve salt tolerance. Our results provide new understanding the potential functions of *4CL* genes in stress tolerance, offer the theoretical genetic basis of walnut varieties adapted to salt stress, and provide an important reference for breeding cultivated walnuts for stress tolerance.

Keywords *4CL* gene family, *Juglans*, Anthracnose resistance, Salt stress

†Jiayu Ma and Dongjun Zuo contributed equally to this work.

*Correspondence:
Huijuan Zhou
ericguanyuzhao@163.com
Peng Zhao
pengzhao@nwu.edu.cn

[1]Key Laboratory of Resource Biology and Biotechnology in Western China, Ministry of Education, College of Life Sciences, Northwest University, Xi'an, Shaanxi 710069, China
[2]Xi'an Botanical Garden of Shaanxi Province, Institute of Botany of Shaanxi Province, Shaanxi Academy of Science, Xi'an, Shaanxi, China
[3]College of Life and Environmental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

Introduction

Persian walnut (*Juglans regia*) and Manchurian walnut (*Juglans mandshurica*) belong to the *Juglans* genus in Juglandaceae [1]. Studies based on simplified genomes, chloroplast genomes, fossil evidence, and biogeographic historical reconstructions have found that these two walnut species diverged 20–31 million years ago [1–4]. *Juglans regia* (2n = 32), English walnut or Persian walnut, is an important oil-seeded perennial woody crop of monoeciousness and the oldest food source and widely cultivated nut-producing plant in the world [1, 5]. *J. regia* is originally from the mountains of Central Asia [6], it is an ancient cross between American and Asian lineages, dating back to the late 1900s [7]. It is distributed in Central, West and South Asia and Europe [1]. *J. regia* trees are light-loving and have strong resistance to drought, cold, and disease [1]. Its seeds are rich in a variety of nutrients needed by the human body, such as oil, vitamins, a variety of trace elements and minerals, etc., which have a certain therapeutic effect on slowing down the aging process and preventing heart disease and diabetes. It can also be made into biodiesel, edible walnut oil, and a variety of industrial chemicals, and is an important raw material for food and industrial applications [8–10]. *Juglans mandshurica* is a temperate deciduous tree whose wood and nuts are highly prized [11]. The fruit has a poor taste and is generally not eaten as a nut, but its seed kernel can be used as a traditional Chinese medicine to warm the kidney and moisten the intestines. In addition, tannin extracts are extracted from green husks, barks, and leaves, bark fibers are used as raw material for paper production. It grows mainly in the Asia region [11–14]. *J. mandshurica* as wild relative of *J. regia*, based on its potential as a reservoir of improved *J. regia* germplasm [11]. It is a hybrid breed with cultivated *J. regia* and can also be used as rootstock for *J. regia* because of its good resistance to stresses [11, 15, 16]. However, walnut production is challenged by environmental stresses, such as salinity, drought, and diseases [10, 17]. Recently, walnuts needed to improve the tolerance of different cultivars and varieties in its breeding program. Resistance studies of abiotic stresses in walnuts have focused on cold [18, 19] and drought tolerance [20, 21]. Mechanisms of salt resistance have been less well studied, whereas land salinization become more and more seriously during global climate changes. Therefore, to overcome the challenges of quality and quantity of walnuts, tolerance rootstocks is a fundamental strategy [22].

Salinity induces osmoregulation and increases antioxidant activities in plants [23, 24]. By applying appropriate stresses to plants, a range of physiological and morphological adaptations were activated, thereby increasing the plant's tolerance to other stresses. Enhancement of antioxidant enzyme activities [25], antioxidant molecules [26], and photo stabilizers [27] under salt treatments were common responses in tolerating other environmental stresses, and thus the imposition of controlled salt stresses is expected to make the plant tolerant to various adverse environmental preparedness. Based on previous studies, moderate salt stress may induce multiple stress tolerance in *J. regia* [23]. There are several studies on germination, physiological, and antioxidative responses of salinity stress in *J. regia* [27, 28]. Recently, the chromosome-level genomes of *J. regia* [5] and *J. mandshurica* [11] have been published, which provides great significance for our comparative study of genome-wide differences of important gene families related to stress responses.

The phenylpropanoid metabolism pathway is essential for plant survival and provides plants with a large number of precursors for secondary metabolites, which contribute to growth and development and external environment stress resistance [29]. 4-Coumarate: CoA ligase (*4CL*) was first extracted by Mansell et al. [30] from cambial regions in *Salix* species. Because of this enzyme's high catalytic specificity for 4-Coumaric acid, it is called 4-Coumarate: CoA ligase. The 4CL have a key role in linking lignin precursors to various other branching pathways and is one of the important enzymes in the phenylpropanoid metabolism pathway. *4CL* members affect plant growth, biosynthesis of phenylpropane derivatives, and environmental stress response, and can effectively regulate and improve plant-environment interactions. Current research suggested that several conserved polypeptide sequences were presented in 4CL protein sequences, including BOX I (SSGTTGLPKGV) and BOX II (GEICIRG) [31]. Among them, BOX I is highly conserved at the N-terminus and is the structural domain that binds to Adenosine monophosphate (AMP) and can directly participate in catalytic reactions [31]. BOX II is not directly involved in catalytic reactions, but the deficiency of this conserved sequence results in an almost complete loss of 4CL enzyme activity [32, 33]. Based on the functions of the proteins encoded by the *4CL* genes, *4CL* can be classified into three types, of which class I is mainly regulated in the biosynthesis of lignin compounds in plants, class II is mainly regulated the formation of flavonoid compounds, and class III is 4CL-like, whose specific function is still unclear [34].

The *4CL* genes have been extensively genome-wide identification studied in plant species, with 13 *At4CL* genes identified in *Arabidopsis thaliana*, 14 *Os4CL* genes in rice [35], 29 *Pbr4CL* genes in *Pyrus bretschneideri* [36], 35 *Eu4CL* genes in *Eucommia ulmoides* [37], 12 *Pg4CL* genes in *Punica granatum* [38] and 12 *Md4CL* genes in apple [39]. Since *4CL* genes are generally regulated in the process of plant response stresses, the studies of the regulatory mechanism and expression level of *4CL* genes

are important for plant molecular biology [40–42]. The *Gh4CL7*-silencing cotton showed more sensitivity to drought treatment, while overexpressed *Gh4CL7 Arabidopsis* lines were more tolerant to drought treatment [40]. Overexpressed *Fm4CL2* was more tolerant to drought in tobacco [41]. Similarly, overexpression of *Fm4CL-like1* improved drought resistance in tobacco [43]. White light, UV irradiation, exogenous ABA, and PEG treatments promoted the up-regulation of the expression of *St4CL6* and *St4CL8* and suppressed the expression of *St4CL5*, which was hypothesized to promote the accumulation of lignin and flavonoids and improve the ability to scavenge ROS through increased gene expression, thus improving the ability of potato to resist abiotic stresses [42]. The expression of the *Pn4CL*s increased with time after infestation with *Phytophthora capsica* and peaked after 24 h [44]. Uhlmann et al. infected potato leaves with *Phytophthora infestans* and then 4CL enzyme activity increased 2-fold 12 h after inoculation [45]. Moreover, after inoculation with *P. capsica*, the expression levels of *Pn4CLs* were all higher in resistance than in susceptible. *Pn4CLs* were involved in the process of pepper resistance to *P. capsica* [46]. Infection of *Arabidopsis* with spores of *Peronospora parasitica* strongly induced *At4CL1* and *At4CL2* mRNA expression [46]. Despite a large number of previous studies, the regulatory mechanisms of *4CL* in response to stresses remain unclear. Identifying the *4CL* gene family and investigating the gene expression of *4CL* genes under stresses in two *Juglans* species may supply clues for further research on the gene function of *4CL* in perennial trees.

In our study, we first systematically identified two *Juglans 4CL* gene families. Then, we analyzed the characterizations of *4CLs*, including locations, collinearity relationships, *cis*-acting elements prediction, and microRNA target predictions. We also performed the gene expressions of *4CL*s using transcriptomic data from multiple organs and biotic stress. Subsequently, the performance of two *Juglans* species and the expression of *4CLs* under salt stress by physiological activity assays and qRT-PCR experiments. All the results provide the basis for the function of *4CLs* in two *Juglans* species, as well as clues on the regulations of *4CLs* in other woody plants.

Materials and methods

Morphology of *Juglans regia* and *Juglans mandshurica*

In this study, we used the Persian walnut (*Juglans regia*) and Manchurian walnut (*Juglans mandshurica*) as plant materials (Fig. S1). The female flowers of *J. regia* are yellow (Fig. S1A) and the male flowers are catkins (Fig. S1B). The leaves are simple pinnately compound, compound leaves alternate, elliptic in shape with smooth margins (Fig. S1C). The fruit is oblate-globose (Fig. S1D). The female flowers of *J. mandshurica* are purplish red or bright red (Fig. S1E) and the male flowers are catkins (Fig. S1F). Leaves are simple pinnately compound, petiole is very short, the leaf shape is oblong, the leaf margin and serrulate, abaxial surface is densely pilose (Fig. 1G). The fruit is ovate or ovoid, with an acute tip, and the infructescence is usually 4-7-fruited (Fig. 1H).

Genome-wide identification of *4CL* genes in *J. Regia* and *J. Mandshurica*

To identify 4CL candidate members in *J. regia* [5] and *J. mandshurica* [11], we used the proteins as query sequences of 13 At4CL members were downloaded from TAIR database (https://www.arabidopsis.org/). We performed the genome-wide identification using BLASTP (E-value < $1E^{-5}$). To determine whether the 4CL candidate members involved the *4CL* gene family, the protein domains of these candidate members were queried in the NCBI CDD [47], Pfam [48], and SMART databases [49], and then the results were selected using excel. Candidate members containing the 4CL domain and AFD_class_I superfamily domain in CDD database, AMP-binding domain and AMP-binding_C domain in Pfam database, and 4CL domain and AFD_CAR_like domain in SMART database were the final members of the *4CL* gene family.

Chromosomal localization and collinearity analysis

Chromosomal localization was visualized for all identified two *Juglans* species *4CL* genes based on gene annotation information using TBTOOLS software [50]. To facilitate subsequent studies, *4CL* genes were renamed according to the order of their position on the chromosome. Analysis of the 4CL genes collinearity between two *Juglans* species and three other selected species (*Arabidopsis*, *O. sativa*, and *P. bretschneideri*) [35, 36], as well as prediction of gene duplication events, was particular using MCScanX software [51]. The non-synonymous substitution (Ka), synonymous substitution (Ks) values, and the ratio of Ka/Ks in homologous genes were calculated using KaKs_Calclator v2.0 software [52].

Physicochemical analyses and *Cis*-acting element prediction

Physicochemical properties and subcellular localization were analyzed on the ExPASy website [53], and WoLF PSORT website (https://wolfpsort.hgc.jp/), respectively. The *cis*-acting elements of the promoter region were predicted from all identified 2000 bp upstream sequences of the *4CL* genes using the PlantCARE online website [54]. Furthermore, the eggNOG-mapper online website (http://eggnog-mapper.embl.de/) was utilized to the identified *4CL* genes for GO (Gene Ontology) enrichment analysis [55].

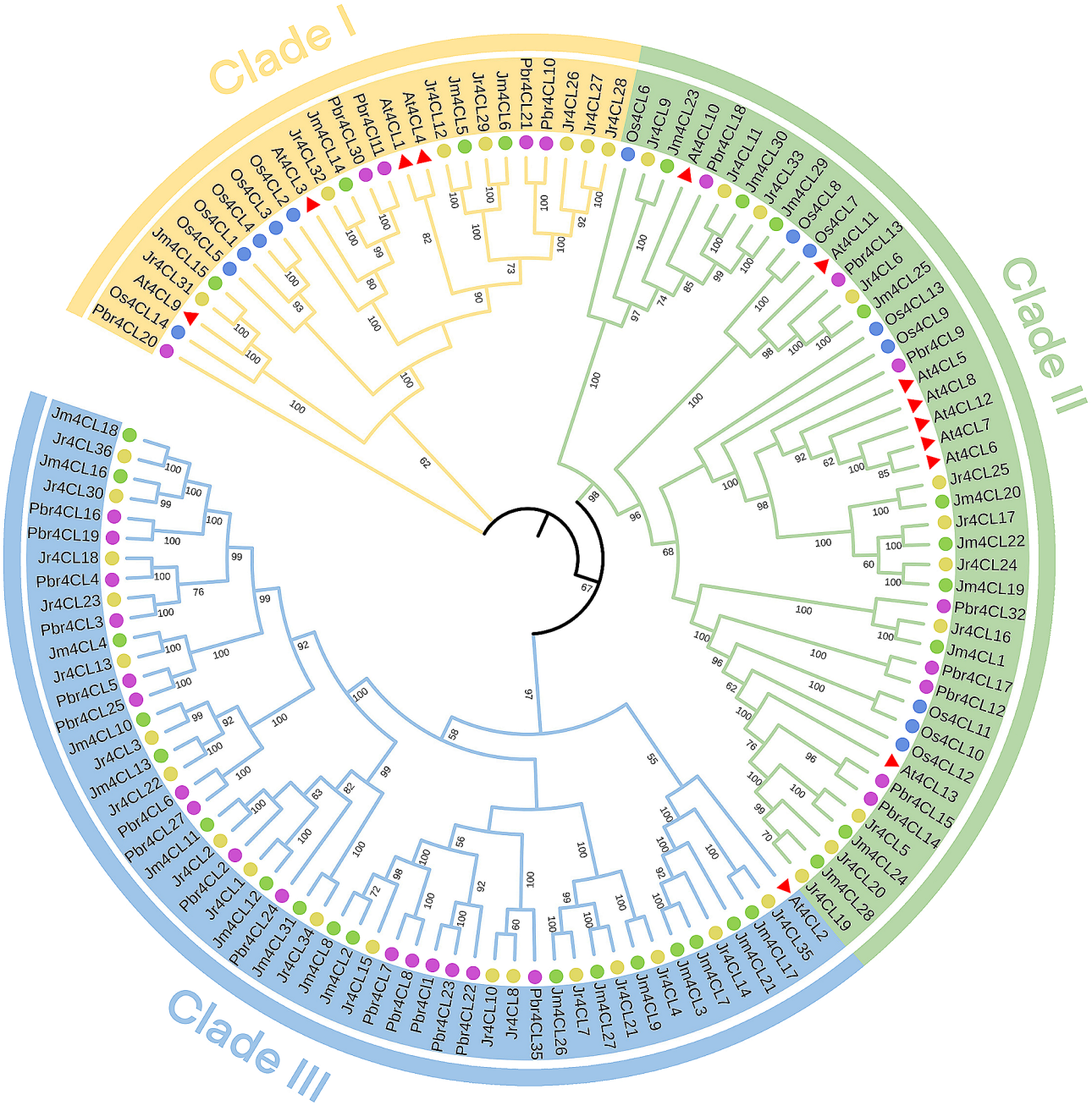

Fig. 1 The phylogenetic tree of 4CL protein of *Arabidopsis thaliana*, *Oryza sativa*, *Pyrus bretschneideri*, *Juglans regia*, and *J. mandshurica*. The red triangles, bule, purple, yellow, and green circles represent *Arabidopsis*, rice, *P. bretschneideri*, *J. regia*, and *J. mandshurica*, respectively

Phylogenetic and characteristic analysis

The 4CL protein sequences of five species (*J. regia, J. mandshurica, P. bretschneideri*, *Arabidopsis*, and *O. sativa*) [11, 35, 36] were constructed as a phylogenetic tree using MEGA v11 software (Maximum Likelihood method; bootstrap: 1000) [56]. We aligned the multiple sequences using ClustalX [57]. Beautification using the online website iTOL [58]. Based on the information provided by the CDD database in NCBI [47], using TBTOOLS [50] to show the distribution of characteristics domains on 4CL proteins of two walnut species. Analyzing the gene structure using the GSDS online website [59].

Protein-protein interaction analysis and microRNA target prediction

We followed the protein-protein interaction analysis in STRING database (http://string-db.org) using all identified 4CL members' protein sequences as query sequences and the *Arabidopsis* protein database as a reference. Visualization using Cytoscape software default parameters

[60]. MicroRNA targeting prediction for all identified *4CL* members using default parameters from the psRNA-Target online website [61].

Plant material collection, treatment, and physiological indicators measurement

We collected the mature leaves of two *Juglans* species' experimental materials. The mature leaves from two *Juglans* species were picked in the middle of July from Qinling National Forest Park in Shaanxi province [62]. Professor Peng Zhao identified *J. mandshurica* and *J. regia* based on the botanical characteristics of the leaves. We obtained permission to collect those plant samples from Qinling National Forest Park in Shannxi province. Salt stress treatments by soaking leaves in 150 mmol/L Na_2SO_4 solution. Removed the leaves after 24 h of salt treatment and soaked in clean water as a control. The obtained leaves were observed phenotypically. Then they were snap-frozen in liquid nitrogen and stored at -80 °C for backup. The voucher specimens of *J. regia* and *J. mandshurica* (deposition accession numbers: NWU2022024 and NWU2022025) were stored at the Evolutionary Botany Laboratory, College of Life Sciences, Northwest University (Xi'an, Shaanxi, China).

Subsequently, we determined the enzyme activities of two *Juglans* species salt-treatment and control groups leaves spectrophotometrically as follows. First, we took fresh leaves and added them to the extract for ice bath homogenization, centrifuged at 8000 g 4℃ and supernatant was taken. Then, superoxide dismutase (SOD), peroxidase (POD), and catalase (CAT) were determined by the spectrophotometric method using SOD, POD, and CAT assay kits (MolFarming, Nanjing China). All physiological indices were measured in three replicates of 0.1–0.2 g of fresh leaves. The mixture containing 0.3 mL of 260 mM methionine, 0.3 mL of 100 μM EDTA-Na2, 0.3 mL of 750 μM NBT, 0.3 mL of 20 μM riboflavin, 1.8 mL of variable-volume phosphate buffer, and extract was assayed for SOD activity at 560 nm. The POD activity was determined by monitoring the absorbance value at 470 nm of 1 mL of extract dissolved in 3 mL of mixed solution (containing 28 μL guaiacol and 19 μL H_2O_2 per 50 mL). The CAT activity was determined by monitoring the absorbance value at 240 nm of 500 μL of extract mixed into a 2.5 mL mixture containing 2 mL phosphate buffer and 500 μL H_2O_2 [63].

Gene expression of *4CLs*

We determined the gene expressions of *4CLs* in two *Juglans* species. The samples from our previous collection included different tissues/organs (leaves, green husks, female and male flowers) [62, 64, 65]. All 24 different tissue/organ samples were sequenced by the Illumina HiSeq X Ten platform (Illumina, San Diego, CA, USA). Then, map all transcriptome clean reads to the reference genome using HISAT2 software [66]. Gene expression levels were calculated using fragments per kb of transcript sequence per million bp sequenced (FPKM) [67]. Calculate FPKM values using FeatureCounts software [68]. Sequencing of walnut fruit disease resistance data was downloaded from the public NCBI database [69]. These contain anthracnose-resistant varieties (F26) and anthracnose-susceptible varieties (F423). The K-means clustering was used to normalize *Jr4CL* genes fruit disease resistance data [70]. The gene expression level heatmaps were drawn using TBtools software [50].

Subsequently, the gene expression of the identified *4CL* genes under salt stress was further explored by qRT-PCR experiments. Each treatment has 3 biological replicates. Total RNA was isolated from leaves of two *Juglans* species using a plant RNA isolation kit (OMEGA, USA). Quality assessment of total RNA based on A260/A280 ratio using the Nanodrop spectrometer (KAIAO, Beijing, China). Reverse transcription to complementary DNA (cDNA) using quality-tested RNA. cDNA template was obtained by reverse transcription using 5× PrimeScript RT Master Mix (Takara) reverse transcriptase. We diluted the cDNA as 5-fold as the template DNA for qRT-PCR experiment. Using 2×Plus SYBR real-time PCR mixture (Biotec) as fluorescent dye, a mixture containing 10 μL 2×plus SYBR real-time PCR mixture, 0.5 μL Primer F, 0.5 μL Primer R, 2 μL cDNA, and 7 μL nuclease-free water was used for qRT-PCR experiments on Bio-Rad CFX96 fluorescence quantitative PCR instrument [62]. The reaction procedure was pre-denaturation at 94 °C for 2 min, denaturation at 94 °C for 15 s, annealing at 58 °C for 15 s, extension at 72 °C for 30 s, and 40 cycles. The reaction was subsequently terminated by plate reading at 95 °C for 5 s after 65 °C dissolution curve capture. *J. regia* β-actin was used as an internal reference gene [62, 71] and the primers were designed with the online Primer3Plus website (https://www.primer3plus.com). The relative expression of all *4CLs* was normalized by the $2^{-\Delta\Delta CT}$ method [72]. All the primer sequences are shown in Table S1.

Results

Identification and phylogenetic relationship of *4CL* genes in *J. regia* and its wild relatives *J. mandshurica*

We identified total of 36 *Jr4CL* genes and 31 *Jm4CL* genes (Table S2). To facilitate subsequent studies, we renamed all genes according to their position order on the chromosome. Table S2 listed the baseline information for all identified *4CL* gene members.

The maximum likelihood phylogenetic tree was performed using 4CL protein sequences of *Arabidopsis* (13), *O. sativa* (14), *P. bretschneideri* (29), *J. regia* (36), and *J. mandshurica* (31). Then, we classified their evolutionary

relationships based on the phylogenetic tree (Fig. 1). All 4CL gene members were divided into three clades (Clade I, II, and III). The largest clade is Clade III, which contained 1 *At4CL*, 18 *Jr4CLs*, 17 *Jm4CLs*, and 16 *Pbr4CLs*, followed by the Clade II, which contained 8 *At4CLs*, 8 *Os4CLs*, 8 *Pbr4CLs*, 11 *Jr4CLs*, and 10 *Jm4CLs*. The remaining *4CL* members were assigned to Clade I, which contained 4 *At4CLs*, 6 *Os4CLs*, 5 *Pbr4CLs*, 7 *Jr4CLs*, and 4 *Jm4CLs*. All four species were distributed in all branches except *Os4CL*, which was distributed in only two branches (Clade I and II). In Clade I-III, three woody plants (*J. regia*, *J. mandshurica*, and *P. bretschneideri*) were clustered, such as (1) *Jm4CL15* and *Jr4CL31*; (2) *Pbr4CL11, Pbr4CL30, Jm4CL14*, and *Jr4CL32*; (3) *Jr4CL12, Jm4CL5, Jr4CL29*, and *Jm4CL6* were clustered in the same branch in Clade I. (4) *Pbr4CL32*, *Jr4CL16* and *Jm4CL1*; (5) *Jr4CL25*, *Jm4CL20, Jr4CL17, Jm4CL22, Jr4CL24*, and *Jm4CL19* were clustered in the same branch in Clade II. (6) *Pbr4CL35, Jr4CL8*, and *Jr4CL10*; (7) *Pbr4CL2, Jr4CL2*, and *Jm4CL11* were clustered in the same branch in Clade III. These results suggested that three woody plants are more closely related.

Physicochemical analyses and subcellular localization prediction of 4CL members

We predicted the physicochemical properties of the 4CL proteins of the two *Juglans* species with the following results (Table 1). The mean length of Jr4CL protein ranged from 523 aa (Jr4CL4) to 694 aa (Jr4CL14), with a mean length of 564 aa. The mean length of Jm4CLs was consistent with that in Jr4CLs, but ranged more widely, from 307 aa (Jm4CL31) to 1083 aa (Jm4CL28). The molecular weights of Jr4CLs varied from 56369.53 kDa (Jr4CL21) to 76612.12 kDa (Jr4CL14) with a mean molecular weight of 61718.09 kDa. The molecular weights of Jm4CLs were greater than those of Jr4CLs, which varied from 34330.35 kDa (Jm4CL31) to 117466.84 (Jm4CL28), with a mean of 61800.35 kDa. In addition, there are 23 and 16 acidic proteins (isoelectric point less than 7) in Jr4CLs (Jr4CL1, Jr4CL, Jr4CL4, Jr4CL6, Jr4CL7, Jr4CL12, Jr4CL13, Jr4CL15, Jr4CL16, Jr4CL18, Jr4CL19, Jr4CL20, Jr4CL21, Jr4CL26, Jr4CL27, Jr4CL28, Jr4CL29, Jr4CL30, Jr4CL31, Jr4CL32, Jr4CL34, Jr4CL35, and Jr4CL36) and Jm4CLs (Jm4CL2, Jm4CL5, Jm4CL6, Jm4CL8, Jm4CL9, Jm4CL14, Jm4CL15, Jm4CL16, Jm4CL17, Jm4CL21, Jm4CL24, Jm4CL25, Jm4CL26, Jm4CL27, Jm4CL28, and Jm4CL31), respectively. It was also found that the Jr4CLs and Jm4CLs clustered in Clade I were all acidic proteins (Fig. 1). Most of the members of Jr4CLs and Jm4CLs had instability indices less than 40 (26 of 36 for Jr4CL and 21 of 31 for Jm4CL), and the total mean hydrophilicity of 18 Jr4CLs (50%) and 21 Jm4CLs (67.7%) was negative, so most of the Jm4CLs were considered to be stable hydrophilic proteins. Additionally, most of the identified 4CL members were located on the plasma membrane, which may be related to the function of 4CL proteins (Table 1).

Chromosomal localization and duplication mode of *4CL*s

For *J. regia*, 36 *4CLs* were randomly and irregularly located in 15 different chromosomes, except for chromosome 6 (Fig. 2A). Chromosome 11 had the highest *Jr4CL* density of 16.67%, followed by chromosome 7 with *Jr4CL* density of 13.89%. Next was chromosome 1, with *Jr4CL* density of 11.11%. Chromosome 2, chromosome 9, and chromosome 13 all contained 3 *Jr4CLs* with a gene density of 8.33%. Chromosome 4, chromosome 8, and chromosome 16 contained 2 *Jr4CLs* with a gene density of 5.56%, respectively. For *J. mandshurica*, 31 *Jm4CLs* were distributed on 14 different chromosomes, with the exception of chromosomes 9 and 15 (Fig. 2B). Chromosome 1 had the largest *Jm4CL* gene density of 16.13%, followed by chromosome 3 with a gene density of 12.90%. Chromosome 2, chromosome 5, and chromosome 11 contained 3 *Jm4CLs* with a gene density of 9.68%. Chromosome 6, chromosome 7, chromosome 8, and chromosome 12 contained 2 *Jm4CLs* with a gene density of 6.45%. The other *4CL*s were situated on distinct chromosomes.

To investigate the reasons for the expansion of the *Jr4CL* and *Jm4CL* gene family, we analyzed the duplication patterns of the *4CL* genes. Gene duplication pattern of the *4CL* genes contained four models, consisting of whole genome duplication (WGD), tandem duplication (TD), proximal duplication (PD), and dispersed duplication (DSD; Table S3) [73]. We found that WGD was the predominant model in both *Juglans* species. WGD duplication patterns accounted for 17 of the 36 *Jr4CLs* (47.22%) and 20 of the 31*Jm4CLs* (64.52%). Nine *4CL*genes in both *J. regia* (*Jr4CL6, Jr4CL8, Jr4CL10, Jr4CL13, Jr4CL15, Jr4CL16, Jr4CL32, Jr4CL34, and Jr4CL35*) and *J. mandshurica* (*Jm4CL1, Jm4CL4, Jm4CL14, Jm4CL15, Jm4CL17, Jm4CL20, Jm4CL25, Jm4CL30, Jm4CL31*) were experienced DSD, respectively. Both *4CL* genes in *Jr4CLs* (*Jr4CL26* and *Jr4CL27*) and *Jm4CLs* (*Jm4CL11* and *Jm4CL12*) underwent TD, respectively. There were 3 *Jr4CLs* (*Jr4CL1, Jr4CL2*, and *Jr4CL28*) experienced PD. Notably, the *Jr4CL9, Jr4CL18, Jr4CL20, Jr4CL25*, and *Jr4CL31* were identified as singleton and did not experience any of the above four duplication patterns.

Conserved domains and gene structures of 4CL members

The results of the maximum likelihood tree were constructed based on two *Juglans* species (Fig. 3A) similar to the phylogenetic tree clustering results constructed using five selected species (Fig. 1). The results indicated that the structural domains of the *4CL* gene family members were conserved to a high degree. All identified 4CL proteins contained 4CL domain and conserved kinase

Table 1 The predicated protein information of 4CL in *Juglans regia* and *J. mandshurica*

Gene name	No. of amino acids	Mol. Wt (Da)	Isoelectric point (pI)	Instability index (II)	Aliphatic index	Grand average of hydropathicity (GRAVY)	Subcellular localization[a]
Jr4CL1	553	60808.00	6.95	42.81	90.72	-0.058	pero
Jr4CL2	555	60992.05	6.83	38.14	86.00	-0.087	pero
Jr4CL3	582	64270.92	7.60	34.50	90.62	-0.136	pero
Jr4CL4	523	56425.31	5.72	34.81	94.21	-0.032	plas
Jr4CL5	561	61038.41	8.19	46.91	99.73	0.051	plas
Jr4CL6	567	62023.66	6.25	35.81	97.94	0.127	plas
Jr4CL7	524	56455.41	6.36	37.83	92.75	-0.014	plas
Jr4CL8	577	63187.75	8.94	43.78	91.82	-0.054	plas
Jr4CL9	542	59471.00	8.73	41.64	97.08	0.034	plas
Jr4CL10	564	62168.99	7.00	43.88	99.40	0.077	plas
Jr4CL11	540	59266.45	8.22	38.75	102.11	0.071	plas
Jr4CL12	544	59499.91	6.06	41.19	96.76	-0.013	chlo
Jr4CL13	565	61938.44	6.79	35.70	92.53	-0.027	pero
Jr4CL14	694	76612.12	7.17	31.63	89.88	-0.083	chlo
Jr4CL15	570	62344.13	5.49	37.67	90.00	-0.125	chlo
Jr4CL16	573	62260.32	5.86	41.61	97.03	-0.020	nucl
Jr4CL17	549	59487.76	8.67	37.30	102.46	0.095	plas
Jr4CL18	590	65202.60	6.80	33.51	82.10	-0.155	chlo
Jr4CL19	557	60340.56	6.25	45.30	97.67	0.069	plas
Jr4CL20	558	60322.60	6.71	42.86	97.87	0.085	E.R.
Jr4CL21	525	56369.53	6.22	34.54	96.82	0.015	plas
Jr4CL22	587	65052.87	8.17	32.67	91.70	-0.147	pero
Jr4CL23	563	61944.21	8.39	32.91	83.85	-0.132	pero
Jr4CL24	550	59855.19	8.63	34.35	99.07	0.044	plas
Jr4CL25	577	63203.98	9.05	38.56	96.46	-0.044	plas
Jr4CL26	550	60633.02	5.95	32.01	98.05	-0.013	chlo
Jr4CL27	543	59784.21	5.54	31.36	99.50	-0.026	chlo
Jr4CL28	542	59606.04	5.61	30.37	99.52	0.001	chlo
Jr4CL29	545	59330.79	5.69	34.70	100.35	0.040	chlo
Jr4CL30	574	62588.95	5.42	33.74	93.97	0.014	chlo
Jr4CL31	558	61012.74	5.81	36.82	97.96	0.018	cyto
Jr4CL32	572	62323.73	5.54	38.05	99.14	0.070	plas
Jr4CL33	544	59498.00	8.80	39.12	100.48	0.079	plas
Jr4CL34	545	60320.58	6.12	34.17	97.06	0.077	cyto
Jr4CL35	661	73839.49	5.98	40.94	84.66	-0.167	nucl
Jr4CL36	573	62371.41	6.86	34.48	97.89	0.084	chlo
Jm4CL1	603	65862.31	7.74	44.42	93.81	-0.130	nucl
Jm4CL2	589	64468.68	6.28	34.78	88.78	-0.206	chlo
Jm4CL3	685	76413.19	8.33	36.05	89.78	-0.077	chlo
Jm4CL4	536	59092.98	9.23	40.47	95.00	0.037	pero
Jm4CL5	544	59472.84	5.87	41.12	96.05	-0.008	chlo
Jm4CL6	540	58924.38	6.50	34.84	98.93	-0.002	chlo
Jm4CL7	582	64207.82	8.51	29.47	90.96	-0.062	chlo
Jm4CL8	568	62025.13	6.62	33.32	85.67	-0.169	chlo
Jm4CL9	523	56473.35	5.75	35.19	93.84	-0.049	plas
Jm4CL10	428	47362.40	7.60	36.43	90.23	-0.164	pero
Jm4CL11	399	44252.13	8.08	42.10	89.37	-0.027	pero
Jm4CL12	396	43952.83	8.53	41.52	90.08	-0.059	pero
Jm4CL13	587	64958.71	7.61	32.69	91.70	-0.143	pero
Jm4CL14	523	56955.42	5.39	37.61	97.99	0.079	plas
Jm4CL15	558	60958.67	5.75	35.88	98.49	0.024	cyto
Jm4CL16	558	61082.29	6.65	30.31	90.34	-0.063	chlo

Table 1 (continued)

Gene name	No. of amino acids	Mol. Wt (Da)	Isoelectric point (pI)	Instability index (II)	Aliphatic index	Grand average of hydropathicity (GRAVY)	Subcellular localization[a]
Jm4CL17	661	73931.66	5.86	41.91	84.66	-0.176	nucl
Jm4CL18	537	58783.23	9.08	37.37	94.47	-0.012	chlo
Jm4CL19	539	58507.49	8.09	37.78	98.94	0.073	plas
Jm4CL20	577	63220.02	9.05	38.36	97.14	-0.033	plas
Jm4CL21	698	76909.30	6.68	37.26	90.64	-0.070	chlo
Jm4CL22	549	59439.76	8.53	38.35	103.53	0.113	plas
Jm4CL23	590	65296.44	8.78	37.21	97.31	0.003	plas
Jm4CL24	530	57363.95	5.92	43.57	100.06	0.097	plas
Jm4CL25	573	62693.39	6.11	37.66	98.27	0.132	plas
Jm4CL26	524	56484.37	6.33	37.26	90.88	-0.029	plas
Jm4CL27	527	56712.87	6.22	34.77	96.26	-0.002	plas
Jm4CL28	1083	117466.84	6.06	44.90	93.01	0.012	plas
Jm4CL29	669	73445.78	8.53	38.44	98.16	-0.008	plas
Jm4CL30	497	54762.36	9.03	40.10	100.97	0.046	E.R.
Jm4CL31	307	34330.35	5.55	46.49	94.92	-0.013	cyto

[a]*Note* pero: peroxisome; plas: plasma membrane; chlo: chloroplast; nucl: nucleus; E.R.: endopiasmic reticulum; cyto: cytoskeleton

structural domains (Fig. 3B). Sequence alignment also suggested that the 4CL proteins of two *Juglans* species were highly conserved in BOX I (SSGTTGLPKGV) and BOX II (GEICIRG) regions (Fig. S2).

Variations of gene structure may affect differences in gene function, so we recognized the gene structure of all identified *4CL*s (Fig. 3C; Table S4). The gene structure results showed that all identified *4CL* genes differed significantly in their structures. The exon count of *Jr4CLs* was from 1 to 23, and the exon count of *Jm4CLs* was from 2 to 23. Among them, 9 *Jr4CLs* contained 6 exons, followed by 7 *Jr4CLs* contained 5 exons and 5 *Jr4CLs* contained 2 exons. In *J. mandshurica*, 7 *Jm4CLs* contained 6 exons, 6 *Jm4CLs* contained 5 exons, and 5 *Jm4CLs* contained 3 exons. *Jr4CL14* and *Jm4CL21* contained the highest number of exons at 23. Notably, individual *4CL* genes contain longer introns, especially Jm*4CL28*.

Synteny analysis of *4CL* genes

The collinearity analysis revealed 11 and 15 paralogous gene pairs in *Jr4CLs* and *Jm4CLs*, respectively (Fig. 4; Table S5). There were a total of 48 *4CL* orthologous gene pairs between *Jr4CLs* and *Jm4CLs* (Fig. 4; Table S5). The number of orthologous gene pairs was much greater than the number of paralogous gene pairs, suggesting a high degree of collinearity in *Jr4CLs* and *Jm4CLs*. Moreover, in two *Juglans* species 8 *Jr4CLs* (*Jr4CL8, Jr4CL10, Jr4CL14, Jr4CL20, Jr4CL23, Jr4CL26, Jr4CL27, Jr4CL28, Jr4CL31*) and 3 *Jm4CLs* (*Jm4CL1, Jm4CL12, Jm4CL15*) without paralogous gene pairs, suggest that they may be *J. regia* or *J. mandshurica* specific genes that did not collinearity. To expose the selection pressure among *4CL* homologous gene pairs, we computed Ka and Ks parameters (Table S5). The Ka/Ks ratios of all *4CL* homologous gene pairs were smaller than 1, indicating that all gene pairs underwent purifying selection and may have evolved relatively slowly.

Subsequently, to investigate the possible evolutionary processes of *4CL*s, we performed collinearity analyses between two *Juglans* species and three selected plants, including one monocotyledon (*O. sativa*) and two dicotyledon species (*Arabidopsis* and *P. bretschneideri*). The *Jr4CLs* and *Jm4CLs* have 20 and 19 orthologous gene pairs with *Arabidopsis*, respectively (Fig. S3A; Table S6). Jr4CLs and Jm4CLs have 8 and 4 orthologous gene pairs with *O. sativa* (Fig. S3B; Table S7), and have 26 and 26 orthologous gene pairs with *P. bretschneideri*, respectively (Fig. S3C; Table S8). In our present study, the number of *4CL* homologous gene pairs was higher in two *Juglans* species and dicotyledon plants than with monocotyledon plants. While in both dicotyledon species, the number of *4CL* homologous gene pairs was higher in two *Juglans* species and *P. bretschneideri* than *Arabidopsis*, suggesting that two *Juglans* species are more closely evolutionarily related to *P. bretschneider* than to *Arabidopsis*. Notably, 2 *Jr4CLs* (*Jr4CL5* and *Jr4CL13*) and 2 *Jm4CLs* (*Jm4CL4*and *Jm4CL24*) had homologous gene pairs with all three selected species, suggesting that these 4 homologous gene pairs may have existed before the differentiation of monocotyledon and dicotyledon plants.

Analysis of *cis*-acting elements and GO annotation for *4CLs*

To understand the potential function of the *4CL* genes in *J. regia* and *J. mandshurica*, we analyzed *cis*-acting elements of the promoter regions and GO functional annotation (Fig. 5). The *cis*-acting elements of the promoter regions were divided into four categories, including plant development and growth, phytohormone response,

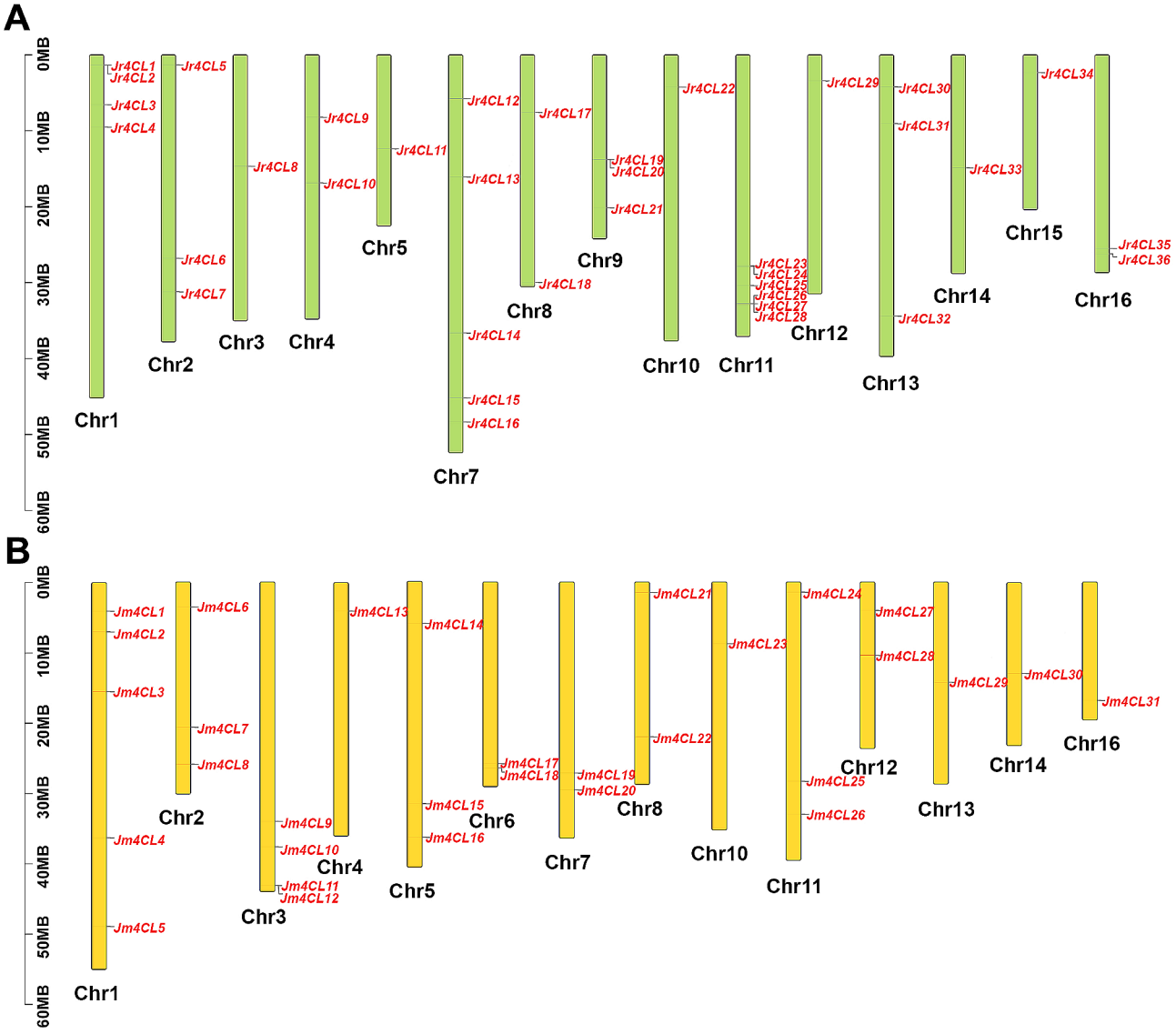

Fig. 2 The locations of *Jr4CL* genes (A) and *Jm4CL* genes (B)

abiotic stress response, and light responsiveness (Fig. 5). Overall, a greater number of *cis*-acting elements were found in *Jr4CLs* than in *Jm4CLs*, suggesting that *Jr4CLs* may be involved in more complex signal transduction pathways. Most of the identified *4CL* genes contain *cis*-acting elements associated with light responsiveness, and the *Jr4CLs* promoter regions had more *cis*-acting elements responsive to light than *Jm4CLs*, suggesting that *Jr4CLs* may be more sensitive to light. It is noticed that we found a significant proportion of identified *4CL* members respond to three *cis*-acting elements associated with Methyl Jasmonate (MeJA) hormone response and Abscisic Acid (ABA) hormone response. In addition, the promoter regions of *4CL cis*-acting elements were mostly associated with abiotic stresses, suggesting that they might be important in abiotic stress resistance of *Jr4CLs* and *Jm4CLs*, especially in anaerobic induction (LTR).

The GO enrichment analysis of all identified *4CL* members showed that the fatty acid biosynthetic process, fatty acid metabolic process, jasmonic acid biosynthetic process, small molecule metabolic process, cellular lipid metabolic process, and ketone biosynthetic process were the top six BPs (Biological processes) subgroups in *J. regia* (Fig. S4A). The microbody and peroxisome were the top two cellular components (CCs) subgroups, while the fatty acid ligase activity and acid-thiol ligase activity were the top two molecular functions (MFs) subgroups (Fig. S4A). For *J. mandshurica*, the jasmonic acid biosynthetic process, jasmonic acid metabolic process, carboxylic acid metabolic process, small molecule metabolic process, and cellular lipid metabolic process were the top

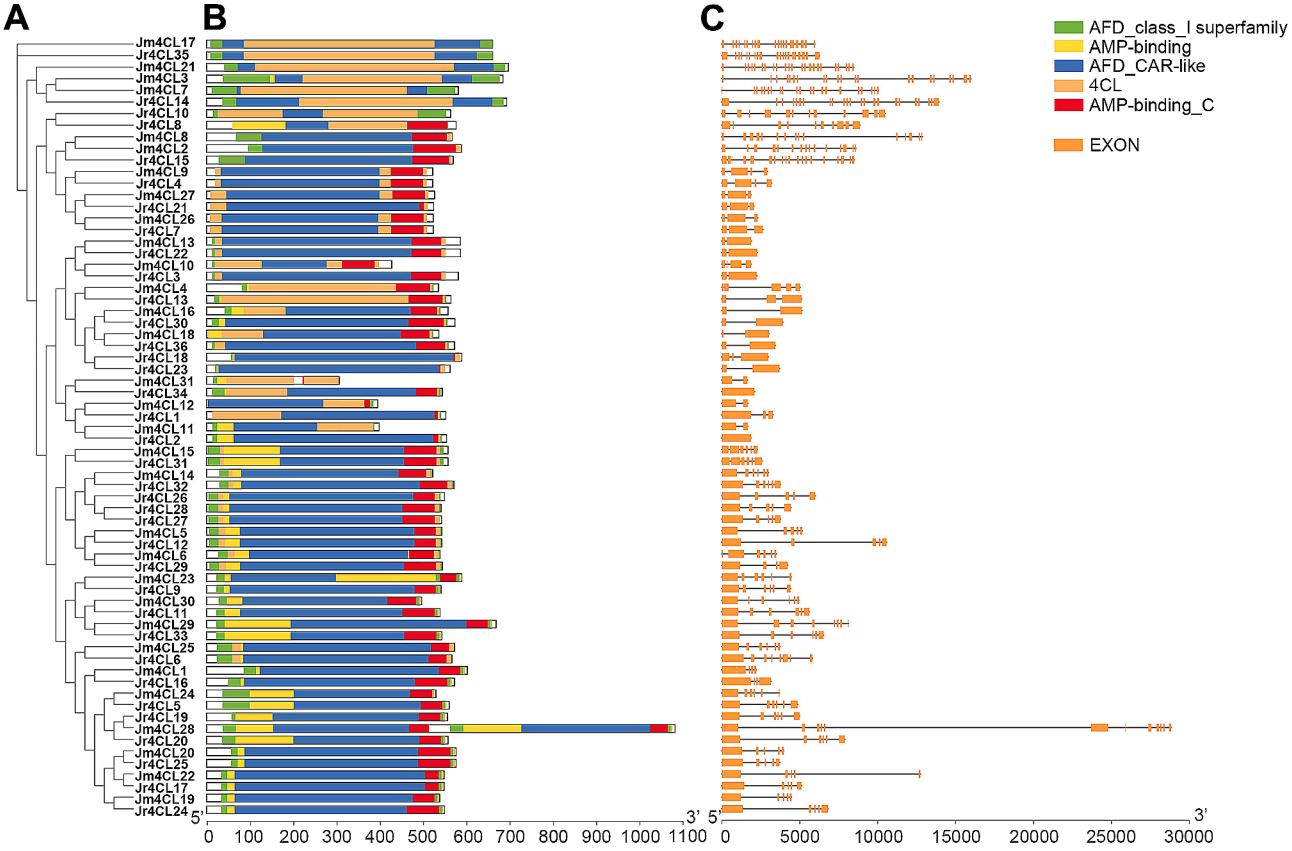

Fig. 3 The characterization of 4CL members. (A) Phylogenetic tree of *4CLs* in two *Juglans* species; (B) Conserved domains of 4CL proteins. (C) Gene structures of *4CLs*. Orange boxes and gray lines indicate exons and introns, respectively

five BPs subgroups, while the microbody and peroxisome were the top two CC subgroups. The top three MF subgroups include the CoA-ligase activity, acid-thiol ligase activity, and ligase activity, forming carbon-sulfur bonds (Fig. S4B).

Prediction of protein-protein interaction and miRNAs targeting of 4CL members

We used the homology mapping method to predict the interactions of 4CL proteins in two *Juglans* species based on the interactions of 4CL proteins in *Arabidopsis* (Table S9). As shown in Fig. 6A, B, both Jr4CLs and Jm4CLs mapped to 11 *Arabidopsis* proteins, and all identified 4CL proteins mainly interacted with CHS proteins. In addition, compared to Jm4CLs, Jr4CLs interacted with more proteins, suggesting that Jr4CLs were involved in more complex signaling pathways.

A total of 260 microRNAs were predicted to target 35 *Jr4CLs* (except *Jr4CL3*) and 218 microRNAs were predicted to target 30 *Jm4CLs* (except *Jm4CL10*, Fig. 6C, Table S10). Among them, 356 miRNAs modulated the expression of identified *4CLs* by cleavage, and 122 miRNAs modulated by translation (Fig. S4). Notably, *Jr4CL34* and *Jm4CL2* were the target genes of 15 and 19 different miRNAs, respectively, and were the genes with the highest number of miRNA targets. In addition, some miRNAs target different *4CL* genes, such as the ath-miR865-3p targets 12 different *4CL* genes at the same time.

Gene expression levels of *Jr4CLs* and *Jm4CLs*

To investigate the expression pattern of *4CL* genes, we analyzed the expression of all identified *4CLs* in four selected organs, including leaves, green husks, male, and female flowers based on transcriptomic data (Fig. 7A and B). According to the transcriptome results, all the identified 36 *Jr4CLs* were expressed in four selected organs. Among them, 14 *Jr4CLs* were highly expressed in leaves, *17 Jr4CLs* in green husks, 8 *Jr4CLs* in female flowers, and *8 Jr4CLs* in male flowers, respectively (Fig. 7A; Table S11). However, for *J. mandshurica*, only 21 *Jm4CLs* were expressed in four selected organs. Of these, there were 6 *Jm4CLs* expressed highly in leaves, 9 *Jm4CLs* in green husks, 8 *Jm4CLs* in male flowers, and 12 *Jm4CLs* in female flowers, respectively (Fig. 7B; Table S12). The different expression patterns in two *Juglans* species suggested that all *Jr4CLs* have roles in the growth and development of the selected organs, while *Jm4CLs* may not function in some organs of *J. mandshurica*. Additionally,

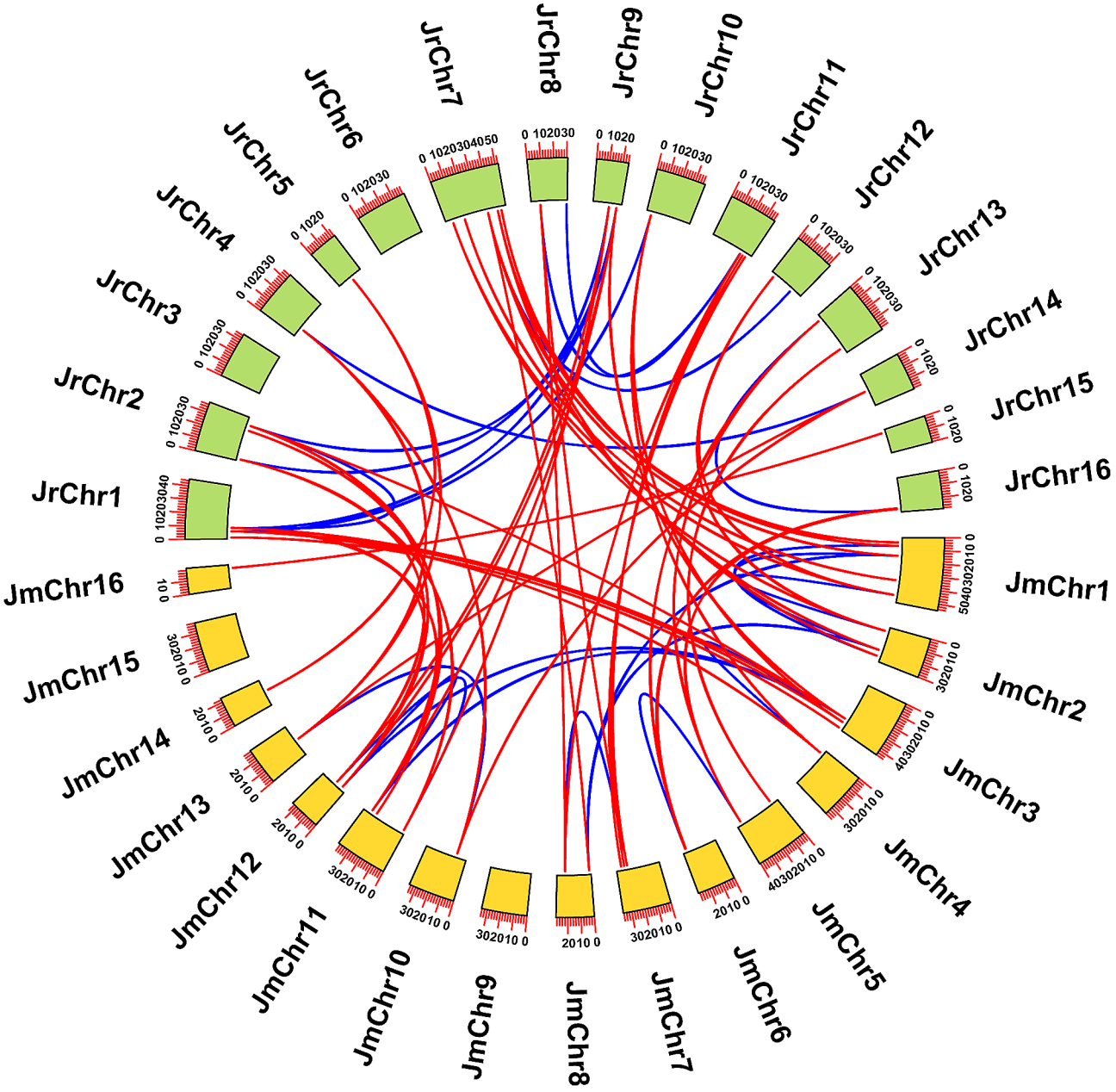

Fig. 4 Collinearity analysis for *4CL* genes among two *Juglans* species. Red and blue lines indicate orthologous and paralogous gene pairs, respectively

6 *Jr4CLs* and 7 *Jm4CLs* were only expressed at high levels in female or male flowers (Fig. 7). It is suggested that these genes might be related to the heterodichogamous.

Genes expression patterns of *Jr4CLs* under biotic stress

To explore the role of *Jr4CLs* in biotic stress (disease) response, we examined the transcriptome gene expression levels of different walnut varieties (F26 and F423) under biotic stress (Fig. 7C; Table S13). Overall, F26 had higher levels of *4CL* gene expression than F423, suggesting that *Jr4CLs* may function in *J. regia* anthracnose resistance. *Jr4CL9* and *Jr4CL13* expression gradually decreased with time after infection with F423 varieties, whereas expression increased with time after infection with F26 varieties, which suggests that they could be associated with the resistance to anthracnose in walnuts.

Based on K-means clustering analysis, the expression trends of both F26 and F423 were classified into 9 subgroups (Fig. S6; Table S14). Most of the *Jr4CLs* showed a gradual increase in expression over 24 h in F26 varieties, reached peak expression at 24 h, followed by a gradual decrease, and then increased again after 72 h (Fig. S5A). In contrast, in the F423 variety, most of the *Jr4CLs*

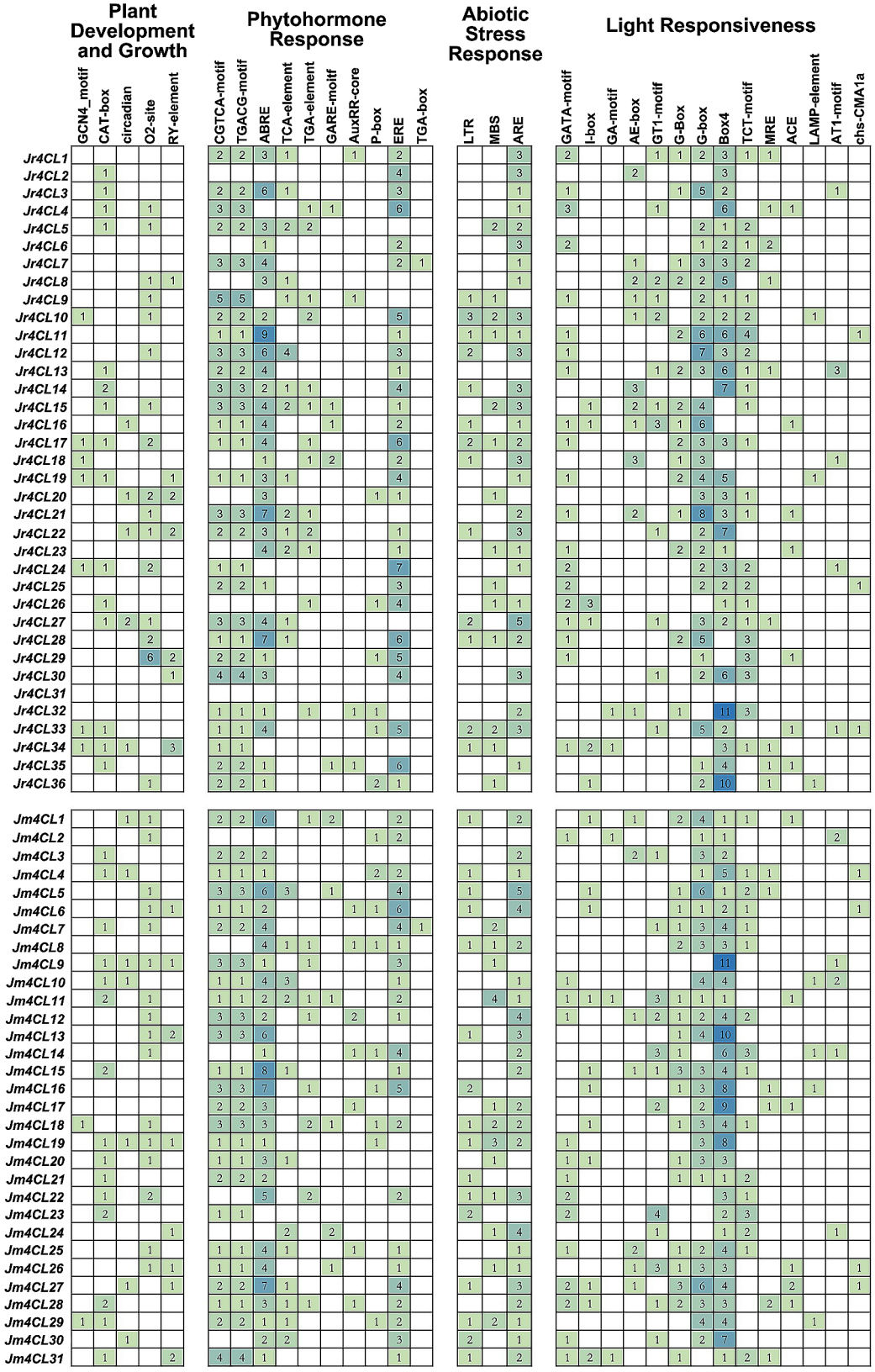

Fig. 5 The analysis of *Cis*-acting elements. The colored numbers indicate the number of *cis*-acting elements

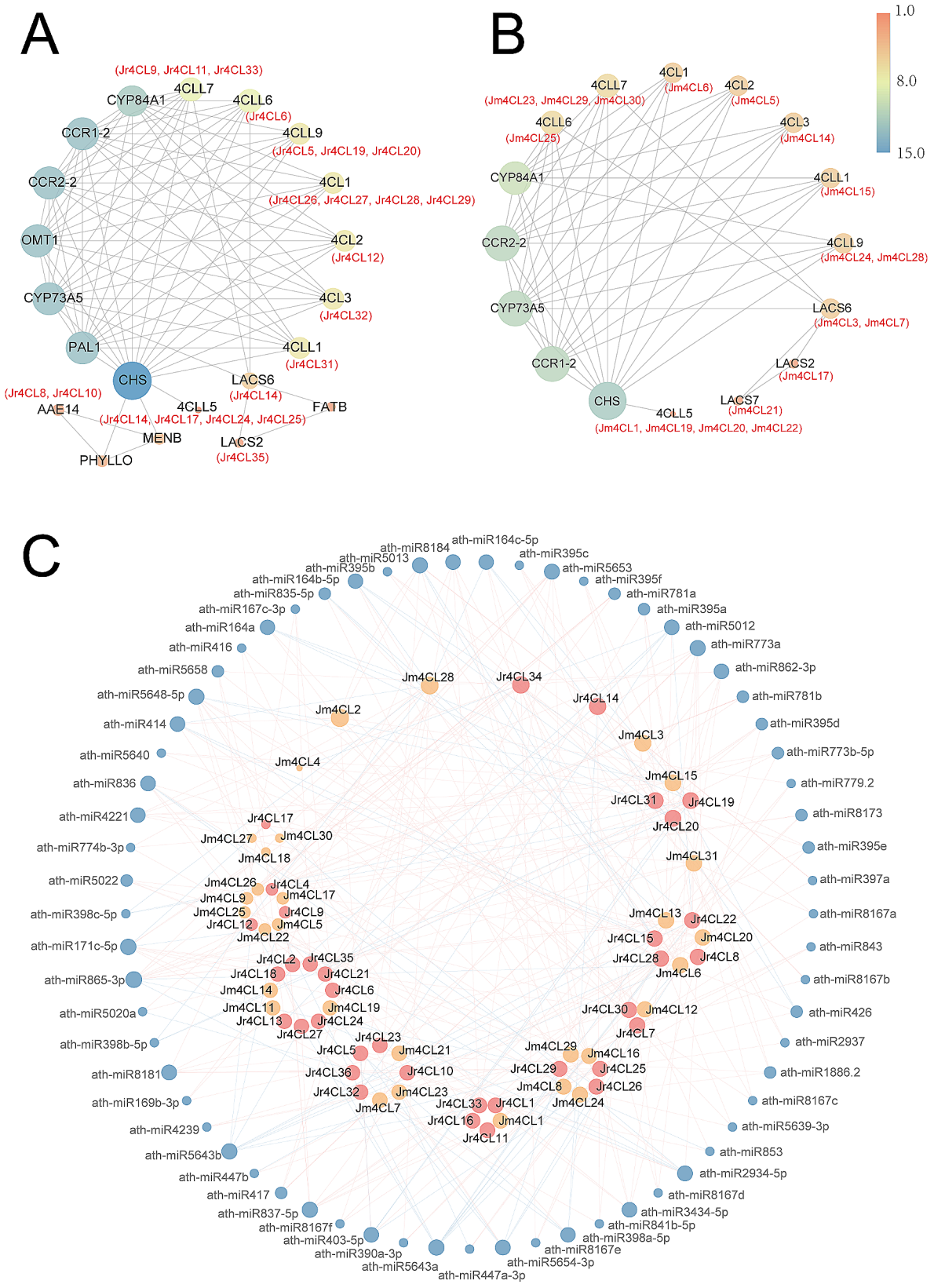

Fig. 6 Protein-protein interaction and microRNA target of 4CL members. (A) Protein-protein interaction of the Jr4CL proteins; (B) Protein-protein interaction of the Jm4CL proteins. Colored circles represent proteins, grey lines indicate interaction; (C) MicroRNA targeting of the *4CL* genes in two *Juglans* species. The blue circle, red circle, and orange circle represent microRNAs, *Jr4CLs*, and *Jm4CLs*, respectively. The size of the circle represents how much of the targeting relationship. Blue lines represent translation and red lines represent cleavage

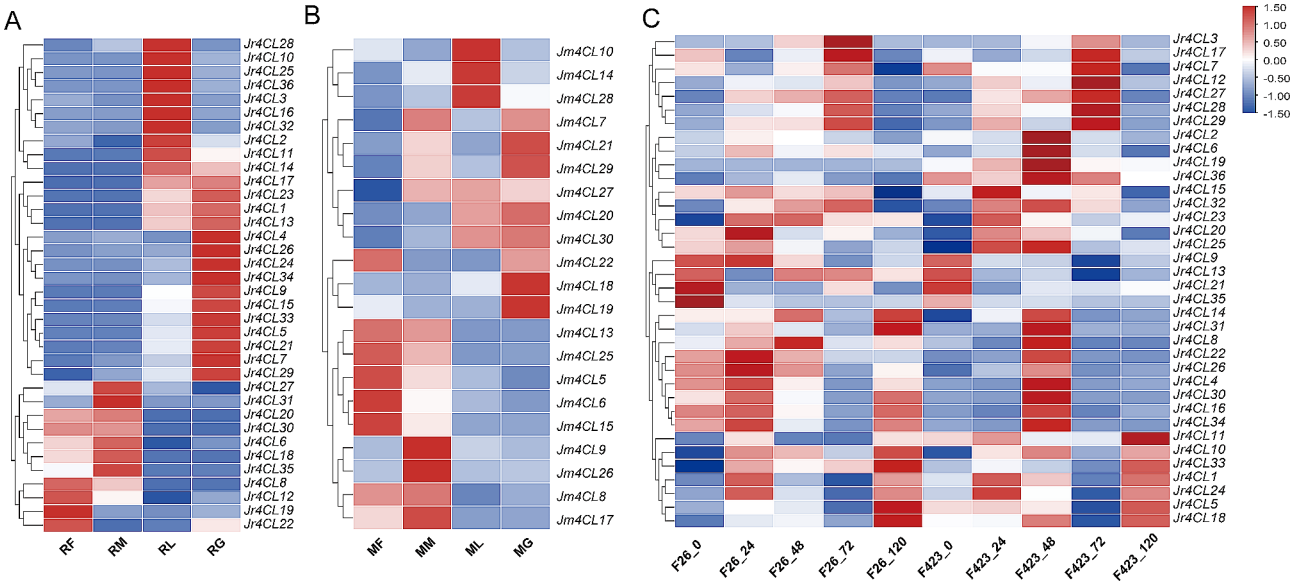

Fig. 7 Gene expression levels of *4CLs*. (A) *Jr4CLs* and (B) *Jm4CLs* in different organs; (C) Gene expression profiles of *Jr4CLs* under biotic stress

reached the highest expression at 48 h, followed by a gradual decrease (Fig. S5B).

Genes expression and antioxidant enzyme activity changes of *4CLs* under salt stress

To better study the differences between two *Juglans* species under salt stress, we separately salt-treated them at the same developmental period leaves and assayed the enzyme activities (Fig. 8). Based on the phenotype results, it can be seen that both *Juglans* species leaves were damaged to varying degrees after salt treatment as compared to the control (Fig. 8A). There were visible black patches on the *J. regia* leaves, and the *J. mandshurica* leaves were significantly wrinkled although did not show black patches compared to those before the salt treatment. Compared to the control, CAT activity increased 1.464 and 1.707 times, and POD activity increased 1.288 and 2.077 times after salt treatment in *J. regia* and *J. mandshurica* leaves, respectively. However, SOD activity decreased 1.383 times in *J. regia* and increased 1.715 times in *J. mandshurica* leaves after salt treatment (Fig. 8B). In addition, the levels of all three enzyme activities were higher in *J. mandshurica* leaves than in *J. regia* leaves. All these results indicate that *J. mandshurica* performed better than *J. regia* under salt stress.

To better understand the potential roles of these genes in salt tolerance in two *Juglans* species, we further analyzed the *4CL* gene expression levels in leaves under salt stress using qRT-PCR (Fig. 9; Fig. S7). There were 28 *Jr4CLs* (except *Jr4CL4*, *Jr4CL8*, *Jr4CL18*, *Jr4CL19*, *Jr4CL29*, *Jr4CL30*, *Jr4CL31*, and *Jr4CL34*) and 19 *Jm4CLs* (except *Jm4CL1*, *Jm4CL2*, *Jm4CL3*, *Jm4CL4*, *Jm4CL11*, *Jm4CL12*, *Jm4CL13*, *Jm4CL15*, *Jm4CL16*, *Jm4CL23*, *Jm4CL24*, and *Jm4CL31*) expressed in leaves, respectively (Fig. 7A and B; Table S11; S12). Among the expressed 28 *Jr4CL* genes, total of 21 *Jr4CLs* (75%) increased the expression level in leaves after salt treatment leaves, while 7 *Jr4CL* genes (25%) decreased the expression level. Among the 19 *Jm4CLs*, 15 *Jm4CLs* (78.95%) increased in expression and 4 *Jm4CLs* (21.05%) decreased in expression after salt treatment. These results suggest that *Jr4CLs* and *Jm4CLs* may respond to salt stress through two different modes (positive and negative) of regulation, of which positive regulation may be predominant. Notably, the expression levels of some *4CL* genes increased dramatically after salt treatment. For example, *Jr4CL17* increased by 30.4 times, *Jr4CL28* increased by 44.6 times, *Jm4CL27* increased by 36.84 times, and *Jm4CL28* increased by 41.299 times, respectively. It is suggested that these *4CLs* may function in *4CL* gene resistance to salt stress.

Discussion

Two *Juglans* species were both economically and ecologically important woody tree species. Recently, the comparative genomic studies of *Juglans* species have become a research hotspot [9, 11, 74]. 4CL is one of the key enzymes in phenylpropanoid metabolism pathway and the final step in phenylpropanoid synthesis pathway [75]. The level of 4CL enzyme activity had a significant impact on the accumulation of compounds, such as flavonoids and lignin. Furthermore, *4CLs* have key roles in plant growth and resistance to environmental stresses from outside [76]. We identified a total of 36 and 31 *4CL* gene members in *J. regia* and *J. mandshurica*, respectively. The

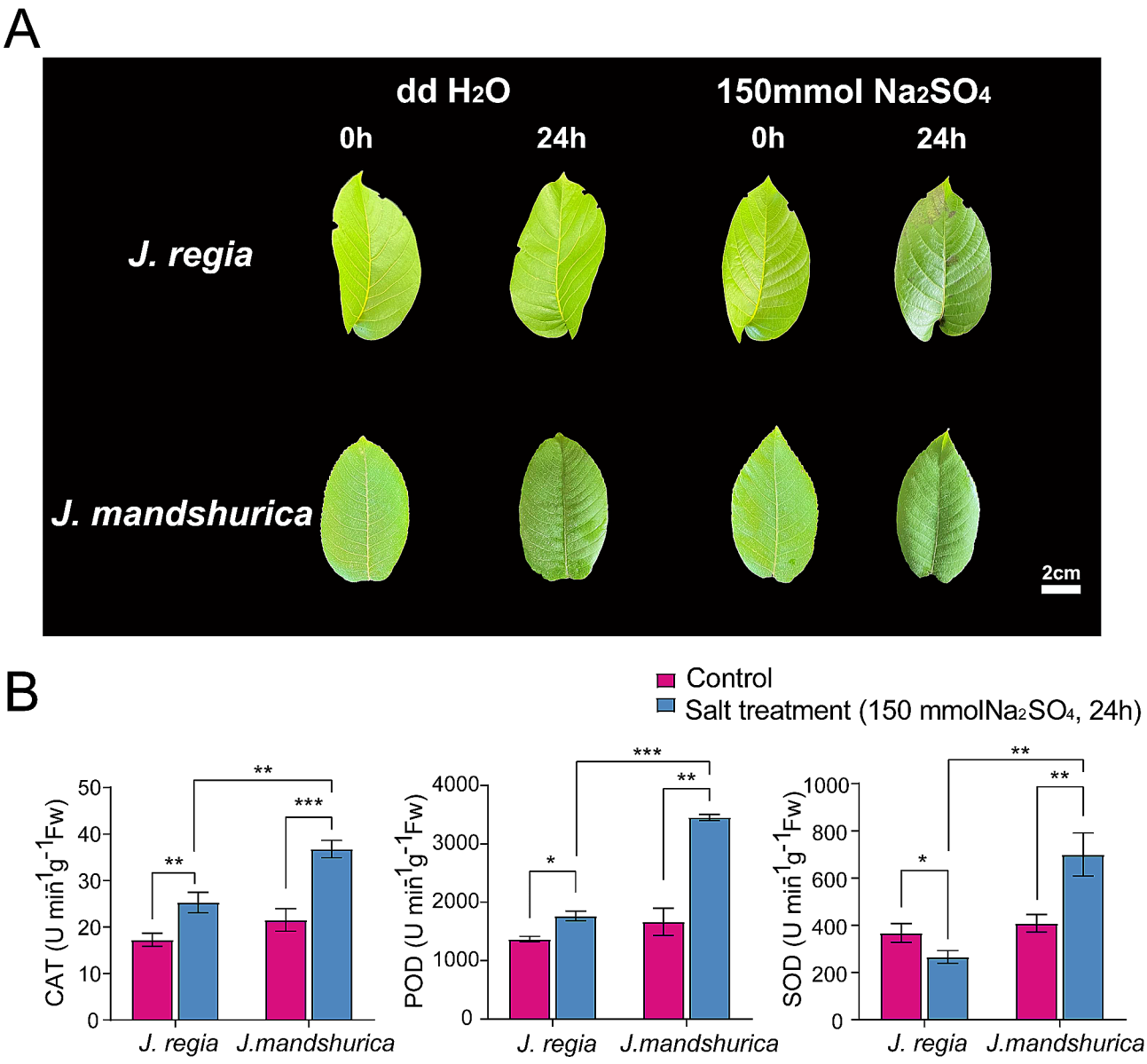

Fig. 8 Characteristics of two *Juglans* species leaves under salt stress. (A) Phenotypes of *J. regia* and *J. mandshurica* leaves under salt stress, where ddH_2O_2 indicated control group, 150 mmol/L Na_2SO_4 indicated salt-treated group; (B) Antioxidant enzyme activities of *J. regia* and *J. mandshurica* leaves under salt stress. CAT: Catalase, POD: Peroxidase, SOD: Superoxide dismutase. The red and blue columns indicated the control and salt-treated groups, respectively. The statistical significance using t-test. *=$p < 0.05$, **=$p < 0.01$, ***=$p < 0.001$

large variation in the number of *4CL* genes in different species [35–39] may be related to the gene duplication events experienced by different species during evolution. WGD is the most common gene duplication pattern, and WGD events may enhance the adaptability of species to their environment [77]. WGD events occurred in most of all identified *4CLs*, of these 47.22% in *Jr4CLs* and 64.52% in *Jm4CLs*. This suggests that WGD events play an important role in duplication events of the *4CL*gene family in two *Juglans* species (Table S3).

According to their phylogenetic relationships, all identified *4CL* genes can be divided into three subgroups (Clade I-III), and *4CLs* of three woody plants were clustered, indicating that the three woody plants are more closely related to each other (Fig. 1). Most *Jr4CLs* and *Jm4CLs* were distributed in the plasma membrane and chloroplast (Table 1). Similarly, *Gh4CLs* were widely distributed [40], but *Cit4CLs* were only localized in the cytoplasm [78], which may be related to the different locations where the *4CLs* function. *Jr4CLs* and *Jm4CLs* were unevenly distributed on the chromosome, but some *4CLs* had high similarity and collinearity, such as Jr4CL1 and Jr4CL2, Jr4CL19 and Jr4CL20, Jm4CL11 and Jm4CL12 (Figs. 2 and 4). Collinearity results showed

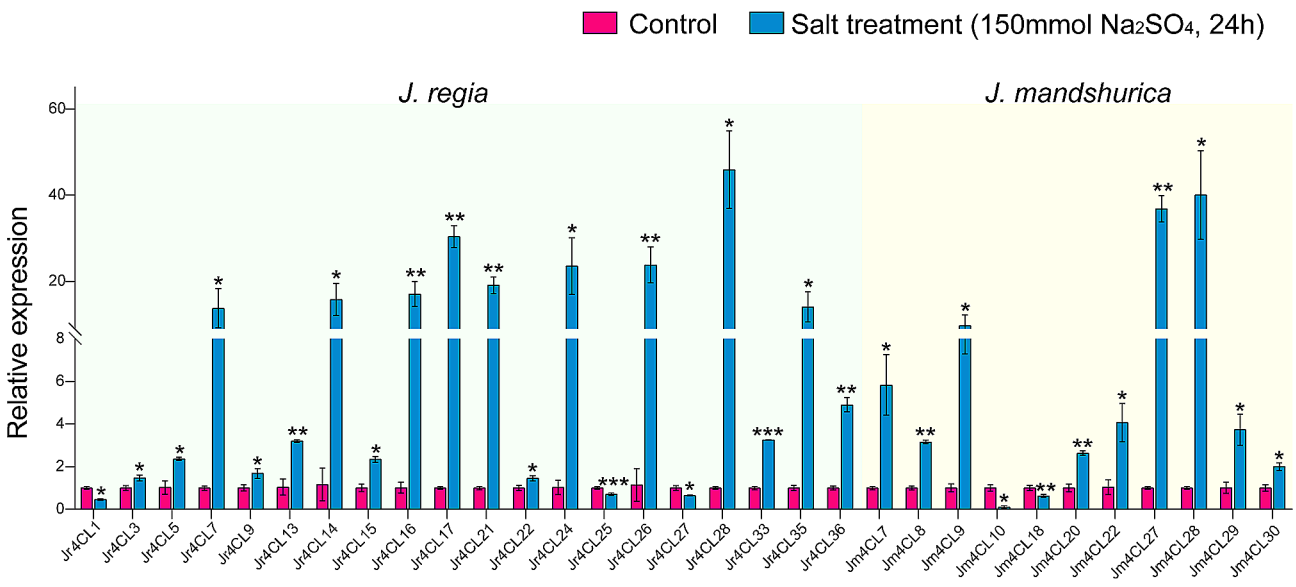

Fig. 9 The qRT-PCR experiments of *4CLs* in *J. regia* and *J. mandshurica* leaves under salt stress. The red and blue columns represent the control and salt-treated groups, respectively. The control group was treated with ddH_2O_2 and the salt-treated group was treated with 150 mmol/L Na_2SO_4 in two *Juglans* species leaves. The green and yellow backgrounds represent *J. regia* and *J. mandshurica*, respectively. The statistical significance using t-test. *=$p < 0.05$, **=$p < 0.01$, ***=$p < 0.001$

that 6 *Jr4CLs* were not collinear with any of the other four species, suggesting that they may be unique genes in *J. regia* (Fig. 4; Fig. S2; Table S5-8). Whereas *Jm4CLs* were collinear with the other four selected species, indicating that *Jm4CLs* may be relatively conserved over the course of evolution. All *Jr4CLs* and *Jm4CLs* homologous gene pairs Ka/Ks were smaller than 1, which shows that all gene pairs underwent purifying selection during evolution (Table S5). These results suggesting that the functions of *4CLs* may have been conserved over the course of evolution. The structural domains of 4CL proteins were highly conserved in two *Juglans* species (Fig. 3B), while all identified 4CL proteins contained conserved Box I and Box II sequences in multiple sequence analysis (Fig. S2), which is consistent with the results of previous studies [31]. However, the gene structures of *4CLs* showed large differences in two *Juglans* species, with exon numbers ranging from 1 to 23 (Fig. 3C; Table S4). The same phenomenons were observed in most species of *4CL* gene families. For example, the number of exons in *Md4CL* [39] and *Eu4CL* gene families [37] also ranged from 1 to 23. An increment in the number of introns in genes may be better than harm to the plant. As non-coding regions, introns could protect genes from mutations and thus better preserve gene function [79, 80]. *Cis*-acting elements are involved in the regulation of gene expression, and a variety of *cis*-acting elements in gene promoters may be associated with different gene functions [64]. A significant population of identified *4CLs* contained *cis*-acting elements in response to hormones such as MeJA, ABA, and abiotic stresses such as anaerobic induction (Fig. 5), indicating that *Jr4CLs* and *Jm4CLs* may be involved in the regulations of MeJA, ABA and anaerobic induction in plants.

Protein-protein interaction prediction results showed that most 4CL proteins had interactions with CHS proteins (Fig. 6A and B). 4CL catalyzes the formation of 4-coumaroyl CoA, which subsequently generates chalcone by chalcone synthase (CHS), the first important enzyme in flavonoid pathway [81]. The phenylpropane metabolic synthesis pathway is active in a wide range of plants when exposed to environmental stresses, providing precursors for the synthesis of flavonoids, lignins, phenolic acids, etc. The research showed that CHS activity was enhanced after exposure to exogenous pathogenic microorganisms [82], suggesting that 4CL may interact with CHS and play a role in defense against plant pathogens. 4CL proteins and CHS proteins co-regulated the production of flavonoid metabolites that regulate seed germination [83] and improve plant tolerance and resistance to adversity stresses [84]. MicroRNAs have become a research hotspot because they regulate gene transcription and thus affect plant productivity in many processes such as growth, development, and environmental stresses [85]. MiRNAs adaptive responses can increase plant survival in environments such as drought, salinity, and pathogens [86]. We found that 35 *Jr4CLs* and 30 *Jm4CLs* were predicted to be the target genes of 260 and 218 *Arabidopsis* miRNAs, respectively (Fig. 6C; Table S10). Among them, cleavage was the main mode of miRNA regulation of *4CL* gene expression.

The expression of *4CLs* in plants is specific in their different tissues/organs, with *At4CL1* and *At4CL2* being the most strongly expressed in seedling roots, while *At4CL3* had high expression levels in flowers [75]. In pomegranate, *Pg4CL1, Pg4CL4, Pg4CL5, Pg4CL6*, and *Pg4CL11* were highly expressed in roots, leaves, flowers, and pericarps. While *Pg4CL7* was highly expressed only in leaf and exocarp [38]. In cotton, different *4CLs* also showed different expression patterns in different organs [40]. *Gh4CL2, Gh4CL15, Gh4CL17, Gh4CL19, Gh4CL23*, and *Gh4CL33* were highly expressed in stems, and *Gh4CL8, Gh4CL30*, and *Gh4CL34* were highly expressed in roots. *Euc4CL17, Euc4CL23, Euc4CL29*, and *Euc4CL31* were highly expressed only in male and female flowers of *Eucommia ulmoides* [37]. Similar phenomena were observed in two *Juglans* species. All *Jr4CLs* were expressed in four organs (Fig. 7A). Nevertheless, only 21 *Jm4CLs were* expressed in selected organs, and the other 10 *Jm4CLs* may function in other tissues or other developmental stages (Fig. 7B). *Juglans* species were heterodichogamous [87, 88]. Interestingly, six *Jr4CL* genes (Jr4CL12, Jr4CL19, Jr4CL22, Jr4CL27, Jr4CL31, and Jr4CL35) and seven *Jm4CL* genes (Jm4CL7, Jm4CL9, Jm4CL21, Jm4CL22, Jm4CL26, Jm4CL27, and Jm4CL29) were highly expressed only in male or female flowers (Fig. 7), suggesting that those *4CL* genes may be related to the heterodichogamous in walnut species, but the regulatory mechanisms involved remain to be further investigated.

Walnut anthracnose, caused by *Colletotrichum gloeosporioides*, is one of the most serious walnut diseases, causing early defoliation of walnut and resulting in reduced fruit production [89]. Chemical control has been the main measure for controlling walnut anthracnose, but it is limited by the pathogen's resistance and the impact on environment, chemical control measures are limited [90–92]. Therefore, the breeding of disease-resistant varieties is of extreme interest. We explored the expression pattern of *Jr4CLs* in F26 and F423 (Fig. 7C; Table S13). We found that the expression level of *Jr4CLs* were generally increased in F26 over F423. Most *Jr4CLs* after infection with F26 showed an increase and then a decrease in expression levels, which increased again after 72 h. Expression levels of most *Jr4CLs* after infection with F423 increased over time at 48 h, and peaked at 48 h, followed by a gradual decrease in expression levels (Fig. S6). These results all suggest that *Jr4CLs* may function in *J. regia* resistance to anthracnose. In previous studies, *J. mandshurica* has better disease resistance than *J. regia*, and *J. mandshurica* is often used as a rootstock for *J. regia* to enhance its disease resistance [11]. Therefore, it is presumable that *Jm4CLs* may have better disease resistance performance compared to *Jr4CLs*, but further studies are needed.

Walnuts are less salt tolerant and more sensitive to salt [28]. After treatment of mature leaves of two *Juglans* species with 150 mmol/L Na_2SO_4 solution for 24 h, *J. mandshurica* had better performance than *J. regia* (Fig. 8A). Consistent with previous studies that concluded that *J. mandshurica* resistance is superior to *J. regia* [11, 14–16, 93, 94]. Under normal growth conditions, plants maintain a certain level of antioxidant enzyme system activity, scavenging the superoxide radicals that are constantly generated, so that the antioxidant enzyme activity and the superoxide radical content in the plant reach a certain equilibrium relationship. However, when plants are under various environmental stresses (such as drought, salt damage, extreme temperatures, pests, and diseases), it will lead to the production of a large number of reactive oxygen species (ROS), resulting in the impairment of the cellular structure and function, which will affect its growth and development, and even lead to its death [95, 96]. The accumulation of ROS induced by environmental stresses prompts plants to eliminate ROS by synthesizing antioxidant enzyme systems through signal transduction [96, 97]. SOD in the antioxidant system is the first line of defense to protect plant cells from oxygen free radicals and scavenge ROS [98]. POD is an endogenous scavenger of ROS in plants under environmental stresses and coordinates with SOD and CAT to scavenge excess free radicals in plants, maintain free radicals in plants at a normal level and enhance plant resilience [99]. When plants are in an adverse environment, antioxidant enzymes maintain high activity, which can keep free radicals and reactive oxygen species at relatively low levels and mitigate the effects on plant cells [100] Therefore, the level of antioxidant enzyme activity can reflect the strength of plant stress tolerance to a certain extent. After salt treatment, the determination of POD, SOD, and CAT activities in treated and control groups of *J. regia* and *J. mandshurica* leaves, to some extent could reflect the degree of response of antioxidant enzyme systems to salt stress in the two *Juglans* species subjected to salt stress. In the present study, the activities of all three antioxidant enzymes in *J. mandshurica* leaves increased significantly after salt treatment (Fig. 8B). However, the activities of CAT and POD in *J. regia* leaves increased significantly after salt treatment, whereas the activity of SOD decreased significantly. SOD is an important enzyme in the plant antioxidant system for scavenging free radicals, which breaks down superoxide anions and defends against cell membrane damage caused by ROS [101, 102], whereas J. regia leaves showed a decrease in SOD activity under salt stress, suggesting that damage caused by excess ROS could no longer be eliminated under this treatment condition (150 mmol/L Na_2SO_4, 24 h), this treatment condition may have reached the limit of *J. regia* leaves to resist salt damage. A similar phenomenon was observed in

Solidago canadensis [103], where CAT and SOD activities increased and POD activity decreased after salt treatment. The CAT and POD contents of strawberries [104] showed a tendency to increase and then decrease with the duration of salt treatment. POD and SOD activities of *Betula platyphylla* increased significantly after salt treatment [105]. In ginger, POD, SOD, and CAT showed significant elevation after salt stress [106]. After salt stress, SOD, CAT, and POD gradually accumulated in salt-tolerant lines of asparagus with the rise of time, while the activities of the three enzymes decreased in salt-susceptible lines at the late stage of treatment [107]. These are the response of plants to protect their normal growth in extreme environments by increasing their antioxidant enzyme system. The increase in SOD, POD, and CAT activities under appropriate salt concentration stress is a self-protection of plants against unfavorable environments, but when there is an excessive accumulation of reactive oxygen radicals, the antioxidant enzymes are reduced by lipid peroxidation of plant cell membranes which makes the cell membranes unstable, and therefore there is a tendency for a decrease in the SOD activity of *J. regia* after salt stress, which is similar to the results of the previous studies [103–107]. In general, the overall levels of increase in activity of antioxidant enzymes in *J. mandshurica* leaves were higher than that of *J. regia*, indicating that *J. mandshurica* leaves have better salt tolerance than *J. regia*, in agreement with the phenotypic result.

Previous studies have shown that 4CL genes are capable of responding to plant adversity stress and that different 4CL gene family members of the same species may differ in their response to stress [108]. Moreover, *4CL*s expression levels were significantly changed under salt stress (Fig. 9; Fig. S7), demonstrated that *4CL* genes of two *Juglans* species may function in response to salt stress. And *4CL* genes may expressed in response to salt stress through positive or negative regulation, with positive regulation predominating. The expression trend of *4CLs* under salt stress was consistent, but the degree of change was different, and the change in the degree of response contributed to better and faster response of *4CL* genes to salt stress. Individual genes showed a sharp increase in expression levels after salt stress, showing that these *4CLs* function in two *Juglans* species' resistance to salt stress. A similar phenomenon was observed in cotton, especially for *Gh4CL21, Gh4CL24, Gh4CL27*, and *Gh4CL31*, which showed a significant increase in expression levels after salt stress [40]. In *Eucommia ulmoides*, all 35 *Euc4CLs* responded to salt stress, and the expression levels of most of the *Euc4CLs* increased significantly after salt treatment, especially *Euc4CL9, Euc4CL17* and *Euc4CL27* [37]. In Mulberry, all four *Ma4CLs* responded to salt stress. All *Ma4CL1-3* showed an overall up-regulation under salt stress. While *Ma4CL4* showed a trend of up-regulation in stems and down-regulation in roots after salt stress [109]. 4CLs were regulated with salt resistance transcription factors such as MYB [110], bHLH [111], NAC [112], bZIP [113], and AP2/ERF [114], especially MYB (Fig. S8), suggesting that these 4CL genes are key candidates for salt tolerance. However, the functions of these *Jr4CLs* and *Jm4CLs* need to be further studied and verified. Walnuts are widely cultivated in Xinjiang Province in China, and central Asia, mostly growing in mountainous and saline areas [1, 5]. *J. mandshurica* has better salt tolerance than *J. regia*, so *J. mandshurica* can be used as a rootstock for walnuts to improve the salt tolerance of walnut fruit trees against the external environment and increase the fruiting rate of walnuts.

Conclusion

In this study, we comprehensively genome-wide identified the *4CL* gene family members in both *Juglans* species. Phylogenetic analysis showed that the *4CL* genes were divided into three branches. Collinearity analysis showed that the *4CL* genes were relatively conserved during evolution, but the gene structures varied widely, which was similar to the results in other plants. Gene expression analysis showed that both *Jr4CLs* and *Jm4CLs* play key roles in the resistance of two *Juglans* species to resistance stresses. Under salt stress treatment, both phenotypic results and antioxidant enzyme activity analyses suggest that *J. mandshurica* had better performance than *J. regia*. Therefore, *J. mandshurica* can be used as a rootstock for *J. regia* to resist diseases and salt damage. Laying the theoretical foundation for walnut germplasm resource enhancement for salt stress. It is also necessary to further explore the unique roles of *Jr4CLs* and *Jm4CLs* in other stresses (e.g., plant nematode disease, extreme temperatures, drought, flooding, etc.).

Abbreviations

4CL	4-Coumarate:CoA ligase
AMP	Adenosine monophosphate
POD	Peroxidase
SOD	Superoxide Dismutase
CAT	Catalase
WGD	Whole genome duplication
TD	Tandem duplication
PD	Proximal duplication
DSD	Dispersed duplication
MeJA	Methyl Jasmonate
ABA	Abscisic acid
CHS	Chalcone synthase

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-024-04899-8.

Supplementary Material 1

Supplementary Material 2

Acknowledgements
Not applicable.

Author contributions
PZ designed and conceptualized the project. JM, DZ, and HL collected samples. JM and HL processed samples. JM, DZ, XZ, HY, LM, MD, and FG performed the data analysis. JM and NZ performed the experiments. JM and DZ wrote the manuscript. HZ and PZ revised the manuscript accordingly. All authors have read and approved the final manuscript.

Funding
This work was supported by the National Natural Science Foundation of China (32370386, 32070372, and 32200295), Science Foundation for Distinguished Young Scholars of Shaanxi Province (2023-JC-JQ-22), Shaanxi key research and development program (2024NC-YBXM-064), Basic Research Project of Shaanxi Academy of Fundamental Science (22JHZ005), China Postdoctoral Science Foundation (2022MD723843 and 2023MD734225), Natural Science Foundation of Shaanxi Province of China (2019JM-008). Science and Technology Program of Shaanxi Academy of Science (2023 K-49, 2023 K-26, and 2019 K-06), Shaanxi Forestry Science and Technology Innovation Key Project (SXLK2023-02-20), Qinling Hundred Talents Project of Shaanxi Academy of Science (Y23Z619F17).

Data availability
The raw data were downloaded from the NCBI Sequence Read Archive (SRA) database under accession number (GSE147083).

Declarations

Ethics approval and consent to participate
This study has been approved by the Chinese government and carried out with the laws of the People's Republic of China. All participants had a license approval letter from the College of Life Sciences, Northwest University. All participants obtained permission to collect *J. regia* and *J. mandshurica* samples from Qinling National Forest Park in Shaanxi province. All methods were carried out according to relevant guidelines and regulations.

Consent for publication
Not applicable.

Conflicts of interest
The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Received: 23 January 2024 / Accepted: 11 March 2024
Published online: 23 March 2024

References

1. Zhao P, Zhou H, Potter D, et al. Population genetics, phylogenomics and hybrid speciation of *Juglans* in China determined from whole chloroplast genomes, transcriptomes, and genotyping-by-sequencing (GBS). Mol Phylogenet Evol. 2018;126:250–65. https://doi.org/10.1016/j.ympev.2018.04.014.
2. Mu XY, Tong L, Sun M, et al. Phylogeny and divergence time estimation of the walnut family (Juglandaceae) based on nuclear RAD-Seq and chloroplast genome data. Mol Phylogenet Evol. 2020;147:106802. https://doi.org/10.1016/j.ympev.2020.
3. Song YG, Fragnière Y, Meng HH, et al. Global biogeographic synthesis and priority conservation regions of the relict tree family Juglandaceae. J Biogeogr. 2020;47(3):643–57. https://doi.org/10.1111/jbi.13766.
4. Zhang Q, Ree RH, Salamin N, et al. Fossil-informed models reveal a boreotropical origin and divergent evolutionary trajectories in the walnut family (Juglandaceae). Syst Biol. 2021;71(1):242–58. https://doi.org/10.1093/sysbio/syab030.
5. Marrano A, Britton M, Zaini PA, et al. High-quality chromosome-scale assembly of the walnut (*Juglans regia* L.) reference genome. GigaScience. 2020;9:giaa50. https://doi.org/10.1093/gigascience/giaa050.
6. Martínez-García PJ, Crepeau MW, Puiu D, et al. The walnut (*Juglans regia*) genome sequence reveals diversity in genes coding for the biosynthesis of nonstructural polyphenols. Plant J. 2016;87(5):507–32. https://doi.org/10.1111/tpj.13207.
7. Zhang BW, Xu LL, Li N, et al. Phylogenomics reveals an ancient hybrid origin of the persian walnut. Mol Biol Evol. 2019;36(11):2451–61. https://doi.org/10.1093/molbev/msz112.
8. Pei D, Lu XZ. Walnut germplasm resources in China, Chinese, Beijing, 2011.
9. Stevens KA, Woeste K, Chakraborty S et al. Genomic variation among and within six *Juglans* species. G3 (Bethesda). 2018; 8:2153–65. https://doi.org/10.1534/g3.118.200030.
10. Yang K, Dong Q, Wu J, et al. Genome-wide analysis of the R2R3-MYB transcription factor gene family expressed in *Juglans regia* under abiotic and biotic stresses. Ind Crops Prod. 2023;198:116709. https://doi.org/10.1016/j.indcrop.2023.
11. Yan F, Xi RM, She RX, et al. Improved *de novo* chromosome-level genome assembly of the vulnerable walnut tree *Juglans mandshurica* reveals gene family evolution and possible genome basis of resistance to lesion nematode. Mol Ecol. 2021;21:2063–76. https://doi.org/10.1111/1755-0998.13394.
12. Bai W, Liao W, Zhang D. Nuclear and chloroplast DNA phylogeography reveal two refuge areas with asymmetrical gene flow in a temperate walnut tree from East Asia. New Phytol. 2010;188:892–901. https://doi.org/10.1111/j.14698137.2010.03407.x.
13. Bai W, Wang W, Zhang D. Contrasts between the phylogeographic patterns of chloroplast and nuclear DNA highlight a role for pollen mediated gene flow in preventing population divergence in an east Asian temperate tree. Mol Phylogenet Evol. 2014;81:37–48. https://doi.org/10.1016/j.ympev.2014.08.024.
14. Hu Z, Zhang T, Gao XX. De novo assembly and characterization of the leaf, bud, and fruit transcriptome from the vulnerable tree *Juglans mandshurica* for the development of 20 new microsatellite markers using Illumina sequencing. Mol Genet Genom. 2016;291:849–62. https://doi.org/10.1007/s00438-015-1147-y.
15. Trouern-Trend A, Falk T, Zaman S. Comparative genomics of six Juglans species reveals disease-associated gene family contractions. Plant J. 2020;102(2):410–23. https://doi.org/10.1111/tpj.14630.
16. Ji LI, Zhang Y, Yang Y. Long-term effects of mixed planting on arbuscular mycorrhizal fungal communities in the roots and soils of *Juglans mandshurica* plantations. BMC Microbiol. 2020;20(1):304. https://doi.org/10.1186/s12866-020-01987-1.
17. Arab MM, Marrano A, Abdollahi-Arpanahi R. Combining phenotype, genotype, and environment to uncover genetic components underlying water use efficiency in Persian Walnut. J Exp Bot. 2020;71(3):1107–27. https://doi.org/10.1093/jxb/erz467.
18. Zhou H, Ma J, Liu H, et al. Genome-wide identification of the CBF gene family and ICE transcription factors in walnuts and expression profiles under cold conditions. Int J Mol Sci. 2024;25:25. https://doi.org/10.3390/ijms25010025.
19. Yuan X, Huang S, Ma H, et al. Differential responses of walnut cultivars to cold storage and their correlation with postharvest physiological parameters. Hortic Environ Biotechnol. 2019;60:345–56. https://doi.org/10.1007/s13580-019-00126-8.
20. Wang B, Zhang J, Pei D, et al. Combined effects of water stress and salinity on growth, physiological, and biochemical traits in two walnut genotypes. Physiol Plant. 2021;172(1):176–87. https://doi.org/10.1111/ppl.13316.
21. Yang G, Zhang W, Liu Z, et al. Both JrWRKY2 and JrWRKY7 of *Juglans regia* mediate responses to abiotic stresses and abscisic acid through formation of homodimers and interaction. Plant Biol. 2017;19:268–78. https://doi.org/10.111 1/plb.12524.
22. Ji X, Tang J, Fan W, et al. Phenotypic differences and Physiological Responses of Salt Resistance of Walnut with four Rootstock types. Plants. 2022;11:1557. https://doi.org/10.3390/plants11121557.
23. Karimi S, Karimi H, Mokhtassi-Bidgoli A, et al. Inducing drought tolerance in greenhouse grown *Juglans regia* by imposing controlled salt stress: the role of osmotic adjustment. Sci Hortic. 2018;239:181–92. https://doi.org/10.1016/j.scienta.2018.05.029.
24. Peteidis A, Therios I, Samouris G, et al. Salinity–induced changes in phenolic compounds in leaves and roots of four olive cultivars (*Olea europaea* L.) and their relationship to antioxidant activity. Environ Exp Bot. 2012;79:37–43. https://doi.org/10.1016/j.envexpbot.2012.01.007.
25. Tavallavi V, Karimi S, Espargham O. Boron enhances antioxidative defense in the leaves of salt-affected *Pistacia vera* seedlings. Horticult J. 2018;87(1):55–62. https://doi.org/10.2503/hortj.OKD-062.
26. Sperling O, Lazarovitch N, Schwartz A, et al. Effects of high salinity irrigation on growth, gas–exchange, and photoprotection in date palms (*Phoenix*

dactylifera L., Cv. Medjool). Environ Exp Bot. 2014;99:100–9. https://doi.org/10.1016/j.envexpbot.2013.10.014.
27. Lotfi N, Vahdati K, Kholdebarin B, et al. Germination, mineral composition, and ion uptake in walnut under salinity conditions. Hortsci. 2009;44(5):1352–7. https://doi.org/10.21273/HORTSCI.44.5.1352.
28. Karimi S, Karimi H, Vahdati K, et al. Antioxidative responses to short-term salinity stress induce drought tolerance in walnut. Sci Hortic. 2020;267:109322. https://doi.org/10.1016/j.scienta.2020.109322.
29. Dong NQ, Lin HX. Contribution of phenylpropanoid metabolism to plant development and plant-environment interactions. J Integr Plant Biol. 2021;63:180–209. https://doi.org/10.1111/jipb.13054.
30. Mansell RL, Babbel GR, Zenk MH. Multiple forms and specificity of cinnamyl alcohol dehydrogenase from cambial regions of higher plants. Phytochemistry. 1976;15:1849–53. https://doi.org/10.1016/s0031-9422(00)88829-9.
31. Schneider K, Hövel K, Witzel K, et al. The substrate specificity-determining amino acid code of 4-coumarate: CoA ligase. PNAS. 2003;100(14):8601–6. https://doi.org/10.1073/pnas.1430550100.
32. Stuible H, Büttner D, Ehlting J, et al. Mutational analysis of 4-coumarate: CoA ligase identifies functionally important amino acids and verifies its close relationship to other adenylate-forming enzymes. Febs Lett. 2000;467(1):117–22. https://doi.org/10.1016/s0014-5793(00)01133-9.
33. Stuible HP, Kombrink E. Identification of the substrate specificity-conferring amino acid residues of 4-Coumarate: coenzyme a ligase allows the rational design of mutant enzymes with new catalytic properties. J Biol Chem. 2001;276:26893–7. https://doi.org/10.1074/jbc.m100355200.
34. Lavhale SG, Kalunke RM, Giri AP. Structural, functional and evolutionary diversity of 4-coumarate-CoA ligase in plants. Planta. 2018;248(5):1063. https://doi.org/10.1007/s00425-018-2965-z.
35. De Azevedo Souza C, Barbazuk B, Ralph SG et al. Genome-wide analysis of a land plant-specific *acyl:coenzymeA synthetase (ACS)* gene family in *Arabidopsis*, poplar, rice and *Physcomitrella* New Phytol. 2008; 179:987–1003. https://doi.org/10.1111/j.1469-8137.2008.02534.x.
36. Cao Y, Fang Z, Li S, et al. Genome-wide identification and analyses of 4CL gene families in *Pyrus Bretschneideri* Rehd. Hereditas (Beijing). 2015;37(07):711–9. https://doi.org/10.16288/j.yczz.15-069.
37. Zhong J, Qing J, Wang Q, et al. Genome-wide identification and expression analyses of the 4-Coumarate: CoA ligase (*4CL*) Gene Family in *Eucommia ulmoides*. Forests. 2022;13:1253. https://doi.org/10.3390/f13081253.
38. Wang Y, Guo L, Zhao Y, et al. Systematic analysis and expression profiles of the 4-Coumarate: CoA ligase (4CL) Gene Family in Pomegranate (*Punica granatum* L). Int J Mol Sci. 2022;23:3509. https://doi.org/10.3390/ijms23073509.
39. Ma Z, Nan X, Li W, et al. Comprehensive genomic identification and expression analysis *4CL* gene family in apple. Gene. 2023;858:147197. https://doi.org/10.1016/j.gene.2023.147197.
40. Sun SC, Xiong XP, Zhang XL, et al. Characterization of the *Gh4CL* gene family reveals a role of *Gh4CL7* in drought tolerance. BMC Plant Biol. 2020;20(1):125. https://doi.org/10.1186/s12870-020-2329-2.
41. Chen X, Su W, Zhang H, et al. *Fraxinus mandshurica* 4-coumarate-CoA ligase 2 enhances drought and osmotic stress tolerance of tobacco by increasing coniferyl alcohol content. Plant Physiol Bioch. 2020;155:697–708. https://doi.org/10.1016/j.plaphy.2020.08.031.
42. Nie T, Sun X, Wang S, et al. Genome-wide identification and expression analysis of the 4-Coumarate: coA ligase gene family in Solanum tuberosum. Int J Mol Sci. 2023;24(2):1642. https://doi.org/10.3390/ijms24021642.
43. Chen X, Wang H, Li X, et al. Molecular cloning and functional analysis of 4-Coumarate: CoA ligase 4 (4CL-like 1) from *Fraxinus mandshurica* and its role in abiotic stress tolerance and cell wall synthesis. BMC Plant Biol. 2019;19:231. https://doi.org/10.1186/s12870-019-1812-0.
44. Fan R, Hu L, Wu B, et al. Cloning and expression analysis of 4-coumarate: coenzyme a ligase gene (*Pn4CL*) in *Piper nigrum*. Chin J Trop Crops. 2020;41(4):737–44. https://doi.org/10.3969/j.issn.1000-2561.2020.04.015.
45. Uhlmann AJ, Ebel J. Molecular cloning and expression of 4-coumarate: coenzyme a ligase, an enzyme involved in the resistance response of soybean (*Glycine max* L.) against pathogen attack. Plant Physiol. 1993;102(4):1147–56. https://doi.org/10.1104/pp.102.4.1147.
46. Jung JH, Kannan B, Dermawan H, et al. Precision breeding for RNAi suppression of a major 4-coumarate: coenzyme a ligase gene improves cell wall saccharification from field grown sugarcane. Plant Mol Biol. 2016;92:505–17. https://doi.org/10.1007/s11103-016-0527-y.
47. Wang J, Chitsaz F, Derbyshire MK, et al. The conserved domain database in 2023. NAR. 2023;51(1):D384–8. https://doi.org/10.1093/nar/gkac1096.
48. Jaina M, Sara C, Williams L, et al. Pfam: the protein families database in 2021. NAR. 2021;49(1):D412–9. https://doi.org/10.1093/nar/gkaa913.
49. Ivica L, Supriva K, Peer B. SMART: recent updates, new developments and status in 2020. NAR. 2021;49(1):D458–60. https://doi.org/10.1093/nar/gkaa937.
50. Chen C, Chen H, Zhang Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020;13(8):1194–201. https://doi.org/10.1016/j.molp.2020.06.009.
51. Wang Y, Tang H, Debarry JD, et al. *MCScanX*: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. NAR. 2012;40(7):e49. https://doi.org/10.1093/nar/gkr1293.
52. Wang D, Zhang Y, Zhang Z, et al. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genomics Proteom Bioinf. 2010;8(1):77–80. https://doi.org/10.1016/S1672-0229(10)60008-3.
53. Gasteiger E, Hoogland C, Gattiker A, et al. Protein Identification and Analysis Tools on the Expasy Serverm. In: Walker JM, editor. The Proteomics protocols Handbook. Humana; 2015. pp. 571–607.
54. Lescot M, Déhais P, Thijs G, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. NAR. 2002;30(1):325–7. https://doi.org/10.1093/nar/30.1.325.
55. Cantalapiedra CP, Hernández-Plaza A, Letunic I, et al. Eggnog-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol. 2021;38(12):5825–9. https://doi.org/10.1093/molbev/msab293.
56. Tamura K, Stecher G, Kumar S. Mol Biol Evol. 2021;38(7):3022–7. https://doi.org/10.1093/molbev/msab120. MEGA11: Molecular Evolutionary Genetics Analysis Version 11.
57. Larkin MA, Blackshields G, Brown NP, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23(21):2947–8. https://doi.org/10.1093/bioinformatics/btm404.
58. Ivica L, Peer B. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. NAR. 2021;49(W1):W293–296. https://doi.org/10.1093/nar/gkab301.
59. Hu B, Jin J, Guo AY, et al. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2015;31(8):1296–7. https://doi.org/10.1093/bioinformatics/btu817.
60. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504. https://doi.org/10.1101/gr.1239303.
61. Dai X, Zhuang Z, Zhao PX. psRNATarget: a plant small RNA target analysis server (2017 release). NAR. 2018;46(W1):W49–54. https://doi.org/10.1093/nar/gky316.
62. Ma J, Zuo D, Ye H, et al. Genome-wide identification, characterization, and expression pattern of the late embryogenesis abundant (LEA) gene family in *Juglans regia* and its wild relatives. J Mandshurica BMC Plant Biol. 2023;23:80. https://doi.org/10.1186/s12870-023-04096-z.
63. Xu C, Li Z, Wang J. Linking heat and adaptive responses across temporal proteo-transcriptome and physiological traits of *Solidago canadensis*. Environ Exp Bot. 2020;175:104035. https://doi.org/10.1016/j.envexpbot.2020.104035.
64. Li M, Ma J, Liu H, et al. Identification and characterization of wall-associated kinase (WAK) and WAK-like (WAKL) gene family in *Juglans regia* and its wild related species *Juglans mandshurica*. Genes. 2022;13(1):134. https://doi.org/10.3390/genes13010134.
65. Liu H, Ye H, Wang J et al. Genome-wide identification and characterization of *YABBY* gene family in *Juglans regia* and *Juglans mandshurica*. Agronomy. 2022; 12:1914. https://doi.org/10.3390/agronomy12081914.
66. Kim D, Langmead B, Salzberg SL. HISAT: A fast spliced aligner with low memory requirements. Nat Methods. 2015;12:357–60. https://doi.org/10.1038/nmeth.3317.
67. Trapnell C, Williams BA, Pertea G, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28:511–5. https://doi.org/10.1038/nbt.1621.
68. Liao Y, Smyth GK, Shi W, featureCounts. An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30:923–30. https://doi.org/10.1093/bioinformatics/btt656.
69. Feng S, Feng H, Liu X, et al. Genome-wide identification and characterization of long non-coding RNAs conferring resistance to *Colletotrichum gloeosporioides* in walnut (*Juglans regia*). BMC Genom. 2021;22:15. https://doi.org/10.1186/s12864-020-07310-6.
70. Yang C, Shen S, Zhou S, et al. Rice metabolic regulatory network spanning the entire life cycle. Mol Plant. 2022;15(2):258–75. https://doi.org/10.1016/j.molp.2021.10.005.

71. Li Y, Luo X, Wu C, et al. Comparative transcriptome analysis of genes involved in anthocyanin biosynthesis in red and green walnut (*Juglans regia* L). Molecules. 2018;23:25. https://doi.org/10.3390/molecules23010025.
72. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(– Delta Delta C(T)) method. Methods. 2001;25(4):402–8. https://doi.org/10.1006/meth.2001.1262.
73. Qiao X, Li Q, Yin H, et al. Gene duplication and evolution in recurring polyploidization-diploidization cycles in plants. Genome Biol Evol. 2019;20:38. https://doi.org/10.1186/s13059-019-1650-2.
74. Zhou H, Yan F, Hao F, et al. Pan-genome and transcriptome analyses provide insights into genomic variation and differential gene expression profiles related to disease resistance and fatty acid biosynthesis in eastern black walnut (*Juglans nigra*). Hortic Res. 2023;10(3):uhad015. https://doi.org/10.1093/hr/uhad015.
75. Lee D, Ellard M, Wanner LA, et al. The *Arabidopsis thaliana* 4-coumarate: CoA ligase (*4CL*) gene: stress and developmentally regulated expression and nucleotide sequence of its cDNA. Plant Mol Biol. 1995;28:871–84. https://doi.org/10.1007/BF00042072.
76. Rao GD, Pan X, Xu F, et al. Divergent and overlapping function of five 4-Coumarate/Coenzyme A ligases from *Populus tomentosa*. Plant Mol Biol Rep. 2015;33:841–54. https://doi.org/10.1007/s11105-014-0803-4.
77. Jiao Y. Double the genome, double the fun: genome duplications in angiosperms. Mol Plant. 2018;11(3):357–8. https://doi.org/10.1016/j.molp.2018.02.009.
78. Shen W, Wang Z, Xue Y, et al. Characterization of4-coumarate: CoAligase (4CL) gene family in Citrus. Acta Horticulturae Sinica. 2019;46(6):1068–78. https://doi.org/10.16420/j.issn.0513-353x.2018-0993.
79. Jo BS, Choi SS. Introns: the functional benefits of introns in genomes. Genomics Inf. 2015;13:112–8. https://doi.org/10.5808/GI.2015.13.4.112.
80. Mukherjee D, Saha D, Acharya D, et al. The role of introns in the conservation of the metabolic genes of *Arabidopsis thaliana*. Genomics. 2018;110:310–7. https://doi.org/10.1016/j.ygeno.2017. 12.003.
81. Sommer H, Saedler H. Structure of the chalcone synthase gene of *Antirrhinum majus*. Molec Gen Genet. 1986;202:429–34. https://doi.org/10.1007/BF00333273.
82. Wanner LA, Li GQ. The phenylalanine ammonialyase gene family in *Arabidoposis Thaliana*. Plant Mol Biol. 1995;27:328–35. https://doi.org/10.1007/bf00020187.
83. Tan H, Man C, Xie Y, et al. A crucial role of GA-regulated flavonol biosynthesis in root growth of Arabidopsis. Mol Plant. 2019;12:521–37. https://doi.org/10.1016/j.molp.2018.12.021.
84. Agati G, Cerovic Z, Pinelli P, et al. Light-induced accumulation of Ortho-dihydroxylated flavonoids as non-destructively monitored by chlorophyll fluorescence excitation techniques. Environ Exp Bot. 2011;73:3–9. https://doi.org/10.1016/j.envexpbot.2010.10.002.
85. Cheng C, Liu F, Sun X, et al. Genome-wide identification of *FAD* gene family and their contributions to the temperature stresses and mutualistic and parasitic fungi colonization responses in banana. Int J Biol Macromol. 2022;204:661–76. https://doi.org/10.1016/j.ijbiomac.2022.02.024.
86. Samynathan R, Venkidasamv B, Shanmugan A, et al. Functional role of microRNA in the regulation of biotic and abiotic stress in agronomic plants. Front Genet. 2023;14:1272446. https://doi.org/10.3389/fgene.2023.1272446.
87. Kimura M, Serwa K, Suyama Y, et al. Flowering system of heterodichogamous *Juglans ailanthifolia*. Plant Species Biol. 2003;18:75–84. https://doi.org/10.1111/j.1442-1984.2003.00088.x.
88. Zhang L, Guo C, Lu X, et al. Flower Development of Heterodichogamous *Juglans mandshurica* (Juglandaceae). Front Plant Sci. 2021;12:541163. https://doi.org/10.3389/fpls.2021.541163.
89. Zhu YF, Yin YF, Qu WW, et al. Morphological and molecular identification of *Colletotrichum gloeosporioides* causing walnut anthracnose in Shandong Province. Acta Hortic. 2014;1050:353–9. https://doi.org/10.17660/ActaHortic.2014.1050.48.
90. Mackenzie SJ, Mertely JC, Peres NA. Curative and protectant activity of fungicides for control of crown rot of strawberry caused by *Colletotrichum gloeosporioides*. Plant Dis. 2009;93(8):815–20. https://doi.org/10.1094/pdis-93-8-0815.
91. Zhang HL, Wang YJ, Zhang CH, et al. Isolation, characterization and expression analysis of resistance gene candidates in pear (*Pyrus spp*). Sci Hortic. 2011;127:282–9. https://doi.org/10.1016/j.scienta.2010.10.016.
92. Seehalak W, Moonsom S, Metheenukul P, et al. Isolation of resistance gene analogs from grapevine resistant and susceptible to downy mildew and anthracnose. Sci Hortic. 2011;128:357–63. https://doi.org/10.1016/j.scienta.2011.01.003.
93. Zhou Z, Han M, Hou M, et al. Comparative study of the leaf transcriptomes and ionoms of *Juglans regia* and its wild relative species *Juglans cathayensis*. Acta Physiol Plant. 2017;39(10):224. https://doi.org/10.1007/s11738-017-2504-8.
94. Chen G, Pi XM, Yu CY. A new naphthalenone isolated from the green walnut husks of *Juglans mandshurica* maxim. Nat Prod Res. 2015;29(2):174–9. https://doi.org/10.1080/14786419.2014.971789.
95. Azooz MM, Youssef AM, Ahmad P. Evaluation of salicylic acid (SA) application on growth, osmotic solutes and antioxidant enzyme activities on broad bean seedlings grown under diluted seawater. Int J Plant Physiol Biochem. 2011;3:253–64. https://doi.org/10.5897/IJPPB11.052.
96. Ahmad P, Hakeem KR, Kumar A, et al. Salt-induced changes in photosynthetic activity and oxidative defense system of three cultivars of mustard (*Brassica juncea* L). Afr J Biotechnol. 2012;11(11):2694–703. https://doi.org/10.5897/AJB11.3203.
97. Meloni DA, Oliva MA, Martinez CA, et al. Photosynthesis and activity of superoxide dismutase, peroxidase and glutathione reductase in cotton under salt stress. Environ Exp Bot. 2003;49(1):69–76. https://doi.org/10.1016/S0098-8472(02)00058-8.
98. Gebauer J, El-Siddig K, Salih AA, et al. *Tamarindus indica* L. seedlings are moderately salt tolerant when exposed to NaCl-induced salinity. Sci Hortic. 2004;103(1):1–8. https://doi.org/10.1016/j.scienta.2004.04.022.
99. Guan B, Yu J, Lu Z, et al. Effects of water-salt stresses on seeding growth and activities of antioxidative enzyme of *Suaeda salsa* in coastal wetlands of the yellow river delta. Environ Sci. 2011;32(8):2422–9. https://doi.org/10.13227/j.hjkx.2011.08.003.
100. Bose J, Rodrigo-Moreno A, Shabala S. ROS homeostasis in halophytes in the context of salinity stress tolerance. J Exp Bot. 2014;65(5):1241–57. https://doi.org/10.1093/jxb/ert430.
101. Wu Z, Wang J, Yan D, et al. Exogenous spermidine improves salt tolerance of pecan-grafted seedlings via activating antioxidant system and inhibiting the enhancement of Na+/K+ ratio. Acta Physiol Plant. 2020;42:83. https://doi.org/10.1007/s11738-020-03066-4.
102. Klein A, Hüsselmann L, Keyster M, et al. Exogenous nitric oxide limits salt-induced oxidative damage in maize by altering superoxide dismutase activity. S Afr J Bot. 2018;115:44–9. https://doi.org/10.1016/j.sajb.2017.12.010.
103. Li Z, Xu C, Wang J. Integrated physiological, transcriptomic and proteomic analyses revealed molecular mechanism for salt resistance in *Solidago canadensis* L. Environ Exp Bot. 2020;179:104211. https://doi.org/10.1016/j.envexpbot.2020.104211.
104. Ghaderi N, Hatami M, Mozafari A, et al. Change in antioxidant enzymes activity and some morpho-physiological characteristics of strawberry under long-term salt stress. Physiol Mol Biol Plants. 2018;24:833–43. https://doi.org/10.1007/s 12298-018-0535-2.
105. Mijiti M, Zhang Y, Zhang C, et al. Physiological and molecular responses of *Betula platyphylla* Suk to salt stress. Trees. 2017;31:1653–65. https://doi.org/10.1007/s00468-017-1576-9.
106. Liu M, Lv Y, Cao B, et al. Physiological and molecular mechanism of ginger (*Zingiber officinale* Roscoe) seedling response to salt stress. Front. Plant Sci. 2023;14:10733434. https://doi.org/10.3389/fpls.2023.1073434.
107. Guo X, Ahmad N, Zhao S, et al. Effect of salt stress on growth and physiological properties of Asparagus seedlings. Plants. 2022;11:2836. https://doi.org/10.3390/plants11212836.
108. Lavhale SG, Kalunke RM, Giri AP. Structural, functional and evolutionary diversity of 4 coumarate CoA ligase in plants. Planta. 2018;248:1063–78. https://doi.org/10.1007/s00425-018-2965-z.
109. Wang CH, Yu J, Cai YX, et al. Characterization and functional analysis of 4-Coumarate: CoA ligase genes in mul-berry. PLoS ONE. 2016;11(5):e01558114. https://doi.org/10.1371/journal.pone.0155814.
110. Zhang P, Wang R, Yang X, et al. The R2R3-MYB transcription factor AtMYB49 modulates salt tolerance in *Arabidopsis* by modulating the cuticle formation and antioxidant defence. Plant Cell Environ. 2020;43:1925–43. https://doi.org/10.1111/pce.13784.
111. Liu D, Li YY, Zhou ZC, et al. Tobacco transcription factor bHLH123 improves salt tolerance by activating NADPH oxidase *NtRbohE* expression. Plant Physiol. 2021;186(3):1706–20. https://doi.org/10.1093/plphys/kiab176.
112. Zhang W, Zhi W, Qiao H, et al. H2O2-dependent oxidation of the transcription factor *GmNTL1* promotes salt tolerance in soybean. Plant Cell. 2023;250. https://doi.org/10.1093/plcell/koad250.

113. Yanagisawa S, Sheen J. Involvement of maize Dof zinc finger proteins in tissue-specific and light-regulated gene expression. Plant Cell. 1998;10:75–89. https://doi.org/10.1105/tpc.10.1.75.
114. Ohme-Takagi M, Shinshi H. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. Plant Cell. 1995;7(2):173–82. https://doi.org/10.1105/tpc.7.2.173.

Publisher's Note