BMC Plant Biology

RESEARCH

Open Access

Function diversification of *CONSTANS-like* genes in *Pyrus* and regulatory mechanisms in response to different light quality

Kefan Cai[1†], Xinyi Li[1†], Dongrui Liu[1†], Sihan Bao[1], Cong Shi[1], Siting Zhu[1], Kai Xu[1*], Xuepeng Sun[1*] and Xiaolong Li[1*]

Abstract

Pear (*Pyrus* L.) is a significant commercial fruit globally, with diverse species exhibiting variations in their flowering periods due to environmental factors. *CONSTANS-like* (*COL*) genes, known from previous studies in *Arabidopsis*, are key regulators of flowering time by sensing photoperiod. However, the evolutionary history and functions of *COL* genes in different pear species remain unclear. In this study, we identified a total of 79 *COL* genes in different pear species, including 12 *COL* genes in *Pyrus bretschneideri* 'DangshanSuli', 9 in *Pyrus ussuriensis* × *hybrid* 'Zhongai 1', 11 in *Pyrus communis* 'Bartlett', 13 in *Pyrus betulifolia*, 18 in *Pyrus pyrifolia* 'Cuiguan', 16 in *Pyrus pyrifolia* 'Nijisseiki'. Analysis of gene structure, phylogenetic tree, and multiple sequences provided valuable insights into the fundamental understanding of *COL* genes in pear. The impact of selection pressure on the *PbrCOLs* in Chinese white pear was assessed using *Ka/Ks*, revealing that the evolution rate of *PbrCOLs* was influenced by purification selection factors. The study also revealed different tissue-specific expression patterns of *PbrCOLs* under varying light quality. Real-time quantitative PCR revealed that under natural light conditions, the expression patterns of *PbrCOL2*, *PbrCOL3*, and *PbrCOL4* are similar to previous studies on *CONSTANS* gene in *Arabidopsis*, with increased expression levels during the day and decreased levels at night. However, *PbrCOL1*, *PbrCOL6*, and *PbrCOL9* exhibit different expression patterns, with decreased expression levels both during the day and at night. After red light treatment, high expression of *PbrCOL3* and *PbrCOL4* was observed at night, while the expression patterns of the other four genes did not show significant changes. Following blue light treatment, the expression peaks of *PbrCOL1* and *PbrCOL6* occurred during the night, showing opposite expression patterns compared to the study in *Arabidopsis*. The overexpression of *PbrCOL3* significantly increase the chlorophyll content in pear seedlings, and its expression significantly affected the expression of other key flowering-related genes. Also, overexpression of *PbrCOL3* resulted in a late-flowering phenotype in *Arabidopsis*. These findings indicate diverse responsive mechanisms and functions of *PbrCOL* genes on flowering time in pear. In conclusion, this study established a foundation for a deeper understanding of the specific roles of *PbrCOLs* in regulating the reproductive development of pear, particularly in the context of the photoperiodic flowering process.

Keywords Pyrus, Evolution, *CONSTANS-like*, Light quality, Chlorophyll content

[†]Kefan Cai, Xinyi Li and Dongrui Liu these authors contributed equally.

*Correspondence:
Kai Xu
xukai@zafu.edu.cn
Xuepeng Sun
xs57@zafu.edu.cn
Xiaolong Li
lixiaolong@zafu.edu.cn

Introduction

Pear are one of the most widely cultivated fruit species due to their rich nutrient content, including dietary fiber, carbohydrates, and vitamins. As a member of the Rosaceae family and the subfamily Maloideae, pears are evolutionarily related to apples as well as other Rosaceae species [1–6]. Numerous species within this family hold economic value and are distributed across diverse climatic regions worldwide. It is well known that fruit trees cultivated in different climate regions are diversified in flowering time due to their long-term evolution. In the contemporary pear industry, multiple environment factors greatly affect pear's vegetative and reproductive growth through a complex regulation mechanism, leading to variantions in the flowering period and resulting in the different trade times in the fruit market. Previous research has uncovered four major pathways that significantly control the flowering time: Photoperiodic pathway [7], Vernalization pathway [8–10], Autonomous pathway [11, 12], and Gibberellin (GA) pathway [13, 14]. Additionally, secondary regulatory pathways such as sugar pathway [15], ambient temperature pathway [16], aging pathway [17], and circadian clock pathway [18] have been identified. Photoperiodic pathway plays a significant role among these pathways, particularly in specific plants such as *Arabidopsis* and rice [19, 20]. It can be classified into three types based on diverse day-length requirements: long-day (LD) condition, short-day (SD) condition, and neutral condition.

The CONSTANS (CO) protein, a main transcriptional regulator in the photoperiodic pathway, plays a crucial role in controlling flowering by perceiving light signals, which was firstly been reported in *Arabidopsis* [21]. *CO*-like (*COL)* genes, homologous to *CO*, have been found to have diverse functions in different plants, and its function was characterized by a long day (LD)-specific late-flowering mutant phenotype and is regulated by the circadian clock component *GI (GIGANTEA)*. Thus, *COL* is a central integrator of both internal circadian clock and the external day-night cycles [22]. Previous study reported that *COL* genes can activate the transcription of *FLOWERING LOCUS T* (*FT*) and *SUPPRESSOR OF OVEREXPRESSION OF COL 1* (*SOC1*) only under long-day condition in *Arabidopsis*, leading to the acceleration of flowering. However, the *COL* homologous can also promote flowering under SD condition in other species. For example, in rice, the *COL* homolog *HEADING DATE 1* (*Hd1*) promoted flowering under SD condition but suppressed flowering under LD condition, suggesting the *COL* homologous gene perform different functions in different plants such as monocotyledon and dicotyledon [23].

In *Arabidopsis*, the overexpression of *AtCOL*3, *AtCOL*7, and *AtCOL*8 can delay flowering time, while in contrast, the overexpression of the *AtCOL*5 gene promotes flowering by enhancing the expression of *FT*, suggesting the functional deversification also appeared among the *COL* homologous [24, 25]. For example, the ectopic expression of *AtCOL* gene in potato can inhibit the potato tuber expansion, while silencing of potato *StCOL* gene can promote the potato tuber expansion. Moreover, *AtCOL*3 regulates root growth, *VviCOL1* plays a major role in bud dormancy, *CrCOL* regulates starch synthesis and cell division, and *AtCOL7* regulates branching. *Ghd2* confers drought sensitivity, *StCOL*1 is involved in tuberization, and *MaCOL*1 regulates fruit ripening [26–32]. Just alike to the functional differentiation, the remarkable different expression patterns of *COL* genes had been reported in various species.

Despite systematic research on the *COL* gene family in various plants, the role of *COL* genes in regulating pear flowering remains unclear. In this study, we utilized fourteen Rosaceae genomes to perform a genome-wide identification of the *COL* gene family members. We provide detailed molecular information about the *COL* gene family in different Rosaceae species, including sequence homology, intron distribution, motif composition, and evolutionary relationships. The expression of *COL* genes in different tissues and organs was obtained from the Pear Expression Database. Our results aim to enhance the understanding of the functional roles and evolution of *COL* genes in Rosaceae species, particularly in pears, and provide insights into the functional verification of candidate genes associated with flowering.

Materials and methods

Plant materials preparation

In this study, three plant materials were utilized: *Pyrus betulifolia*, *Nicotiana benthamiana*, and *Arabidopsis thaliana*. The seeds of *Pyrus betulifolia* were obtained from Nanjing Agricultural University, selecting only the plump, non-cracked seeds. These seeds were then disinfected with a 2% sodium hypochlorite solution and sown in a substrate composed of a 3:1 mixture of organic fertilizer and soil. The seeds were placed in a climate chamber maintained at a constant temperature of 22 °C, relative humidity of 80%, and a photoperiod of 16 h light and 8 h dark. *Arabidopsis thaliana* was sourced from Zhejiang A&F University, and its sowing and cultivation conditions were consistent with those of *Pyrus betulifolia*. *Nicotiana benthamiana*, also obtained from Zhejiang A&F University, was cultivated under conditions of a constant temperature of 25 °C, relative humidity of 80%, and a photoperiod of 16 h light and 8 h dark in a plant growth chamber.

Whole Genome Identification of COL genes in Rosaceae

The genome data of nine Rosaceae species (Pear, Apple, Strawberry, Rose, Apricot, Japanese Apricot, Peach, Raspberry, and Plum) were downloaded via the GDR database (https://www.rosaceae.org/). We then used two methods to identify the genes of *COL* family. Firstly, to survey the *COL* members in the nine Rosaceae species (including six pear accessions), we selected the typical B-box and CCT domains and downloaded the Hidden Markov Model (HMM) files of B-box and CCT domains in Pfam (http://pfam.xfam.org/), which was used to identify candidate *COL* genes in nine Rosaceae species by using HMMSearch software (https://github.com/madscientist01/hmmsearch). Secondly, the protein sequences of the *COL* family from *Arabidopsis* were subsequently downloaded through Plant Transcription Factor Database (PlantTFDB, https://planttfdb.gao-lab.org/). Using the *COL* protein sequences in *Arabidopsis* as a query, a BLAST search (E-value < $1E^{-5}$) was performed by using the full-length amino acid sequences of *COL*s and *COL* of *Arabidopsis*. As a result, the overlapped genes of two methods were used for further analyses. The genes were named according to their Latin scientific name and ordered by the E-value results of Blast and HMMSearch.

SMART (http://smart.emblheidelberg.de/) and InterProScan tool (http://www.ebi.ac.uk/Tools/pfa/iprscan/) are utilized to verify the completeness of the *COL* domain and to perform a functional analysis of the corresponding protein sequences. A total of 154 genes were identified as putative members of *COL* genes in nine Rosaceae species. In addition, we used the Expasy tool (https://web.expasy.org/protparam/) to predict protein molecular weights (Mw) and isoelectric points (pI) for the nine species.

Multiple sequence alignment and phylogenetic tree construction

Multiple sequence alignment of *COL* protein sequences of ten plant species (nine Rosaceae species and *Arabidopsis thaliana*) was performed by ClustalW [33] in the MEGA11 [34] software package. Further, the multiple sequence alignment result was used to constructe an unrooted phylogenetic tree using MEGA-X software [35] with the neighbor-joining method (NJ) and 1000 bootstraps were set, and then used the iTOL website (https://itol.embl.de/) to visualize the constructed phylogenetic tree.

Gene structure and conserved motifs analyses

The gene structure of the *COL* genes in Rosaceae species was visualized with the TBtools tool [36]. Phylogenetic analysis was performed using the *COL* homologous sequences of several plant species together. The COL protein sequences were obtained from the GDR database and NCBI. The main motifs of *COL* were characterized by the Multiple Em for Motif Elicitation (MEME, https://meme-suite.org/meme/) and the domain of *COL* were searched by the NCBI website. Then, the schematic diagrams of protein domain structure and conserved motif were illustrated with the TBtools software.

Analysis of the cis-acting elements

We extracted all promoter sequences of the COL genes using the "getfasta" function in TBtools to obtain 2000 bp sequences upstream of the transcriptional start sites (TSS, putative promoter regions). Subsequently, we utilized the PlantCARE website (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/) to predict the cis-regulatory elements present in the *COL* gene promoter regions.

Computation of *Ka*/*Ks* values

TBtools sooftware was used to perform the nucleotide sequence alignment of *COL*s. Non-synonymous and synonymous substitution rates (denoted as *Ka* and *Ks*, respectively) were implemented by *Ka*/*Ks* Calculator program.

Quantitative real-time PCR analysis (qRT-PCR)

In this experiment, we selected *Pyrus betulifolia* grown at Zhejiang A&F university. In late October, we selected branches with good growth conditions and full buds, trimmed them to a length of 10–12 cm with 2–3 bud points, and conducted hydroponic cultivation. The branches were placed in a growth chamber with a temperature of 28 °C and a light cycle of 16 h light/8 h darkness. Approximately 2–3 weeks later, the branches would bud, and after the leaves unfolded, branches with relatively consistent growth conditions were selected and divided into three groups, each group containing 5 branches. The group I was placed in a growth chamber with a temperature of 22 °C, a light cycle of 16 h light/8 h darkness, and natural light quality as the control. The group II was placed in a growth chamber with a temperature of 22 °C, a light cycle of 0 h light/24 h darkness, with additional installation of red light bulbs (wavelength 620–630 nm, power 10W) as red light treatment. The group III was placed in a growth chamber with a temperature of 22 °C, a light cycle of 0 h light/24 h darkness, with additional installation of blue light bulbs (wavelength 450–460 nm, power 10W) as blue light treatment. The time of placement in the growth chamber was considered as 0 h, and leaf samples were taken every 4 h, a total of 6 times. The leaf samples were immediately frozen in liquid nitrogen and stored in a −80 °C freezer, and then total RNA

was extracted using a modified cetyltrimethylammonium bromide (CTAB) method.

The full-length cDNAs were synthesized using the PrimeScript RT reagent Kit (TaKaRa, Otsu, Japan), and primers were designed using the NCBI website (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) (Table S2). Then, qRT-PCR was performed on the CFX96 Real-Time PCR System (Applied Biosystems, Foster City, CA, United States) using the SYBR Premix Ex Taq Kit (TaKaRa, Otsu, Japan) according to the manufacturer's protocol. The *PbrActin* served as an internal control. The relative expression of genes was calculated using the $2^{-\Delta\Delta CT}$ method. All the above samples were executed independently in triplicate.

Determination of chlorophyll content

The chlorophyll content of pear leaves, as determined by the SPAD-502, is expressed in SPAD units. These units represent a relative value of chlorophyll content calculated based on the degree of light absorption at specific wavelengths by the leaves. The chlorophyll content was measured in multiple leaves from the same plant, with three measurements taken for each leaf. Subsequently, the data from various parts of these three leaves were compiled to create a boxplot, illustrating the chlorophyll content of leaves from different treatment groups of pear seedlings.

Different vectors construction of PbrCOL3 gene

In this experiment, SnapGene (version 6.0.2) was utilized for primer design, selecting the p1300-35S-GFP-BS2 overexpression vector, which contains a GFP tag and restriction enzyme sites for XbaI and BamHI. The coding DNA sequence (CDS) of *PbrCOL3* (Accession number: LOC103953229) was constructed at the N-terminal end of the GFP tag. Additionally, the TRV2 vector was chosen for the VIGS (Virus-Induced Gene Silencing) experiment, where the SnabI restriction site was used to insert the specific domain (208 bp) of *PbrCOL3* into the vector.

RNA was extracted from the young leaves of pear trees and reverse transcribed into cDNA, followed by PCR amplification. The reaction mixture included 1 μL of 200 ng/μL cDNA, 1 μL of 100 μmol/L upstream and downstream primers, 12.5 μL of high-fidelity enzyme (Vazyme, catalog number: P510, 2 × Phanta Flash Master Mix), and 9.5 μL of DEPC-treated water. The reaction proceeded with the following program: 30 s at 98 °C for initial denaturation, followed by 30 cycles of 10 s at 98 °C, 5 s at 58 °C, and 6 s at 72 °C, concluding with a final extension at 72 °C for 1 min. The PCR products were analyzed using 1% agarose gel electrophoresis, followed by purification and recovery of the target gene, which was then ligated to the p1300-35S-GFP-BS2 vector and TRV2 vector. For the agarose gel electrophoresis images of specific gene fragment amplification and E. coli positive detection, please refer to Figure S6.

Genetic transformation experiments of pear seedling and Arabidopsis

Each clone and overexpression vector was propagated in *Escherichia coli DH5α* competent cells (Tsingke, catalog number: TSC-C14, DH5α Chemically Competent Cell), followed by transformation into *Agrobacterium tumefaciens GV3101*(Coolaber, catalog number: CC405, GV3101 Chemically Competent Cell). Positive clones were screened via PCR. Subsequently, *Agrobacterium* that contained the overexpression vector was then used to infect *Arabidopsis thaliana*, following the method of Zhang et al. [37], to obtain transgenic *Arabidopsis* overexpressing the *PbrCOL3* gene. The VIGS experiment was conducted according to the method reported by Xiao et al. [38].The transformed *Agrobacterium* was cultured in LB medium supplemented with rifampicin (25 mg/L) and kanamycin (50 mg/L) at 28 °C with shaking for 16–18 h, followed by centrifugation at 4000 rpm for 10 min. The pellet was resuspended in infiltration buffer [10 mM MES (pH 5.6), 10 mM $MgCl_2$, and 150 μM acetosyringone] and incubated for 1 h.

Pear seedlings were sown in sterilized substrate (composed of a 3:1 mixture of organic fertilizer and soil) and maintained in an artificial climate incubator at 25 °C under a photoperiod of 16 h light and 8 h dark for 35 days. *Agrobacterium* containing the overexpression and transient silencing vectors was injected into the upper four leaves of the pear plants. Chlorophyll content was measured 3 to 7 days post-infection (specific time indicated in the figure legend), followed by photographic documentation and sampling; the samples were preserved at −80 °C for subsequent qRT-PCR analysis.

Wild-type *Arabidopsis* seeds were sown in sterilized substrate (composed of a 3:1 mixture of organic fertilizer and soil) and cultivated under the same greenhouse conditions. Once the primary inflorescence reached 7–10 cm in length, it was removed to promote the development of lateral inflorescences. Inoculation was performed prior to the opening of the petals, with each flower immersed in the *Agrobacterium* suspension for 3–5 s, repeated three times. After a 24-h dark incubation, the plants were maintained under a photoperiod of 16 h light and 8 h dark until seed collection and subsequent re-sowing.

Data statistical analyses

The IBM SPSS Statistics 22 software was used to analysis the experiment data. All data were shown as mean ± standard deviations (SD). A one-way analysis of variance (ANOVA) was applied to assess the statistical

significance of results, with the Student's t-test to compare means at a $p < 0.05$. Graphics were performed with GraphPad Prism software (v.10.2.0).

Results

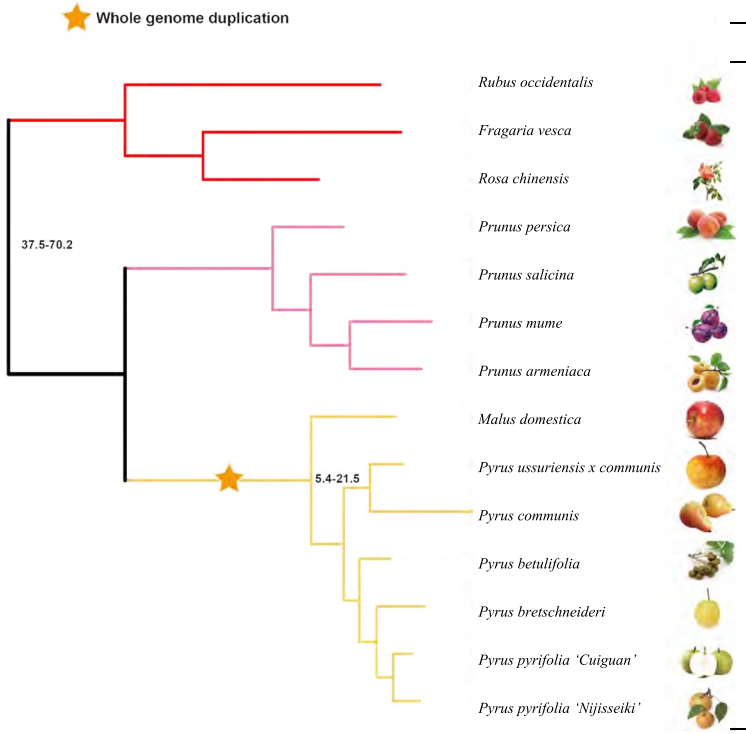
Identification of COL genes in nine Rosaceae species

To identify the *COL* members in the Rosaceae species, the BLAST and HMMSearch software were used to survey the potential genes in the genomes of different species. The typical B-box and CCT domain Hidden Markov Model (HMM) files downloaded from Pfam database were used for the genome wide search of *COL* genes. A total of 154 distinct genes was identified as putative members of nine Rosaceae species, including 9 genes from raspberry, 8 genes from rosa, 7 genes from strawberry, 9 genes from peach, 9 genes from apricot, 8 genes from plum, 9 genes from Japanese apricot, 16 genes from apple, 9 genes from 'Zhongai 1', 11 genes from 'Bartlett', 13 genes from *Pyrus betulifolia*, 17 genes from 'DangshanSuli', 18 genes from 'Cuiguan' pear, 16 genes from 'Nijisseiki' (Fig. 1). Lineage-specific Whole Genome Duplication (WGD) was found to have occurred in the ancestor of all of different pears and apples, which may have resulted in the larger number of *COL* genes in pears and apple.

Then, the ExPASy Proteomics Server website was used to predict the molecular weight (Mw) and isoelectric point (pI) of the COL protein sequences of Chinese white pear (Table S1). The length of the COL protein sequences ranged from 340 (Pbr036464.1, Pbr019957.1) to 859 (Pbr026954.1) amino acids. The protein mass ranged from 37.8 kD (Pbr036464.1) to 93.6 kD (Pbr026954.1), and protein pIs ranged from 5.06 (Pbr038936.1) to 7.87 (Pbr023570.1) (Table S1).

Phylogeny and multiple sequence alignment analyses of COLs

To fully identify the evolutionary relationship of *COL* homologous from the *P. bretschneideri* 'DangshanSuli' and other Rosaceae species, a phylogenetic tree was constructed using 171 COL amino acid sequences derived from 15 plant species, including 17 in *Arabidopsis* and 154 in 14 diverse Rosacea species. The topology of phylogenetic tree suggested that these COL proteins could be subdivided into three groups (Fig. 2), which are consistent with the previous studies [39]. As depicted in the tree, four *COL* homologs of *P. bretschneideri*, *PbrCOL7*, *PbrCOL8*, *PbrCOL9*, and *PbrCOL11*, are classified under group I; five homologs, *PbrCOL1*, *PbrCOL2*, *PbrCOL3*, *PbrCOL4*, and *PbrCOL5*, are categorized under group II;



Common name	Chromosomes	Genome genes	CO-like genes
Raspberry	14	33286	9
Strawberry	14	28588	7
Rose	14	45469	8
Peach	16	35389	9
Apricot	16	24448	9
Japanese apricot	16	29706	9
Plum	16	40067	8
Apple	34	45352	16
Zhongai 1	34	36485	9
Bartlett	34	37445	11
Duli	34	59552	13
DangshanSuli	34	42180	12
Cuiguan	34	42559	18
Nijisseiki	34	44879	16

Fig. 1 Phylogenetic tree and genome information for fourteen Rosaceae species. The species tree was constructed through Orthofinder software and visualized using MEGA 11. The genome data of fourteen Rosaceae species was downloaded from GDR, Genome Database for Rosaceae (http://www.Rosaceae.org/). The pentagram (left) indicates the occurrence of WGD. The values on the left indicate species divergence time (unit: MYA)

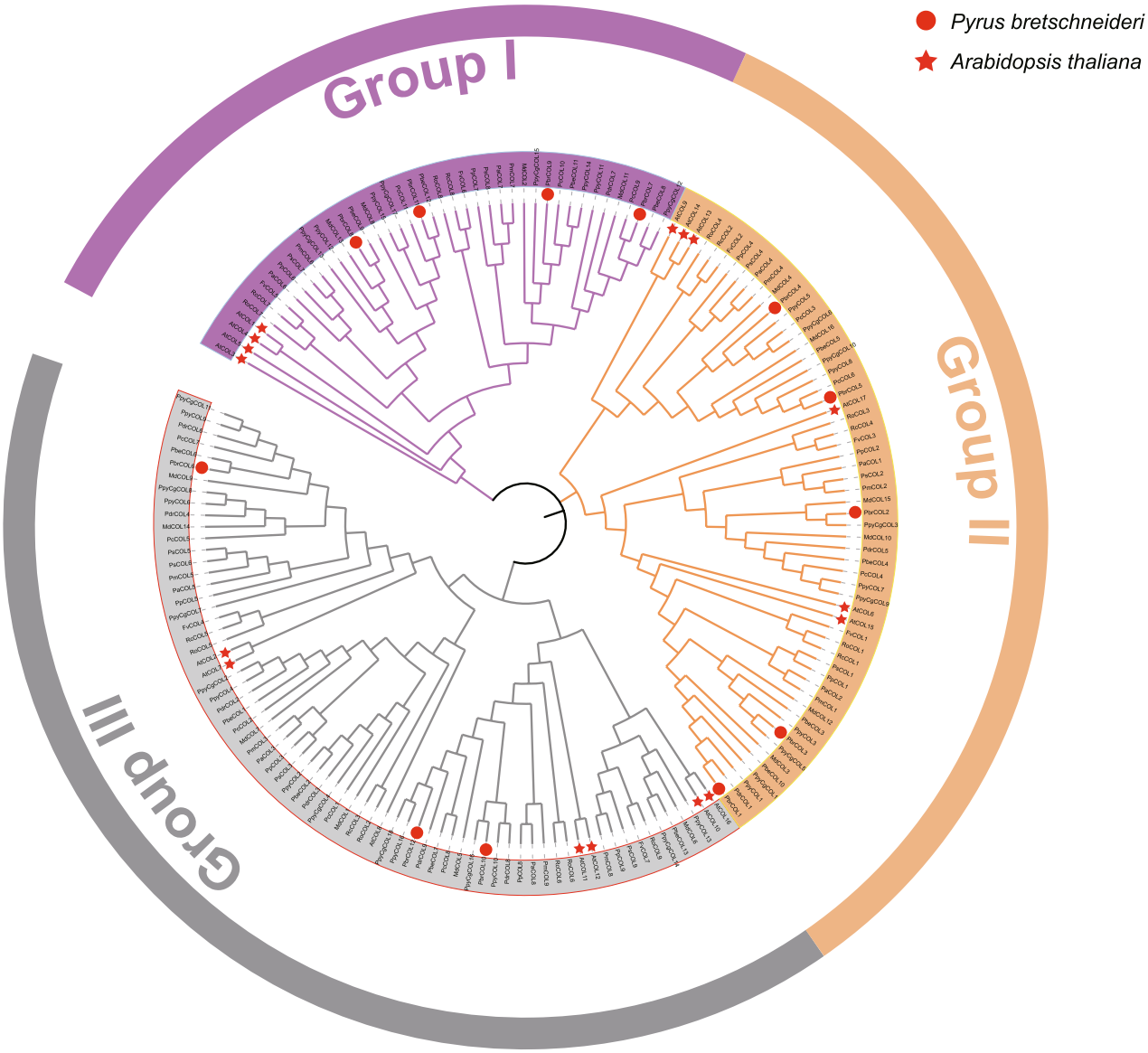


Fig. 2 Phylogenetic analysis of CONSTANS-like (COL) homologs in fourteen Rosaceae species and *Arabidopsis thaliana*. The clades of groups I, II, and III are marked in purple, orange, and silver, respectively. *Pyrus bretschneideri CONSTANS-like* genes (*PbrCOLs*) are indicated by red round points. *Arabidopsis thaliana COL* genes (*AtCOLs*) are indicated by red pentagram. At: *Arabidopsis thaliana*; Fv: *Fragaria vesca*; Md: *Malus.domestica*; Pdr: *Pyrus ussuriensis x communis*; Pc: *Pyrus Communis*; Pbe: *Pyrus betulifolia*; Pbr: *Pyrus bretschneideri*; Ppycg: *Pyrus pyrifolia* Cuiguan; Ppy: *Pyrus pyrifolia*. Pp: *Prunus persica*; Ps: *Prunus salicina*; Pm: *Prunus mume*. Pa: *Prunus armeniaca*; Rc: *Rosa chinensis*; Ro: *Rubus occidentalis*

whereas the remaining three homologs, *PbrCOL6*, *PbrCOL10*, and *PbrCOL12*, are placed in group III. The identity of 12 PbrCOL protein sequences ranged from 18.06% to 93.19% (Figure S1).

Protein structure and evolutionary relationship could be illustrated by multiple sequence alignment. The alignment results of the amino acid sequences of 'DangshanSuli', 'Cuiguan' and *Arabidopsis* showed that these proteins contained at least one B-box domain and one CCT domain (Figure S2). Furthermore, the amino acid sequence of B-box domain can be divided into two different specific domains: B-box1 (B1) and B-box2 (B2), according to their specificity of zinc ion-binding amino acid residues. Besides the B1 and B2 domains, some additional domain, such as B-box Superfamily (BSF), were also found in COL proteins of Rosaceae species. B-box domain is one type of conserved zinc fingers region, which owned the consensus C-X2-C-X7-9-C-X2-D-X4-C-X2-C-X3-4-H-X4-8-H (X can be any amino acid). The similar domain composition was found in all

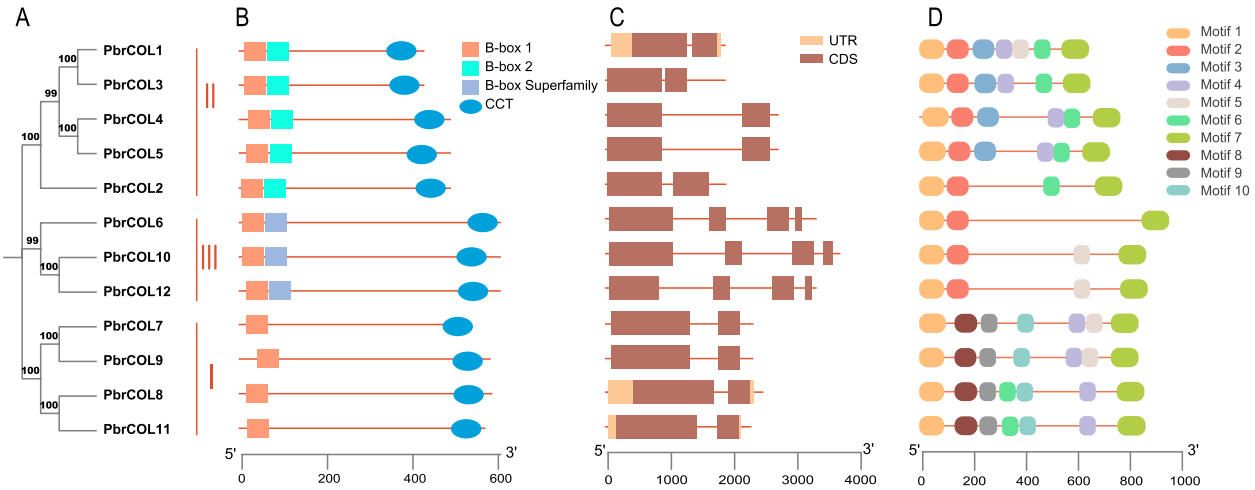


Fig. 3 Genomic and protein structures of PbrCOLs. **A** Phylogenetic analysis of 12 *PbrCOLs* genes which are categorized into three groups. The number indicates the bootstrap value. **B** The domain distribution of PbrCOLs. The CCT, B-box1, B-box2, and B-box superfamily are represented by light blue ovals, orange boxes, light green boxes, and blue–gray boxes, respectively. **C** The gene structure: CDS and UTR are represented by purple boxes and orange boxes, respectively. **D** The distinct 10 motifs distribution of PbrCOLs proteins. Every motif is represented by different colors which marked at top-right, respectively

PbrCOL proteins which is composed of approximately 40 residues. CCT domain is also highly conserved in *PbrCOL* genes, which contains approximately 45 amino acid residues. The domain composition perfectly aligns with the clustering of the phylogenetic tree (Fig. 2). Members in group I (*PbrCOL*7, *PbrCOL*8, *PbrCOL*9, and *PbrCOL*11) own only one B-box domains and one CCT domain (B1 + CCT); members in group II (*PbrCOL*1, *PbrCOL*2, *PbrCOL*3, *PbrCOL*4, and *PbrCOL*5) own two B-box domain (B1 and B2) and one CCT domain (B1 + B2 + CCT); while for members in group III (*PbrCOL*6, *PbrCOL*10, and *PbrCOL*12), besides a typical B1-type B-box domain and one CCT domain, the diverged zinc-finger structure domain called B-box Superfamily was also detected (B1 + BSF + CCT).

Gene structure, domain, motif, and Cis-acting elements analysis of PbrCOLs

The gene structure, particularly the distribution of introns, plays a crucial role in gene alternative splicing, enabling a gene sequence to generate different proteins. As depicted in Fig. 3, *PbrCOL* genes were categorized into three distinct groups, and a notable correlation between gene structure and phylogeny emerged among the groups, suggesting that proteins within the same group may potentially have similar functions. Specifically, two exons and one intron were identified in all genes belonging to group I (*PbrCOL*7, *PbrCOL*8, *PbrCOL*9, and *PbrCOL*11) and group II (*PbrCOL*1, *PbrCOL*2, *PbrCOL*3, *PbrCOL*4, and *PbrCOL*5). In contrast, all genes in group III possess four exons and three introns. The variation in the number of introns among *PbrCOL* genes may have contributed to differences in their genomic length.

The domain and motif identification of PbrCOL proteins were conducted using the NCBI and MEME websites. As a result, we identified 10 conserved motifs in the PbrCOL proteins and visualized the information using TBtools. It is evident that the domains and motifs of PbrCOL proteins are relatively conserved. Notably, all proteins contained Motif 1 and Motif 7. Interestingly, Motif 2 was exclusively detected only in the group II (B1 + B2 + CCT) and group III (B1 + BSF + CCT), and was replaced by Motif 8, Motif 9, and Motif 10 in the group I (B1 + CCT), indicating that this may be a highly conserved functional domain in PbrCOL. The analysis of conserved motifs provides robust evidence that strongly supports the clustering of the phylogenetic tree.

The function and expression of a gene are related to the type of cis-acting element contained in the upstream promoters [40]. In our study, we identified a total of 1145 cis-acting elements in the promoter regions of 12 *PbrCOL* genes. There are diverse functions of these cis-acting elements, including 140 cis-acting elements related to hormone such as MeJA, abscisic acid, gibberellin, and auxin, and 40 stress-related cis-acting element of low temperature and drought. Here, we selected a total of 166 light-related cis-acting elements for 12 gene promoters of *PbrCOL* in the *P. bretschneideri* 'DangshanSuli'. We then focused on 7 varieties of important cis-acting elements for further analysis, including light responsive elements, cis-acting regulatory elements involved in light responsiveness, MYB binding sites involved in light

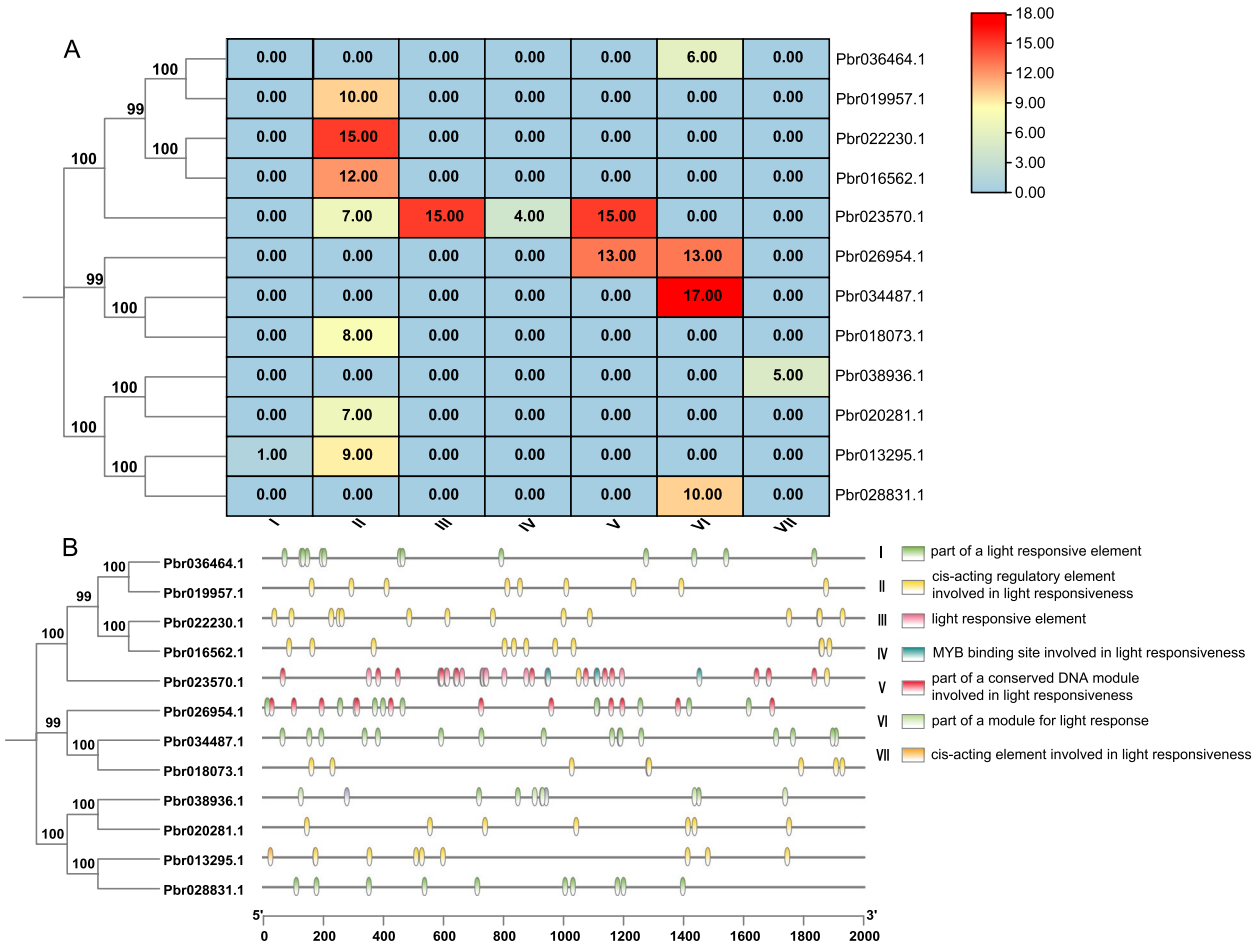


Fig. 4 **A** The heatmap of cis-element that distributed on the 2-Kb promoter sequences of *PbrCOLs*. The number in the rectangle indicates the total number of cis-element in different categories. (B) Distribution of the major light-related cis-elements in the 2-Kb promoter sequences of 12 *PbrCOL* genes. Different cis-elements are represented by different symbols as indicated

responsiveness, and conserved DNA modules involved in light responsiveness. Notably, all *PbrCOL* genes contained light-responsive elements, and type II cis-acting element accounts for the largest proportion, which is about 40.96%. (Fig. 4). This result provides further evidence for the *COL* genes respond to light stimuli.

Recurrent tandem and proximal duplication occurred following whole-genome duplication

The *Ks* (synonymous substitutions per site) value is commonly used to estimate the evolutionary dates of genome or gene duplication events. Here, we used the *Ks* value to estimate the duplication of *COL* gene pairs from each of Rosaceae species. In *P. bretschneideri*, two whole-genome duplication (WGD) events were detected including an ancient WGD event, which corresponds to the paleo-hexaploidization (γ) event shared by core eudicots that took place ~ 140 Mya (Ks ~ 1.5–1.8) [41] and the recent WGD that is inferred to have occurred 30–45 Mya (Ks ~ 0.15–0.3) [1]. The *Ks* values of the majority of WGD-derived *COL* gene pairs ranged from 0.10 to 4.37 in our study (Figure S3), suggesting that these genes may have derived from both ancient and recent WGD event. It is noteworthy that the majority of *COL* gene pairs belonged to the WGD and transposed duplication (TRD) categories. Interestingly, the *Ks* value of both events is more or less the same, indicating that *COL* genes underwent frequent small-scale gene duplications after ancient or recent genome duplication events, making them older in age. However, the *COL* gene pairs derived from dispersed duplication event are not found in apple and pear. The higher *Ks* values of *COL* gene pairs that derived from WGD suggesting that these *COL* genes of Rosaceae species have been retained from the ancient eudicot γ duplication event.

To further detect which selection process drove the evolution of the *COL* gene family in nine Rosaceae species, *Ka*/*Ks* values of paralogous *COL* gene pairs were

calculated, which represents the ratio between the non-synonymous replacement rate (*Ka*) and the synonymous replacement rate (*Ks*) of two protein-coding genes. This ratio can determine whether there is selection pressure acting on the protein-coding gene. $Ka/Ks > 1$ indicated positive selection, which means the non-synonymous substitution is beneficial to the species and has been preserved by positive selection of the environment; $Ka/Ks = 1$ demonstrated neutral evolution; and $Ka/Ks < 1$ suggested negative (or purifying) selection, which means most non-synonymous substitutions are harmful, and most non-synonymous substitutions are purified. In our study, *Ka*/*Ks* ratios for all *COL* gene pairs in each species showed that all of them were lower than one, indicating that purifying selection was the primary force driving the evolution of *COL* family gene in Rosaceae species (Fig. 5). *COL* genes derived from TRD showed relatively high *Ka*/*Ks* ratios in the species investigated, and TRD often leads to the formation of pseudogenes, while other types of duplication lead to rapid expansion of the plant genome. Overall, the results from the *Ks* and *Ka*/*Ks* analysis support the hypothesis that WGD and TRD are the main duplication event promoting the evolution of *COL* genes in Rosaceae species and contributing to the formation of homologous *COL* genes.

Chromosomal distribution and synteny analysis of COL genes in Rosaceae species

To investigate the evolutionary relationship of *COL* genes in nine Rosaceae species, we analyzed the genomic distribution and synteny of *COL* genes on the chromosomes. The distribution of *COL* genes on chromosomes exhibited relatively consistent patterns within certain species groups. For instance, in subgroup I ($n = 7$), which includes raspberry, rosa, and strawberry, there is at least one *COL* gene located on chromosome 4, indicating that the homologous *COL* gene pairs may have originated from this chromosome. Similarly, a similar distribution was observed in subgroup II ($n = 8$) and III ($n = 17$), with at least one *COL* gene located on chromosome 5. Apple and pears possess more *COL* genes, which are distributed on chromosomes 6, 13, 14, 15, 16, and 17, ranging from 5 to 17 genes. The differentiation between subgroups I, II, and III suggests that WGD has led to an increase in *COL* homologous genes in contemporary Rosaceae species.

In *P. bretschneideri* 'DangshanSuli', we identified 12 homologous *PbrCOL* gene pairs (Fig. 6), which contained all 12 *PbrCOL* genes. Further, the Dupgen-finder software was used to perform duplication event type analysis, and we found that most (83.33%) of *PbrCOL* genes (*PbrCOL*1, *PbrCOL*3–5, *PbrCOL*7–12) originated from a WGD event, while *PbrCOL*2 originated from TRD event, and *PbrCOL*6 originated from tandem duplication event. It is notable that the *COL* homologous gene pairs remarkably decreased in *Pyrus ussuriensis* × *hybrid* 'Zhongai 1', indicating the hybridization may purify the *COL* homologous gene pairs. Additionally, two homologous *COL* gene pairs were identified in raspberry, apricot, and peach; while only one *COL* gene pair was identified in rosa, strawberry, plum, and Japanese apricot (Figure S4).

Red and Blue Light influence the chlorophyl content and expression of PbrCOL genes

The transcriptome data of different tissues (leaf bud, flower bud, stem, adult leaf, juvenile leaf, stamen, fruit samples collected from 15, 25, and 55 days after flowering) was obtained from the Pear Expression Database (PearEXP, http://www.peardb.org.cn/) to detect the expression level of *PbrCOLs* (Fig. 7A). For the members of group I, *PbrCOL3* and *PbrCOL6* exhibited similar expression patterns in different tissues and showed relatively higher expression levels in most of pear tissues. In contrast, the lowest expression level was detected for the members of group II including *PbrCOL5*, *PbrCOL7*, *PbrCOL8*, *PbrCOL10*, *PbrCOL11* and *PbrCOL12*. Except for *PbrCOL8* and *PbrCOL11*, which had relatively high expression in juvenile leaf, the other *PbrCOL* genes in group II showed relatively lower expression level in most of tissues. The specific and various expression profiles of *PbrCOL* genes in different tissues and fruit developmental stages indicated the diverse functions. This observation lays a potential foundation for further research on the *COL* gene functions.

Previous studies uncovered the fact that the expression of the *COL* genes regulate the flowering time through a sophisticated mechanism to sense the light condition in most plant species. In *Fragaria vesca* (woodland strawberry), for instance, the expression of most *FveCOL* genes can be regulated by circadian clock and showed different diurnal expression patterns under SD or LD condition [42]. The research on the *COL* genes of *Arabidopsis* revealed that the wavelength of light significantly influences the expression of *COL* genes. In fact, *phyB*, the red-light receptor gene, delays flowering mainly by destabilizing the *COL* protein and leads to decreased expression of a florigen gene, *flowering locus T* (*FT*) [43]. Additionally, *CRYs,* the blue light receptor genes, contribute to light entrainment of the circadian clock, which affects the expression pattern of the clock-controlled gene *COL* [44]. Thus, we use *Pyrus betulifolia* as the plant material to perform the growth treatment with natural light (CK), red light, and blue light under LD condition to explore the function of *COL* genes in pear. As leave are the main organ through which the plant perceives the light signal, we used the healthy adult *Pyrus betulifolia* leave to measure the chlorophyll content of the

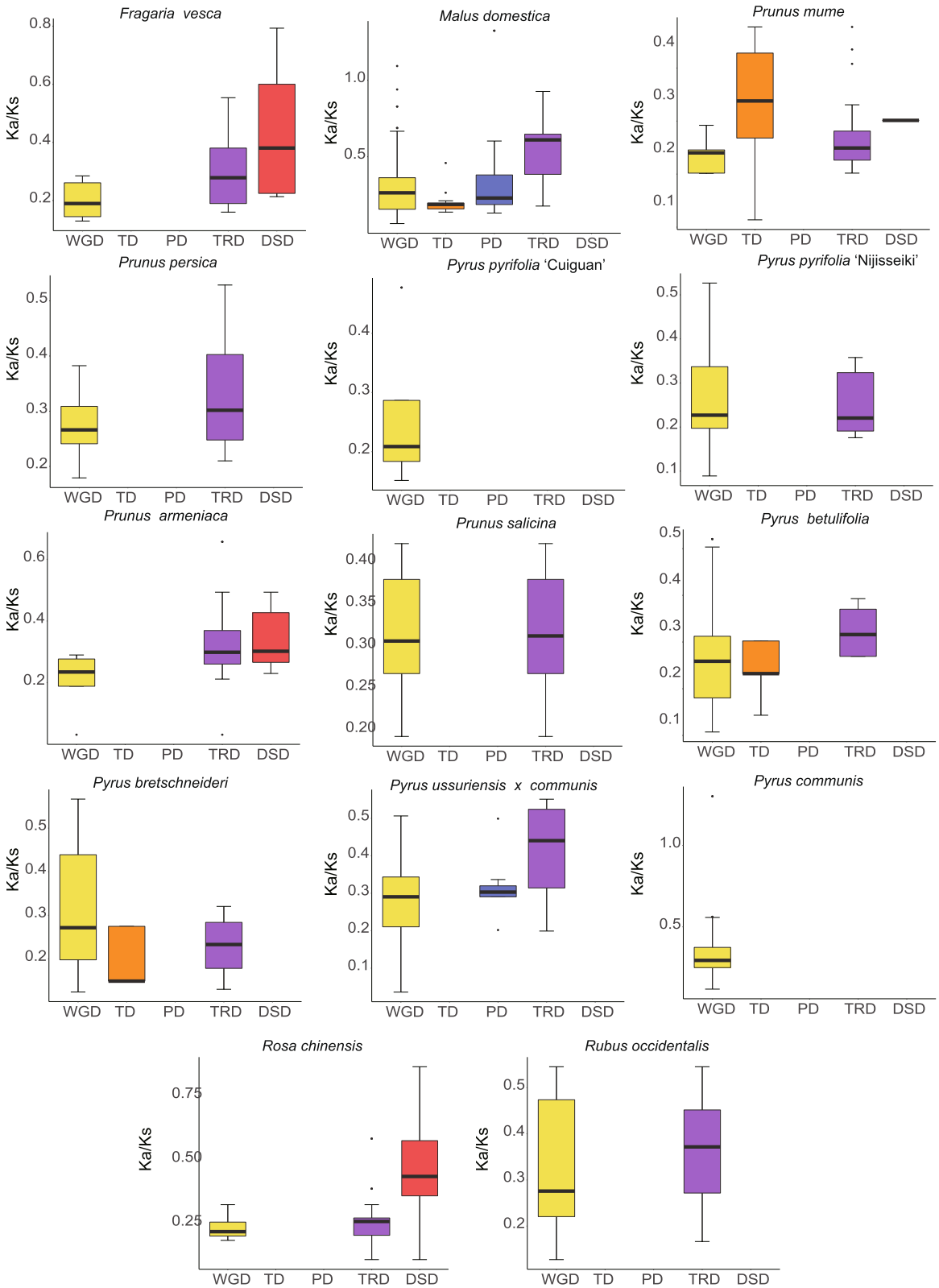


Fig. 5 Ka/Ks distribution of different modes of duplicated *COL* gene pairs in fourteen Rosaceae species. WGD: Whole-genome duplication; TD: Tandem duplicate; PD: Proximal duplication; TRD: transposed duplication; DSD: Dispersed duplication. The five modes of duplicated *COL* gene pairs are represented by yellow, orange, blue, purple, and red color, respectively

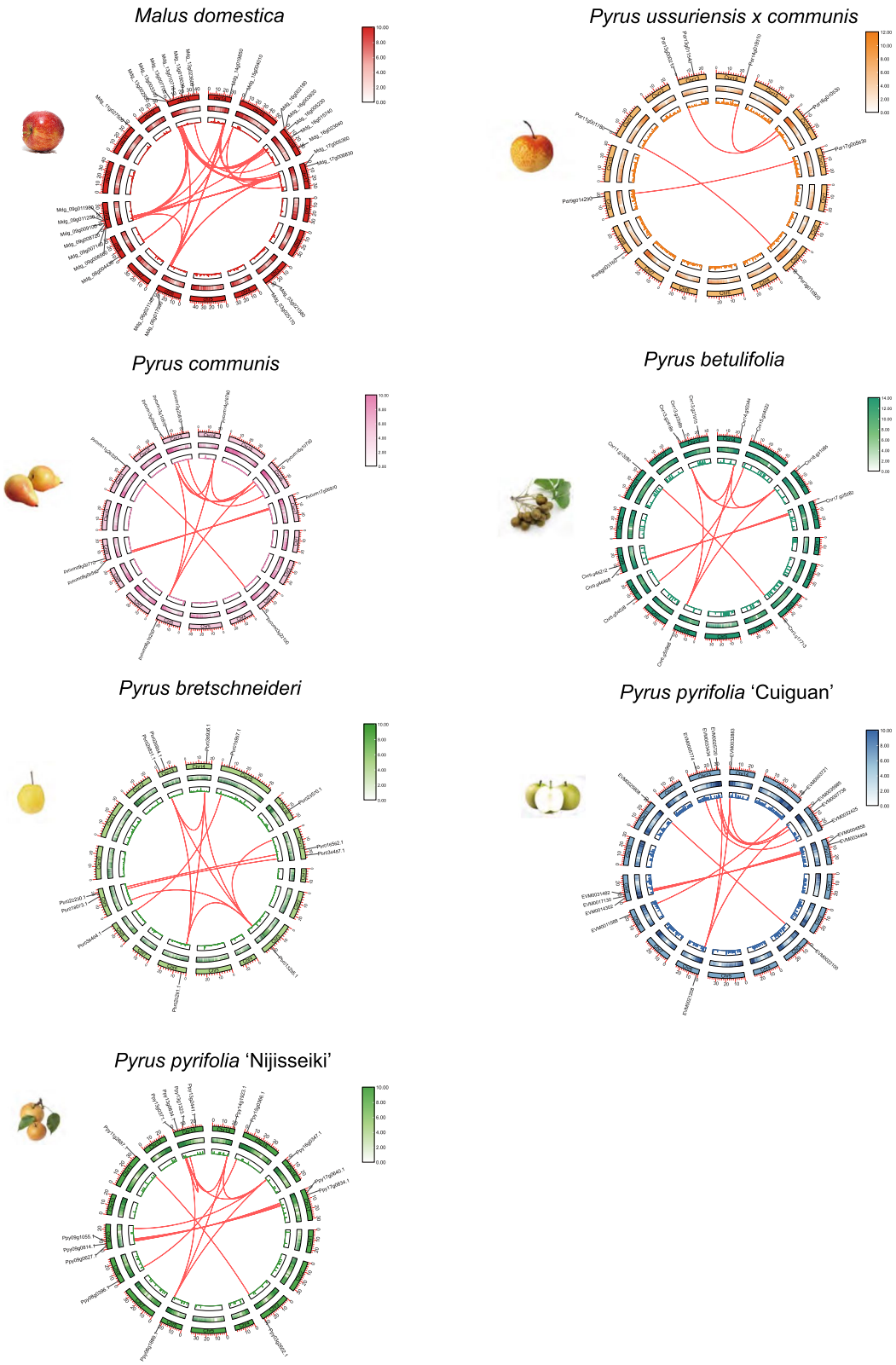


Fig. 6 Chromosomal localization and syntenic relationships of *COL* genes in apple and different pear genomes. *COL* genes are mapped on different chromosomes and syntenic gene pairs are linked by red lines

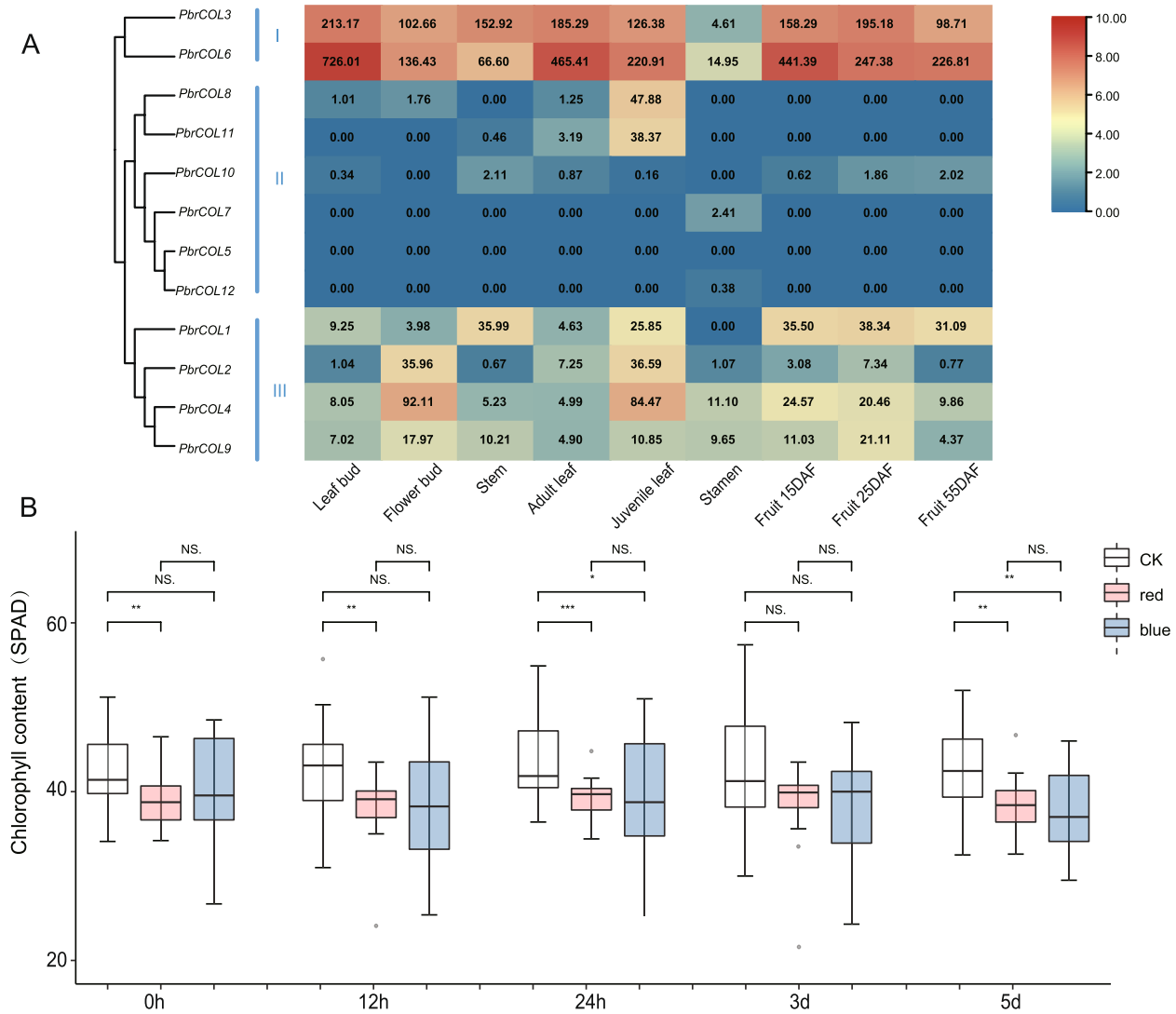


Fig. 7 **A** The heatmap of expression level of 12 *PbrCOLs* genes in different tissues. The transcriptome data was downloaded from PearEXP, Pear Expression Database (http://www.peardb.org.cn/). **B** The chlorophyll content of the pear leave of three treatment groups. Red box represents the red light, blue box represents the blue light, and white box represents the natural light

three treatments at the same time. The results showed that the chlorophyll content decreased significantly after 12 h (h) of red light irradiation, and the chlorophyll content of leaves irradiated with blue light was no different from that of the control after 12 h, whereas significantly decreased after 24 h. The chlorophyll content of leaves decreased after both types of light treatment, and there was no significant difference in leaf chlorophyll content after 12 h and 24 h. It is worth considering that after three days, there was no significant difference in the chlorophyll content of leaves in the three treatments. However, after five days, the chlorophyll content of leaves treated with red and blue light showed a significant decline compared with the control (Fig. 7B).

Further, *PbrCOL1*, *PbrCOL2*, *PbrCOL3*, *PbrCOL4*, *PbrCOL6*, and *PbrCOL9* were chosen for qRT-PCR analysis based on their relatively higher expression in adult leaf tissue. The results showed that the *PbrCOL2*, *PbrCOL3*, and *PbrCOL4* exhibited the similar expression pattern under nature light treatment, the expression peak appeared at the 12:00 a.m., and gradually decreased until the midnight (0:00 a.m.) (Fig. 8). This is consistent with the expression pattern of *COL* gene in *Arabidopsis* [45]. However, *PbrCOL1*, *PbrCOL6*, and *PbrCOL9* showed a different pattern of gradually decreased. For the treatment with red light, all of *PbrCOL* genes, except *PbrCOL4*, showed a remarkable increase at their expression peak. It is noteworthy that the expression patterns of

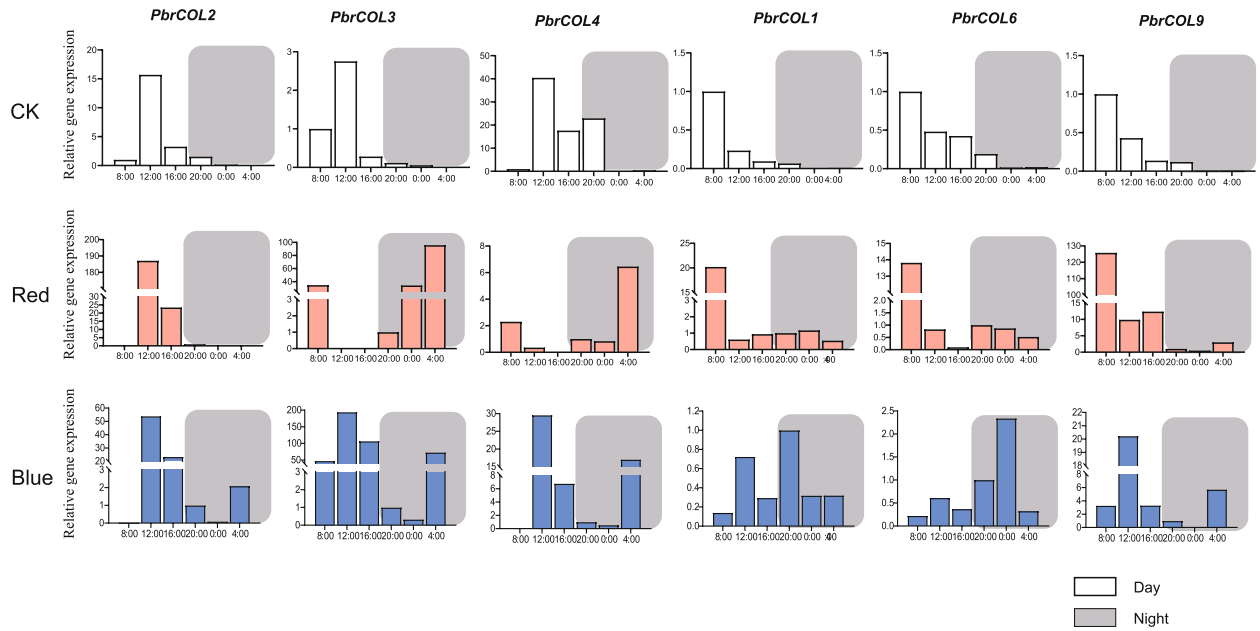


Fig. 8 Expression patterns of six *PbrCOL* genes verified by qRT-PCR analysis at the six time points. The left y-axis represents the relative expression levels obtained by qRT-PCR analysis. The times (8:00–4:00) represent six different time points of a day. The white rectangle indicates the natural white light treated group. The red rectangle indicates the red light treated group. The blue rectangle indicates the blue light treated group. The silver rectangle indicates the night

PbrCOL3 and *PbrCOL4* are completely contrary, while the other *PbrCOL* genes maintain a similar pattern. Further, under the blue light treatment, the expression patterns of *PbrCOL2*, *PbrCOL3*, and *PbrCOL4* are similar to the nature light, yet there are a second peak appeared at 4:00 a.m. in the morning. Additionally, the expression peak of *PbrCOL1* and *PbrCOL6* appeared at 8:00 p.m. and 0:00 a.m., suggesting the blue light may promote the *COL* stability in the dark condition.

The overexpression of PbrCOL3 in the Arabidopsis and Pyrus betulifolia seedling

Further, we constructed the 35S overexpression vectors to achieve stable transformations in *Arabidopsis thaliana* and transient transformations in tobacco and pear seedlings by using *Agrobacterium*-mediated genetic transformation method. Subcellular localization observations were conducted through injections into the leaves of *Nicotiana benthamiana*, revealing that *PbrCOL3* is predominantly expressed in the nucleus as compared with the empty vector controls (Fig. 9A), which is consistent with previous findings of the homologous gene *At5G24930* (*COL4*) in *Arabidopsis* [46].

Through the transient transformation of pear seedlings via *Agrobacterium* injection and subsequent observation of their phenotypic changes, it was discovered that upon injection with *Agrobacterium* that containing the overexpression vector of *PbrCOL3*, the Chlorophyll content in pear seedlings significantly increased on the fifth day, while the other groups showed no significant difference (Fig. 9B and C), suggesting that *PbrCOL3* positively regulates the synthesis of chlorophyll. Subsequently, based on previous studies on the flowering pathways in *Arabidopsis*, we selected seven genes that play pivotal roles in the regulation of flowering: *CO* [21], *CRY2* [23], *FT* [47, 48], *PHYA* [49, 50], *PHYB* [51, 52], and *GI* [53].

(See figure on next page.)

Fig. 9 Transient and stable transformation of the *PbrCOL3* gene. **A** The subcellular localization results were conducted in tobacco by using *Agrobacterium*-mediated transformation. **B** Wild-type pear seedlings, as well as pear seedlings injected with *Agrobacterium* containing overexpression vectors and transient silencing vectors. WT: wild type. OE-empty: empty overexpression vector. OE-PbrCOL3: overexpression vector with PbrCOL3. VIGS-empty: empty gene silencing vector. VIGS-PbrCOL3: gene silencing vector with PbrCOL3. **C** Changes in chlorophyll content in the leaves of wild pear seedlings under control and various treatment groups on Day 0 and Day 5. **D** The qRT-PCR result of six key flowering-related genes. **E** The picture of wild type and PbrCOL3 overexpressing *Arabidopsis thaliana*. **F** After extracting DNA from *Arabidopsis*, agarose gel electrophoresis was performed

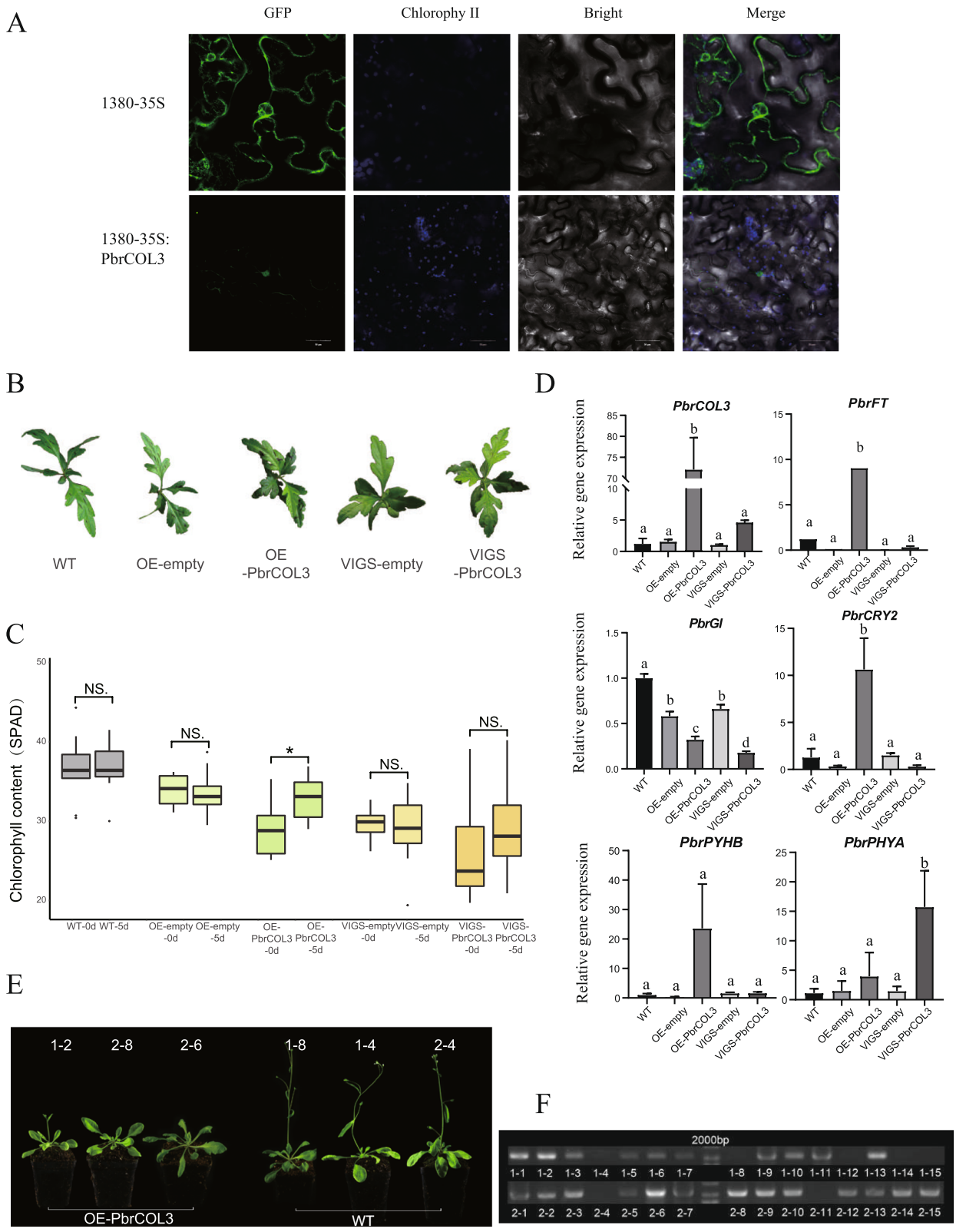


Fig. 9 (See legend on previous page.)

By performing a BLAST sequence alignment, we identified the homologous genes of these *Arabidopsis* genes within the pear genome. Then, we performed qRT-PCR analysis to observe the expression levels of these flowering-related key genes in pear seedlings that injected with transient overexpression and transient silencing vector of *PbrCOL3*. The results indicated that *PbrCRY2*, *PbrFT*, and *PbrPHYB* exhibited similar expression patterns, displaying a significant increase of expression level in pear seedlings overexpressing *PbrCOL3*. In contrast, *PbrPHYA* demonstrated a notable increase in expression in pear seedlings subjected to transient silencing of *PbrCOL3*. *PbrGI* was found to have the highest expression level in the control group, while its expression generally decreased across the other treatment groups, showing the lowest relative expression in the transient silencing treatment group (Fig. 9D). These results indicated that the expression of gene *PbrCOL3* significantly affected the expression levels of other key flowering-related genes.

To further investigate the effect of *PbrCOL3* on plant flowering, we utilized the floral dip method to introduce this gene into *Arabidopsis*. After the *Arabidopsis* plants have produced seeds following the treatment, the seeds are sown on selective medium. Subsequently, normal-growing seedlings are selected and transplanted into a substrate. Leaf tissue is then harvested for DNA extraction, followed by agarose gel electrophoresis. Interestingly, overexpression of *PbrCOL3* resulted in a late-flowering phenotype in *Arabidopsis* (Fig. 9E and F). This result is consistent with previous research on the *COL4* gene in *Arabidopsis* [46], suggesting that the *PbrCOL3* gene may perform similar functions and mechanisms in *Arabidopsis*. However, compared to the quantitative results observed in the leaves of young pear seedlings, the expression levels of *PbrFT* in the leave increased, indicating that this gene may promote flowering in perennial pear plants, In contrast, *PbrFT* exhibited different functions in the annual plant *Arabidopsis*, implying that functional differentiation between annual and perennial plants may account for this discrepancy, as studies have reported different flowering regulatory pathways between perennial and annual plants [54].

Discusstion

COL genes have been widely reported in many plant species, such as *Arabidopsis* [24], rice [30], soybean [55], populus [56], and barley [57], and have been found to significantly impact the regulation of flowering time, plant development, and resistance to abiotic stress [58]. However, a comprehensive genome-wide investigation of the COL family in pear and other Rosaceae species has not been reported to date.

In this study, a genome-wide exploration of the *COL* gene family identified a total of 154 genes, including 12 in 'DangshanSuli', 11 in 'Bartlett', 9 in 'Zhongai 1', 13 in *Pyrus betulifolia*, 18 in 'Cuiguan' pear, 16 in 'Nijisseiki', 16 in apple, 9 in raspberry, 7 in strawberry, 8 in rose, 9 in peach, 9 in apricot, 9 in Japanese apricot, and 8 in plum. The protein structural framework is crucial for predicting the proper functioning of proteins, and all 154 *COL* genes contain distinct B-box and CCT domains in their protein sequences. Our analysis revealed that the number of *COL* genes in apple and pear is almost double that of the other Rosaceae species, suggesting that WGD plays an essential role in the expansion of *COL* genes. The physical and chemical prediction (e.g., MW, number of amino acids, isoelectric points) of *COL* genes showed that the 12 PbrCOL proteins varied slightly, indicating that COL genes have remained stable during recent evolutionary processes.

The high-quality genome data of Rosaceae species in GDR allowed for a comprehensive analysis of evolutionary process in pear and other Rosaceae species. We reconstructed the phylogenetic tree of *COL* family genes in pear using the Neighbor-Joining methods, and three distinct subfamilies were determined. The results from the analysis of conserved motifs, gene features, and gene structures of *COL* family genes strongly supported the classification results obtained from the phylogenetic analysis. Previous analysis also showed that *COL* genes family of *Arabidopsis* is subdivided into three groups [39]. The gene structure analysis of *PbrCOLs* from 'DangshanSuli' showed highly consistency with the previous study [42]. Further, we found that the characteristics of *COL* genes were similar within each subfamily and varied among different subfamilies. The results of synteny analysis indicated that genetic constitution of the *COL* genes is quite conserved in those plant species during the evolutionary process. However, the hybrid pear (*Pyrus ussuriensis*×*communis*) exhibited a remarkable decrease in *COL* homologous gene pairs, suggesting that the hybridization may affect the homologous gene by an unknown mechanism.

Different types of gene duplication events, including WGD, tandem, proximal, transposed and dispersed duplications, are the main driving forces for gene family expansion in eukaryotes. WGD events can generate large numbers of duplicate genes in a very short period. Pear and apple experienced a recent lineage-specific WGD, whereas strawberry, sweet cherry, Japanese apricot, black raspberry, and peach did not undergo this duplication event.

The *COL* gene family has been reported and classified as the transcription factors that are involved in the regulation of flowering time in the photoperiod signaling pathway. Additionally, the *COL* gene family plays an important role in a wide range of biological processes, including cell

and seedling growth, tuberization, and resistance of abiotic stress [58, 59]. Research on *Arabidopsis* revealed that different layers of *COL* regulation at the transcriptional and posttranslational level are crucial for the ability of *Arabidopsis* to perceive and respond to photoperiod [60]. Moreover, different light qualities play antagonistic roles in the photoperiodic flowering response: while blue and far-red light (F-RL) promote flowering, red light (RL) delays it [57].

The function of the *COL* gene family in plant development can be hypothesized by analyzing their expression pattern. Therefore, we investigated the expression pattern of *PbrCOL* genes in the pear adult leaf in natural, red, and blue light conditions. As the result showed in the Fig. 8, under natural light, the expression patterns of *PbrCOL2*, *PbrCOL3*, and *PbrCOL4* are consistent with the LD photoperiod, while *PbrCOL1*, *PbrCOL6*, and *PbrCOL9* showed contrasting expression pattern, indicating different responsive mechanism and functional diversions of *PbrCOL* genes. During the day, the accumulation of CO reaches its peak, while the E3 ubiquitin ligase COP1/SPA complex can degrade the CO protein, reducing its abundance and inhibiting flowering. Based on previous studies on photoreceptors, blue light can activate CRY2, leading to the inactivation of the COP1/SPA complex, thereby increasing the stability of CO protein and promoting flowering [61]. By treating pear seedling leaves with blue light, it was found that the expression levels of *PbrCOL1*, *2*, *3*, *4*, *6*, and *9* genes were very low during the night under natural light conditions, but after blue light exposure, higher expression levels could be detected at night. This suggests that blue light may also play a similar role in pears, but further experiments are needed to confirm this.

In *Arabidopsis*, the control of flowering time is not limited to the CO protein alone; other members of the *COL* family also regulate flowering through different, overlapping, and antagonistic functions [62, 63]. Some members of the *COL* family function to inhibit flowering under both long day and short day conditions. *AtCOL4*, the highest sequence homology in Arabidopsis of *PbrCOL3*, functions to inhibit flowering. Researchers have found that *AtCOL4* and *AtCO* co-localize in the nucleus of cells, suggesting that they may be antagonistic genes. The precise balance of their expression regulates flowering in plants [46]. Based on the Ka/Ks value ratio of homologous gene pairs in the *PbrCOL* family, it is indicated that the *COL* genes in pears show evolutionary conservation, potentially suggesting the presence of antagonistic gene pairs between *COL* and *CO* in pears. Among the identified *PbrCOL* family members, *PbrCOL4* and *PbrCOL5* exhibit the highest sequence homology with *AtCO*. Further experimental verification is needed to determine if *PbrCOL4* and *PbrCOL5* co-localize with *PbrCOL3* and if protein–protein interactions exist to elucidate potential antagonistic effects.

Conclusion

In summary, a total of 154 *COL* genes were identified in fourteen Rosaceae genomes, with 12 *PbrCOL* genes originating from *P. bretschneideri* 'DangshanSuli'. Based on phylogenetic tree, gene structure, and conserved motif analyses, the *COL* family genes were classified into three subgroups (groups I-III). Evolution analysis of *COLs* suggests that lineage-specific genome duplication events in the ancestors of pear and apple played a significant role in the expansion of the *COL* gene family, and frequent small-scale gene duplications also contributed to the expansion of the *COL* gene family in Rosaceae species. The evolutionary analysis reveals that purifying selection was the primary force driving the evolution of *COL* genes. Transcriptome and qRT-PCR analysis indicated that *PbrCOL* genes were responsive to different light quality. These findings lay a foundation for further investigations into the molecular function of *COL* genes in Rosaceae species.

Abbreviations

COL	CONSTANS-like
GA	Gibberellin
LD	Long-day
SD	Short-day
Gl	GIGANTEA
FT	Flowering Locus T
SOC1	Suppressor of Overexpression of CO 1
Hd1	Heading Date 1
At	Arabidopsis thaliana
Mw	Molecular weights
pI	Isoelectric points
NJ	Neighbor-joining
TSS	Transcriptional start sites
ANOVA	A one-way analysis of variance
WGD	Whole Genome Duplication
TRD	Transposed duplication
TD	Tandem duplicate
PD	Proximal duplication
DSD	Dispersed duplication
RL	Red light
F-RL	Far-red light
Ka	Non-synonymous replacement rate
Ks	Synonymous replacement rate

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06325-z.

Additional file 1: Figure S1 The heatmap of the identity value of 12 *PbrCOLs* gens.

Additional file 2: Figure S2 Analysis of ZF-BBOX and CCT domains of COL gene family in *Arabidopsis thaliana*, *Pyrus bretschneideri*, and *Pyrus pyrifolia* Cuiguan.

Additional file 3: Figure S3 Ks distribution of different modes of duplicated COL gene pairs in fourteen Rosaceae species. WGD: Whole-genome duplication; TD: Tandem duplicate; PD: Proximal duplication; TRD: transposed duplication; DSD: Dispersed duplication. The five modes of duplicated COL gene pairs are represented by yellow, orange, blue, purple, and red color, respectively.

Additional file 4: Figure S4 Chromosomal localization and syntenic relationships of *COL* genes in seven Rosaceae species. A: raspberry, rosa, and strawberry. B: apricot, peach, plum, and Japanese apricot. *COL* genes

are mapped on different chromosomes and syntenic gene pairs are linked by red lines.

Additional file 5: Figure S5 Figure S5 Gel electrophoresis images during the construction of overexpression vector and VIGS vector. A: Using the NCBI database to predict the structural domains of PbrCOL3, it was confirmed that the unique CCT domain of the COL family is located in the amino acid sequence position 271-321. Therefore, this sequence was used as the insert fragment for the VIGS vector. B: The cDNA reverse transcription amplified the target fragment of the VIGS vector, with a length of 208bp. Lane 1 showed the 500bp DNA marker, while lanes 2 and 3 displayed the amplified fragment. C: After the recombinant plasmid was transformed into DH5α E. coli, a positive screening was conducted. Lanes 2-16 represented independent 15 single colonies, while lanes 1 and 17 showed the 500bp DNA marker. D: The cDNA reverse transcription amplified the full-length CDS sequence of the overexpression vector (1020bp, with the stop codon removed). Lane 1 displayed the 2000bp DNA marker, lanes 2-6 showed the amplified fragment of the target gene PbrCOL3, and lanes 7-11 contained other genes. E: After the recombinant plasmid was transformed into DH5α E. coli, a positive screening was conducted. Lanes 2-11 represented independent 10 single colonies, while lane 1 showed the 2000bp DNA marker.

Additional file 6.

Additional file 7: Table S1 Sequence analysis of 12 PbrCOLs. MW: molecular weight; pI: Isoelectric point; GRAVY: Grand average of hydropathicity.

Additional file 8: Table S2The list of primers designed for qRT-PCR.

Acknowledgements
The author of this paper is very grateful for the help provided by the following experts and personnel: Professor Xu Kai identified the pear seedling materials used in the experiment, Professor Sun Xuepeng provided the software and hardware needed for the experimental analysis and put forward valuable comments on the conclusion.

Authors' contributions
This research was done with the collaboration of all authors. K. X and X. S supervised and planned the study. X. L advised experiments and mainly participated in project administration, funding acquisition, manuscript writing, review and editing. K. C are mainly involved in the writing of original draft, data visualization, data validation, data curation, and conceptualization. X. L and D. L contributed to writing and correlation analysis. S. B, C. S and S. Z contributed to plant material cultivation, management, and data analysis.

Funding
This study was supported by "14th Five-Year Plan" fruit breeding project of Zhejiang Province: pear new varieties breeding (2021C02066-5), National Natural Science Foundation of China (32302486) and the Scientific Research and Development Foundation of Zhejiang A & F University (2022LFR048).

Data availability
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable. This manuscript does not involve researching about humans or animals.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Key Laboratory of Quality and Safety Control for Subtropical Fruit and Vegetable, Ministry of Agriculture and Rural Affairs, College of Horticulture Science, Zhejiang A&F University, Hangzhou 311300, Zhejiang, China.

Received: 5 December 2024 Accepted: 28 February 2025
Published online: 10 March 2025

References

1. Wu J, Wang Z, Shi Z, Zhang S, Ming R, Zhu S, et al. The genome of the pear (Pyrus bretschneideri Rehd). Genome Res. 2013;23:396–408.
2. The International Peach Genome Initiative, Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, et al. The high-quality draft genome of peach (Prunus persica) identifies unique patterns of genetic diversity, domestication and genome evolution. Nat Genet. 2013;45:487–94.
3. Shirasawa K, Isuzugawa K, Ikenaga M, Saito Y, Yamamoto T, Hirakawa H, et al. The genome sequence of sweet cherry (Prunus avium) for use in genomics-assisted breeding. DNA Res. 2017;24:499–508.
4. VanBuren R, Bryant D, Bushakra JM, Vining KJ, Edger PP, Rowley ER, et al. The genome of black raspberry (*Rubus occidentalis*). Plant J. 2016;87:535–47.
5. Li Y, Pi M, Gao Q, Liu Z, Kang C. Author Correction: Updated annotation of the wild strawberry Fragaria vesca V4 genome. Hortic Res. 2019;6:90.
6. Zhang Q, Chen W, Sun L, Zhao F, Huang B, Yang W, et al. The genome of Prunus mume. Nat Commun. 2012;3:1318.
7. Corbesier L, Gadisseur I, Silvestre G, Jacqmard A, Bernier G. Design in *Arabidopsis thaliana* of a synchronous system of floral induction by one long day. Plant J. 1996;9:947–52.
8. Burn JE, Smyth DR, Peacock WJ, Dennis ES. Genes conferring late flowering inArabidopsis thaliana. Genetica. 1993;90:147–55.
9. Burn JE, Bagnall DJ, Metzger JD, Dennis ES, Peacock WJ. DNA methylation, vernalization, and the initiation of flowering. Proc Natl Acad Sci USA. 1993;90:287–91.
10. Clarke JH, Dean C. Mapping FRI, a locus controlling flowering time and vernalization response in Arabidopsis thaliana. Molec Gen Genet. 1994;242:81–9.
11. Martinez-Zapater JM, Somerville CR. Effect of Light Quality and Vernalization on Late-Flowering Mutants of *Arabidopsis thaliana*. Plant Physiol. 1990;92:770–6.
12. Goto N, Kumagai T, Koornneef M. Flowering responses to light-breaks in photomorphogenic mutants of Arabidopsis thaliana, a long-day plant. Physiol Plant. 1991;83:209–15.
13. Wilson RN, Heckman JW, Somerville CR. Gibberellin Is Required for Flowering in *Arabidopsis thaliana* under Short Days. Plant Physiol. 1992;100:403–8.
14. Blázquez MA, Green R, Nilsson O, Sussman MR, Weigel D. Gibberellins Promote Flowering of Arabidopsis by Activating the *LEAFY* Promoter. Plant Cell. 1998;10:791–800.
15. Moghaddam MRB, Ende WVD. Sugars, the clock and transition to flowering. Front Plant Sci. 2013;4:22.
16. Wigge PA. Ambient temperature signalling in plants. Curr Opin Plant Biol. 2013;16:661–6.
17. Zhu Q-H, Helliwell CA. Regulation of flowering time and floral patterning by miR172. J Exp Bot. 2011;62:487–95.
18. Nagel DH, Kay SA. Complexity in the Wiring and Regulation of Plant Circadian Networks. Curr Biol. 2012;22:R648–57.
19. Putterill J, Robson F, Lee K, Simon R, Coupland G. The CONSTANS gene of arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell. 1995;80:847–57.
20. Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, et al. Hd1, a Major Photoperiod Sensitivity Quantitative Trait Locus in Rice, Is Closely Related to the Arabidopsis Flowering Time Gene CONSTANS. Plant Cell. 2000;12:2473–83.
21. Koornneef M, Hanhart CJ, Van Der Veen JH. A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana. Molec Gen Genet. 1991;229:57–66.

22. Mizoguchi T, Wright L, Fujiwara S, Cremer F, Lee K, Onouchi H, et al. Distinct Roles of *GIGANTEA* in Promoting Flowering and Regulating Circadian Rhythms in Arabidopsis. Plant Cell. 2005;17:2255–70.
23. Valverde F, Mouradov A, Soppe W, Ravenscroft D, Samach A, Coupland G. Photoreceptor Regulation of CONSTANS Protein in Photoperiodic Flowering. Science. 2004;303:1003–6.
24. Datta S, Hettiarachchi GHCM, Deng X-W, Holm M. *Arabidopsis* CONSTANS-LIKE3 Is a Positive Regulator of Red Light Signaling and Root Growth. Plant Cell. 2005;18:70–84.
25. Robson F, Costa MMR, Hepworth SR, Vizir I, Pin˜eiro M, Reeves PH, et al. Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants. The Plant Journal. 2001;28:619–31.
26. Zhang J, Hu Y, Xu L, He Q, Fan X, Xing Y. The CCT domain-containing gene family has large impacts on heading date, regional adaptation, and grain yield in rice. J Integr Agric. 2017;16:2686–97.
27. Almada R, Cabrera N, Casaretto JA, Ruiz-Lara S, González VE. VvCO and VvCOL1, two CONSTANS homologous genes, are regulated during flower induction and dormancy in grapevine buds. Plant Cell Rep. 2009;28:1193–203.
28. Deng X, Fan X, Li P, Fei X. A Photoperiod-Regulating Gene *CONSTANS* Is Correlated to Lipid Biosynthesis in *Chlamydomonas reinhardtii*. Biomed Res Int. 2015;2015:1–12.
29. Wang H, Zhang Z, Li H, Zhao X, Liu X, Ortiz M, et al. CONSTANS-LIKE 7 regulates branching and shade avoidance response in Arabidopsis. J Exp Bot. 2013;64:1017–24.
30. Liu J, Shen J, Xu Y, Li X, Xiao J, Xiong L. *Ghd2,* a *CONSTANS* -like gene, confers drought sensitivity through regulation of senescence in rice. EXBOTJ. 2016;67:5785–98.
31. González-Schain ND, Díaz-Mendoza M, Żurczak M, Suárez-López P. Potato CONSTANS is involved in photoperiodic tuberization in a graft-transmissible manner. Plant J. 2012;70:678–90.
32. Chen J, Chen J, Wang J, Kuang J, Shan W, Lu W. Molecular characterization and expression profiles of MaCOL1, a CONSTANS-like gene in banana fruit. Gene. 2012;496:110–7.
33. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23:2947–8.
34. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol. 2021;38:3022–7.
35. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol. 2018;35:1547–9.
36. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol Plant. 2020;13:1194–202.
37. Zhang X, Henriques R, Lin S-S, Niu Q-W, Chua N-H. Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. Nat Protoc. 2006;1:641–6.
38. Xiao Z, Xing M, Liu X, Fang Z, Yang L, Zhang Y, et al. An efficient virus-induced gene silencing (VIGS) system for functional genomics in Brassicas using a cabbage leaf curl virus (CaLCuV)-based vector. Planta. 2020;252:42.
39. Crocco CD, Botto JF. BBX proteins in green plants: Insights into their evolution, structure, feature and functional diversification. Gene. 2013;531:44–52.
40. Cui Y, Cao Q, Li Y, He M, Liu X. Advances in *cis* -element- and natural variation-mediated transcriptional regulation and applications in gene editing of major crops. J Exp Bot. 2023;74:5441–57.
41 Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, et al. The genome of the domesticated apple (Malus × domestica Borkh). Nat Genet. 2010;42:833–9.
42. Zhao X, Yu F, Guo Q, Wang Y, Zhang Z, Liu Y. Genome-Wide Identification, Characterization, and Expression Profile Analysis of CONSTANS-like Genes in Woodland Strawberry (Fragaria vesca). Front Plant Sci. 2022;13:931721.
43. Endo M, Tanigawa Y, Murakami T, Araki T, Nagatani A. Phytochrome-Dependent Late-Flowering accelerates flowering through physical interactions with phytochrome B and CONSTANS. Proc Natl Acad Sci USA. 2013;110:18017–22.
44. Wang X, Jiang B, Gu L, Chen Y, Mora M, Zhu M, et al. Publisher Correction: A photoregulatory mechanism of the circadian clock in Arabidopsis. Nat Plants. 2021;7:1516–1516.
45. Kim SY, Yu X, Michaels SD. Regulation of *CONSTANS* and *FLOWERING LOCUS T* Expression in Response to Changing Light Quality. Plant Physiol. 2008;148:269–79.
46. Steinbach Y. The Arabidopsis thalianaCONSTANS-LIKE 4 (COL4) – A Modulator of Flowering Time. Front Plant Sci. 2019;10:651.
47. Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, et al. FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of *Arabidopsis*. Science. 2007;316:1030–3.
48. Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T. A Pair of Related Genes with Antagonistic Roles in Mediating Flowering Signals. Science. 1999;286:1960–2.
49. Johnson E, Bradley M, Harberd NP, Whitelam GC. Photoresponses of Light-Grown phyA Mutants of Arabidopsis (Phytochrome A Is Required for the Perception of Daylength Extensions). Plant Physiol. 1994;105:141–9.
50. Mockler T, Yang H, Yu X, Parikh D, Cheng Y, Dolan S, et al. Regulation of photoperiodic flowering by *Arabidopsis* photoreceptors. Proc Natl Acad Sci USA. 2003;100:2140–5.
51. Cerdán PD, Chory J. Regulation of flowering time by light quality. Nature. 2003;423:881–5.
52. Guo H, Yang H, Mockler TC, Lin C. Regulation of Flowering Time by *Arabidopsis* Photoreceptors. Science. 1998;279:1360–3.
53. Park DH, Somers DE, Kim YS, Choy YH, Lim HK, Soh MS, et al. Control of Circadian Rhythms and Photoperiodic Flowering by the *Arabidopsis GIGANTEA* Gene. Science. 1999;285:1579–82.
54. Hyun Y, Vincent C, Tilmes V, Bergonzi S, Kiefer C, Richter R, et al. A regulatory circuit conferring varied flowering response to cold in annual and perennial plants. Science. 2019;363:409–12.
55. Awal Khan MA, Zhang S, Emon RM, Chen F, Song W, Wu T, et al. CONSTANS Polymorphism Modulates Flowering Time and Maturity in Soybean. Front Plant Sci. 2022;13:817544.
56. Chen H, Zhang S, Du K, Kang X. Genome-wide identification, characterization, and expression analysis of CCT transcription factors in poplar. Plant Physiol Biochem. 2023;204:108101.
57. Griffiths S, Dunford RP, Coupland G, Laurie DA. The Evolution of *CONSTANS* -Like Gene Families in Barley, Rice, and Arabidopsis. Plant Physiol. 2003;131:1855–67.
58. Zhang B, Feng M, Zhang J, Song Z. Involvement of CONSTANS-like Proteins in Plant Flowering and Abiotic Stress Response. IJMS. 2023;24:16585.
59. Khatun K, Debnath S, Robin AHK, Wai AH, Nath UK, Lee D-J, et al. Genome-wide identification, genomic organization, and expression profiling of the CONSTANS-like (COL) gene family in petunia under multiple stresses. BMC Genomics. 2021;22:727.
60. De Clercq I, Van De Velde J, Luo X, Liu L, Storme V, Van Bel M, et al. Integrative inference of transcriptional networks in Arabidopsis yields novel ROS signalling regulators. Nat Plants. 2021;7:500–13.
61. Zuo Z, Liu H, Liu B, Liu X, Lin C. Blue Light-Dependent Interaction of CRY2 with SPA1 Regulates COP1 activity and Floral Initiation in Arabidopsis. Curr Biol. 2011;21:841–7.
62. Cheng X, Wang Z. Overexpression of COL9, a CONSTANS-LIKE gene, delays flowering by reducing expression of CO and FT in Arabidopsis thaliana. Plant J. 2005;43:758–68.
63. Li F, Sun J, Wang D, Bai S, Clarke AK, Holm M. The B-Box Family Gene STO (BBX24) in Arabidopsis thaliana Regulates Flowering Time in Different Pathways. PLoS ONE. 2014;9(2):e87544.

Publisher's Note