Genome-wide identification and expressional analysis of carotenoid cleavage oxygenase (CCO) gene family in *Betula platyphylla* under abiotic stress

Jiajie Yu[1†], Yiran Wang[1†], Heming Bai[1], Xiang Zhang[1] and Ruiqi Wang[1*]

Abstract

Background Carotenoid cleavage oxygenases (CCOs) are a group of enzymes that catalyze the oxidative cleavage of carotenoid molecules. These enzymes widely exist in plants, fungi, and certain bacteria, and are involved in various biological processes. It would be of great importance and necessity to identify CCO members in birch and characterize their responses upon abiotic stresses.

Results A total of 16 *BpCCOs*, including 8 *BpCCDs* and 8 *BpNCEDs* were identified in birch, and phylogenetic tree analysis showed that they could be classified into six subgroups. Collinearity analysis revealed that *BpCCOs* have the largest number of homologous genes in *Gossypium hirsutum* and also have more homologous genes in other dicotyledons. In addition, promoter analysis revealed that the promoter regions of *BpCCOs* contained many abiotic stress-related and hormone-responsive elements. The results of qRT-PCR showed that most of the *BpCCOs* were able to respond significantly to ABA, PEG, salt and cold stresses. Finally, the prediction of the interacting proteins of BpCCOs by STRING revealed several proteins that may interact with BpCCOs and be involved in plant growth and development/abiotic stress processes, such as HEC1 (bHLH), ATABA1, ATVAMP714, etc.

Conclusion In this study, the CCO members were identified in birch in a genome-wide scale. These results indicate that *BpCCO* genes may play important roles in the abiotic stress responses of birch plants.

Keywords Carotenoid cleavage oxygenase (CCO), CCD, NCED, Abiotic stress, *Betula platyphylla*

†Jiajie Yu and Yiran Wang contributed equally to this work.

*Correspondence:
Ruiqi Wang
870831997@nefu.edu.cn
1State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry
University, Heilongjiang Harbin 150040, China

Background

Carotenoids are a class of isoprenoid compounds found in a wide range of photosynthetic and non-photosynthetic organisms (e.g. bacteria and fungi) and play an important role in the growth and development process [1, 2]. For example, xanthophylls and violaxanthin function as components of plant light-harvesting protein complexes [3]. In addition, carotenoids in chloroplasts are responsible for light absorption, electron transfer, and the elimination of triplet oxygen and superoxide anion during plant photosynthesis [4, 5]. On the other hand, the carotenoid-derived product zeaxanthin aldehyde can be converted to the phytohormone abscisic acid (ABA), which plays an important role in a number of biological processes such as plant abiotic stress response and seed development [6, 7]. In the carotenoid metabolic pathway, carotenoid cleavage oxygenases (CCOs) specifically catalyze the cleavage of conjugated double bonds in the polyene chain of the carotenoid molecule, resulting in the formation of a variety of deacylated carotenoids and their derivatives [6, 8]. In higher plants, based on the epoxy structure of substrate, CCOs can be further classified into carotenoid cleavage dioxygenases (CCDs) and nine-cis-epoxycarotenoid dioxygenases (NCEDs) [9]. CCDs are mainly responsible for the oxidative cleavage of carotenoids in higher plants, leading to the biosynthesis of biologically smaller apocarotenoids [5]. NCEDs can catalyze the cleavage of the 11,12 double bond of violaxanthin (C_{40}) or neoxanthin (C_{40}) to form xanthoxin (C_{15}), and this catalytic reaction carried out by NCEDs is considered to be the rate-limiting step in ABA biosynthesis [10, 11].

In recent years, with the accumulation of the studies on plant *CCO* family, *CCO*s have been found in several species. For instance, nine known *CCO* members have been found in *Arabidopsis* [12–14], 13 in poplar [14], 19 in *Nicotiana tabacum* [15], 16 in *Forsythia suspensa* [16], etc [17–19]. Meanwhile, *CCO* family members were found to be actively involved in growth and developmental regulation and stress response processes in plants. For example, *AtCCD7* and *AtCCD8* are capable of sequentially oxidizing β-carotene to produce monocerolactone hormone synthesis precursors, which play important roles in biological processes such as meristem formation, lateral root formation, seed germination, and response to drought and salt stresses [20–23]. In plants, the first carotenoid-cleaving enzyme involved in precursor flavin biosynthesis, NCED, was derived from an abscisic acid hormone-deficient maize (*Zea mays*) mutant. The mutant embryos were found to have normal sensitivity to ABA, and the detached leaves from mutant seedlings showed significantly higher rates of water loss than wild-type leaves [24]. In *Glycine max*, the expression levels of *CCO*s were found to change significantly upon salt,

drought, low and high temperature stress, suggesting that CCOs may be involved in multiple abiotic stress response processes [25]. The overexpression of sweet potato *carotenoid cleavage dioxygenase 4* (*IbCCD4*) reduced the salt tolerance of the transgenic *Arabidopsis* [26]. In cotton, virus-induced gene silencing of *GaNCED3a*, the ortholog of *GhNCED3a_A/D*, reduced the resistance of the transgenic plants not only to drought but also to salt treatment, and also led to significantly reduced proline content, high malondialdehyde content and high-water loss rate [27]. *Arabidopsis* overexpressing *Crocus sativus CCD4b* had longer roots and more lateral roots, as well as increased tolerance to salt, dehydration and oxidative stress compared to wild type [28]. Furthermore, *Brassica oleracea CCD1* and *CCD4* are highly responsive to both drought and salt stress [29]. In *Malus domestica*, the expression level of an *CCO* gene changed significantly under both salt and drought stress, suggesting that *MdCCO* is involved in the abiotic stress response process in apple [30]. The above results indicate that CCO family genes have important biological functions in plant growth and development and in response to abiotic stress. However, there has been no study to date to identify the CCO family genes in birch (*Betula platyphylla*) at the genome-wide level.

In this study, *Betula platyphylla* was selected as the research material. It is a widespread tree species in eastern Asia, born on mountain slopes or in forests at altitudes of 400–4100 m, with strong adaptability. Birch especially preferring moist soils, and is a pioneer tree species in secondary forests. Therefore, birch is used as an ideal plant material for researches on plant responses to abiotic stresses. Hu et al. found that the transgenic birch plants showed better tolerance to salt and osmotic stress upon the overexpression of *BpNAC012* gene [31]. Guo et al. identified the *BpIMYB46* gene from birch and found it to be involved in the salt and osmotic tolerance [32]. In Wang et. al.'s research, the expression of an *BpNAC* gene, *BpNAC90*, was responsive to drought stress. The overexpression of this gene conferred improved drought tolerance to the transgenic birch plants [33]. With the development of genome sequencing technology, the whole genome sequencing has been accomplished for birch in recent years [34]. Based on a current clear genomic background, a total of 16 members of the *Betula platyphylla* CCO (BpCCO) gene family have been identified with bioinformatics methods. Then, the analyses on gene structure, chromosomal localization, collinearity, and promoter activity were also performed. Additionally, we analyzed the expression patterns of *BpCCO*s under abiotic stresses using qRT-PCR. The protein interaction network of BpCCOs was also proposed. This study provides new insights into the evolutionary relationships of the *BpCCO* family, serving as a valuable reference for

investigating the biological functions of *CCO*s and the abiotic stress response of forest tree.

Methods

Identification of the CCOs in *Betula platyphylla*

To obtain the genome information of *Betula platyphylla*, the genome, coding sequences (CDS) and protein data of *Betula platyphylla* were searched and downloaded from the Phytozome v13.1 database (https://phytozome.jgi.doe.gov/pz/portal.html). A BLASTP search was performed and the Hidden Markov Model (HMM) [35] file (PF03055) was retrieved from the Pfam protein family database (http://pfam.xfam.org). The putative members of the carotenoid cleavage oxygenase family in birch were identified using HMMER v3.1. The sequences of these putative members were then further checked using Inter-Pro (http://www.ebi.ac.uk/interpro).

After obtaining the BpCCO family members, their amino acid sequence characteristics such as molecular weight, isoelectric point, amino acid number, aliphatic index and hydrophilic average (GRAVY) score were analyzed using ExPASy (http://www.expasy.org/).

Phylogenetic relationships and genestructures of BpCCOs

The conserved motifs of BpCCOs were identified by using online tool MEME 5.0 (http://meme-suite.org). After the alignment of the protein sequences of BpCCOs by using Clustal X [36], the phylogenetic tree was constructed via MEGA 11 software [37, 38] (number of bootstrap replicates=1000) with the Maximum likelihood (ML) method. Online software GSDS2.0 (Gene Structure Display Server: http://gsds.cbi.pku.edu.cn/) was used to analyze the exon and intron distribution patterns of *BpCCO*s, which were visualized by using Tbtools-II (v1.120) [39].

Chromosomal localization and collinearity visualization

The location information of the *BpCCO*s was retrieved from Phytozome database and mapped onto the chromosome using Tbtools-II (v1.120). To analyze the syntenic relationships between the *CCO*s from birch and other plant species, the *CCO* gene sequences from four dicots (*Salix purpurea*, *Populus trichocarpa*, *Arabidopsis thaliana*, *Gossypium hirsutum*) and one monocot (*Oryza sativa*) were retrieved from phytozome database. Gene duplication events were investigated using the MCScanX (Multicollinearity Scanning Toolkit) software [40] with default parameters, and were visualized using TBtools [39]. Tbtools-II (v1.120) was used to calculate the non-synonymous (Ka) and synonymous (Ks) ratio for duplicated gene pairs.

Cis-acting element analysis

The sequence 2000 bp upstream of the transcription start site (TSS) of each *BpCCO* was extracted from the Phytozome database. *Cis*-elements were predicted with PLACE [41] and visualized with TBtools.

GO (gene ontology) annotation of *BpCCO*s

GO annotation of *BpCCO*s was conducted using Tbtools-II (v1.120) software. The protein sequences were uploaded to Tbtools-II (v1.120), and then through BLAST in Phytozome database. The results were visualized and downloaded. All parameters were set as default.

Collection of miRNA targeting *BpCCO*s

The coding sequences of *BpCCO*s were uploaded to the psRNATarget database (http://plantgrn.noble.org/psRNATarget) to search for the miRNAs that target the *BpCCO*s. The targeting relationships were further predicted by graphical illustration. The results were visualized by Cytoscape (v3.0) software.

Construction of protein-protein interaction network

Interacting proteins of BpCCOs were predicted using STRING online software [42] and linear relationships in the predicted results were visualised using Cytoscape [43].

Yeast transformation and transcriptional activation activity verification

Yeast transformation assay was conducted to identify the transcriptional activation activity of BpCCD4 and BpNCED5. The coding sequences of *BpCCD4* and *BpNCED5* were cloned and then combined with pGBKT7 vector, respectively, by using two enzyme digestion. The recombinant vector (pGBKT7-BpCCD4, pGBKT7-BpNCED5), positive control (pGBKT7-53/pGADT7-T) and negative control (pGBKT7) were transformed into yeast competent cells (Y2H), respectively. The transformed yeast fluids were cultured on the nutrition-deprived media (SD/-Trp, SD/-Trp/-His/-Ade) for the test of growing state.

Yeast two hybrid assay

According to the results of protein-protein interaction network, two combinations (BpCCD4/BpABA1, BpNCED5/BpVAMP714) were selected for further verification. The coding sequences of *BpABA1* and *BpVAMP714* were cloned and then combined with pGADT7 vector, respectively. The vector combinations pGBKT7-BpCCD4/pGADT7-BpABA1, pGBKT7-BpNCED5/ pGADT7-BpVAMP714, pGBKT7-53/ pGADT7-T (positive control) and pGBKT7-LAM/ pGADT7-T (negative control) were transformed into yeast competent cells (Y2H), respectively. The

transformed yeast fluids were cultured on the nutrition-deprived media (SD/-Trp/-Leu, SD/-Trp/-Leu/-His/-Ade) for the test of growing state.

Salt-resistant yeast transformation

Salt-resistant yeast transformation assay to explore the function of *BpCCO*s in salt tolerance. The coding sequences of *BpCCD7*, *BpNCED2.5* and *BpNCED5* were cloned and combined with pYES2-NTB vector, respectively, with seamless cloning method. The recombinant vectors, pYES2-NTB-BpCCD7, pYES2-NTB-BpNCED2.5 and pYES2-NTB-BpNCED5, as well as negative control (pYES2-NTB) were transformed into salt-resistant yeast (INVSc1). The transformed yeast fluids were cultuerd on nutrition-deprived yeast media (SD/-Ura) with different concentration of salt (NaCl). The concentration gradient of salt was set 0, 0.3, 0.6, 1.0, 1.3, 1.6 and 2.0 M. The growth state of yeast indicates salt tolerance.

Plant materials and treatments

The plant materials (not in a publicly available herbarium) used in this study were white birch (*Betula platyt-phylla*) preserved by the State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University (Harbin, China). This plant species was firstly identified by Chen et al. via transcriptomic analysis in 2021 [44]. The deposition numbers of the genomic data of white birch are 679 for phytozome database and 78630 for NCBI database. They were cultivated with vegetative reproduction method and grown in a greenhouse with a 16-h light/8-h dark photoperiod at 25℃. For tissue-specific analysis, seedlings were grown in hydroponic culture (1/2 MS medium containing 25 g/L sucrose, 0.02 mg/L NAA and 0.4 mg IBA) for 60 d. After homogenization, plant samples including root, stem and leaf were collected, immediately frozen in liquid nitrogen and kept in a refrigerator at -80℃ for further experimental use.

The seedlings for ABA, PEG6000 or NaCl treatment were also grown in hydroponic culture for 60 d. For ABA treatment, we adjusted the concentration of ABA in hydroponic culture to 100 μM. For the PEG6000 treatment, we adjusted the concentration of PEG6000 in the hydroponic culture to 20% (W/V). For salt treatment, we adjusted the concentration of NaCl in the hydroponic culture to 200 mM. Plant samples (leaf) were collected at 0 h, 3 h, 6 h, 12 h, 24 h and 48 h of each treatment according to the result of tissue-specific analysis. For cold treatment, the treatment conditions were 4℃ for 3 h, seedlings were first grown in hydroponic culture for 60 d and then transplanted to soil for further growth for 30 days. After cold treatment, plant leaves were collected. Prior to sample preservation, all collected roots were rinsed with deionized water to remove the nutrient solution and then the excess water was absorbed with absorbent tissue. All these samples had three biological replicates.

qRT-PCR validation

Total RNA of the plant samples was extracted using the Mega Pure Plant RNA Kit (Msunflowers Biotech Co.,Ltd, Beijing, China). The quality of RNA was examined by gel electrophoresis and concentration measurement. The cDNA was then synthesised using a reverse transcription kit (PrimeScriptTM RT Reagent Kit, Takara Bio, Kusatsu, Japan). Primers used in qRT-PCR were designed according to the downloaded full-length cDNA sequences of *BpCCO*s, and *18 S* ribosomal RNA was selected as the internal reference gene. qRT-PCR was performed using THUNDERBIRD Next SYBR qPCR Mix (TOYOBO, Osaka, Japan). The reaction system was 20 μL and the conditions were as follows: predenaturation at 94℃ for 30 s, denaturation at 94℃ for 5 s, renaturation at 58℃ for 15 s, extension at 72℃ for 10 s, step 2 to 4 for 45 cycles, melting curve for 6 s. The reaction was performed on an Applied Biosystems 7500 Fast Real-Time PCR System. Each reaction had 3 biological replicates. The relative expression levels of *BpCCO*s were calculated using the $2^{-\Delta\Delta Ct}$ method [45].

Statistical analysis

Regular calculations were carried out with the assistance of Excel software. Significant differences were analyzed by one-way ANOVA followed by Duncan's multiple range test.

Result

Phylogenetic and structural analysis of CCO family members

A total of 16, 23, and 9 *CCO*s were obtained from *Betula platyphylla*, *Populus trichocarpa* and *Arabidopsis thaliana*, respectively, and a phylogenetic tree was generated based on their sequences. *CCO*s can be classified into two subfamilies, namely CCDs (II-VI) and NCEDs (I). Based on the previous studies on the *Arabidopsis CCO* family [46], CCDs were further classified into five subgroups (I-VI) (Fig. 1). Among them, subgroups II, III, IV and V corresponded to *CCD4*, *CCD1*, *CCD7* and *CCD8* in *Arabidopsis*, while members of subgroup VI appeared only in poplar and birch, suggesting that members of this subgroup may have appeared after the differentiation of herbaceous and woody plants. Among all the subgroups, subgroup I contained the highest number of genes (17 genes) and IV contained the lowest number of genes (3 genes).

The analysis revealed that the length of these proteins varied between 97 (BpNCED2.3) and 657 (AtNCED9) amino acids in all three species (Table 1). The molecular weights ranged from 10866.49 (BpNCED2.3) to 73107.61 Da (BpCCD1.1). The pI of the CCO protein ranged from
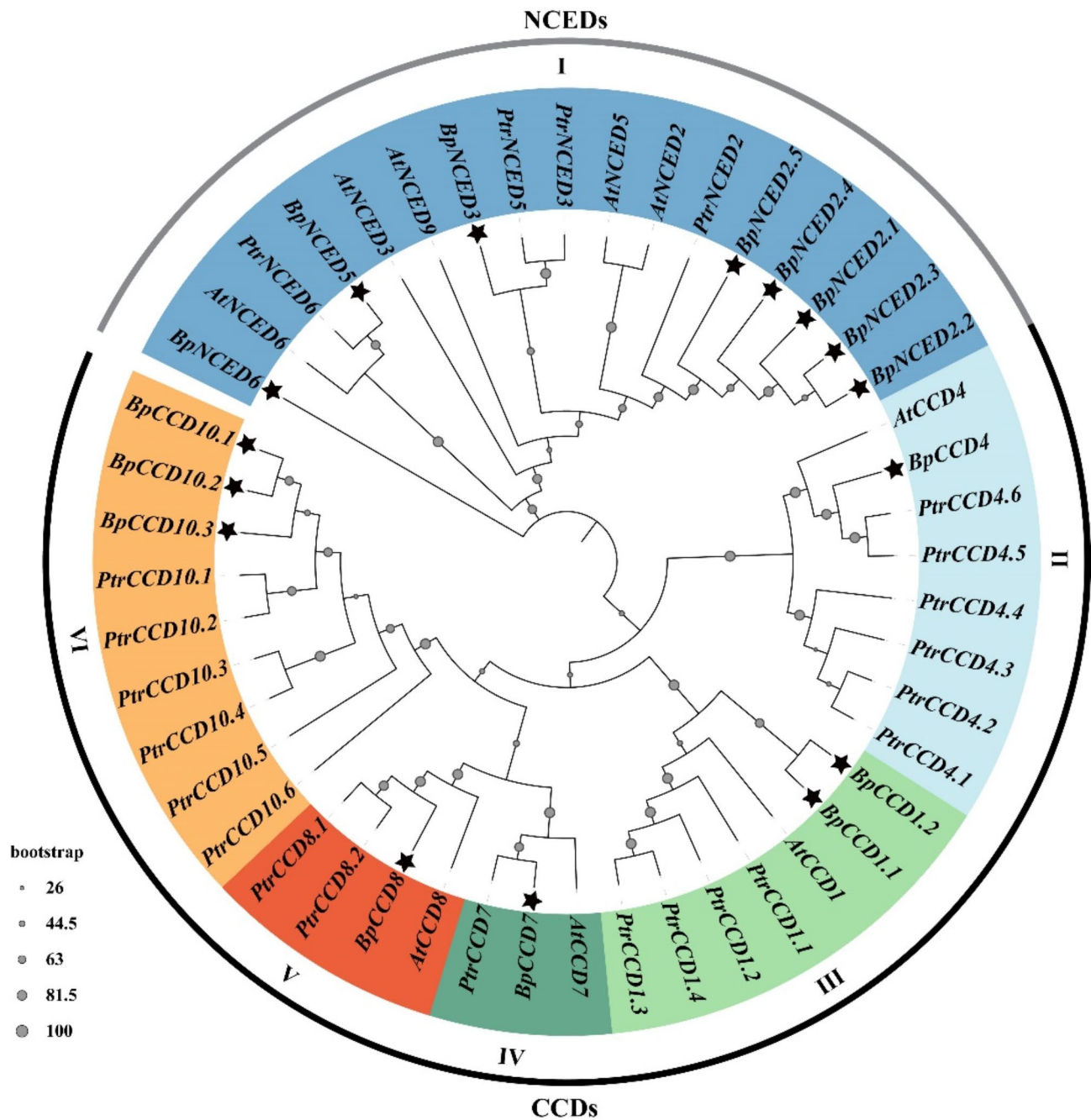
Fig. 1 Phylogenetic analysis of *CCO*s in *Betula platyphylla* (*Bp*), *Arabidopsis thaliana* (*At*), and *Populus trichocarpa* (*Ptr*). Maximum likelihood (ML) method with 1,000 bootstrap replicates was applied to draw the phylogenetic tree via MEGA11 software. The tree was divided into two subfamilies or six subgroups; each color represents one subgroup. Black stars indicate *BpCCO*s

4.67 (BpNCED2.4) to 9.73 (PtrCCD4.1), with 34 out of 48 CCO proteins being acidic (pI < 7). The aliphatic index of the CCO protein ranged from 65.84 (BpNCED2.4) to 90.48 (PtrCCD4.1), which indicates that most CCO proteins were stable. The grand average of hydropathicity (GRAVY) value ranged from −0.725 (BpNCED2.1) to -0.112 (BpNCED6). The GRAVY value of all CCO proteins from the three species is less than 0, which indicates that all these proteins exhibit hydrophilicity.

Exon-intron and motif patterns of *BpCCO*s

Based on the sequence information of *BpCCO*s, the gene structure and motif pattern of each member were analyzed using TBtools and MEME software, respectively (Fig. 2). Generally, genes within the same evolutionary

Yu *et al. BMC Genomics*   (2024) 25:872

Page 6 of 21

Table 1 Basic information of *CCO* genes in birch, *Arabidopsis* and Poplar

|	Gene ID	Name	Length (aa)	Molecular Weight (Da)	Theoretical pI	Aliphatic Index	GRAVY
Betula platyphylla	BPChr13G17223	*BpCCD1.1*	642	73107.61	8.83	80.3	-0.282
	BPChr13G17227	*BpCCD1.2*	473	53390.62	6.77	81.37	-0.195
	BPChr01G21247	*BpCCD10.1*	127	14548.63	9.09	87.56	-0.213
	BPChr06G26044	*BpCCD10.2*	378	42469.4	5.79	87.38	-0.278
	BPChr01G21232	*BpCCD10.3*	434	49373.75	5.42	80.35	-0.316
	BPChr02G19610	*BpCCD4*	614	67053.36	6.22	80.67	-0.181
	BPChr13G16033	*BpCCD7*	607	68239.68	6.18	81.7	-0.226
	BPChr01G05228	*BpCCD8*	554	61283.85	6.2	84.12	-0.239
	BPChr04G18364	*BpNCED2.1*	207	23308.56	9.24	70.14	-0.725
	BPChr08G05426	*BpNCED2.2*	205	23,513	8.75	76.05	-0.478
	BPChr02G18988	*BpNCED2.3*	97	10866.49	8.91	76.39	-0.469
	BPChr03G18237	*BpNCED2.4*	154	17540.82	4.67	65.84	-0.402
	BPChr07G30066	*BpNCED2.5*	602	66395.55	7.31	80	-0.345
	BPChr04G18342	*BpNCED3*	598	66437.81	8.22	81.89	-0.307
	BPChr06G29428	*BpNCED5*	606	67455.66	8.95	84.13	-0.283
	BPChr06G29332	*BpNCED6*	473	51925.89	6.55	84.4	-0.112
Arabidopsis thaliana	AT3G63520	*AtCCD1*	538	60908.11	6.05	84.78	-0.25
	AT4G19170	*AtCCD4*	595	65601.73	6.42	83.38	-0.218
	AT2G44990	*AtCCD7*	629	70850.3	6.3	75.44	-0.394
	AT4G32810	*AtCCD8*	570	63956.91	6.65	81.89	-0.326
	AT4G18350	*AtNCED2*	583	65066.57	6.56	86.6	-0.268
	AT3G14440	*AtNCED3*	599	65856.66	5.9	79.58	-0.283
	AT1G30100	*AtNCED5*	589	65337.05	5.6	79.61	-0.322
	AT3G24220	*AtNCED6*	577	63821.02	5.92	85.06	-0.237
	AT1G78390	*AtNCED9*	657	73015.17	6.59	81.39	-0.315
Populus trichocarpa	Potri.009G060500	*PtrCCD1.1*	567	63954.46	6.31	78.99	-0.327
	Potri.001G265400	*PtrCCD1.2*	551	62118.07	5.84	80.64	-0.338
	Potri.001G265600	*PtrCCD1.3*	545	61486.53	5.87	79.72	-0.311
	Potri.001G265900	*PtrCCD1.4*	545	61261.28	5.83	80.09	-0.317
	Potri.018G042650	*PtrCCD10.1*	599	66781.67	5.47	81.62	-0.25
	Potri.006G239402	*PtrCCD10.2*	627	71108.85	6.25	77.85	-0.326
	Potri.018G043000	*PtrCCD10.3*	607	68841.99	5.18	75.91	-0.366
	Potri.018G043500	*PtrCCD10.4*	607	68703.88	5.38	75.91	-0.372
	Potri.006G239200	*PtrCCD10.5*	521	58691.51	5.31	77.22	-0.355
	Potri.006G239300	*PtrCCD10.6*	605	68530.71	5.81	75.21	-0.38
	Potri.009G152000	*PtrCCD4.1*	294	33125.71	9.73	90.48	-0.172
	Potri.009G152300	*PtrCCD4.2*	567	63079.23	7.91	84.02	-0.191
	Potri.009G152200	*PtrCCD4.3*	592	66353.73	7.27	78.48	-0.272
	Potri.009G151900	*PtrCCD4.4*	595	66742.2	7.95	78.25	-0.278
	Potri.005G069100	*PtrCCD4.5*	611	66937.24	6.15	83.01	-0.199
	Potri.019G093400	*PtrCCD4.6*	612	66973.29	6.52	80.8	-0.215
	Potri.014G056800	*PtrCCD7*	613	69070.55	8.07	80	-0.339
	Potri.006G238500	*PtrCCD8.1*	557	61890.3	5.75	78.06	-0.313
	Potri.018G044100	*PtrCCD8.2*	557	62003.58	5.87	81.26	-0.288
	Potri.011G084100	*PtrNCED2*	597	66845.1	8.79	79.01	-0.403
	Potri.001G393800	*PtrNCED3*	599	66461.48	6.41	77.76	-0.371
	Potri.011G112400	*PtrNCED5*	592	65395.26	6.58	76.59	-0.375
	Potri.003G176300	*PtrNCED6*	602	66043.5	6.39	81.74	-0.223
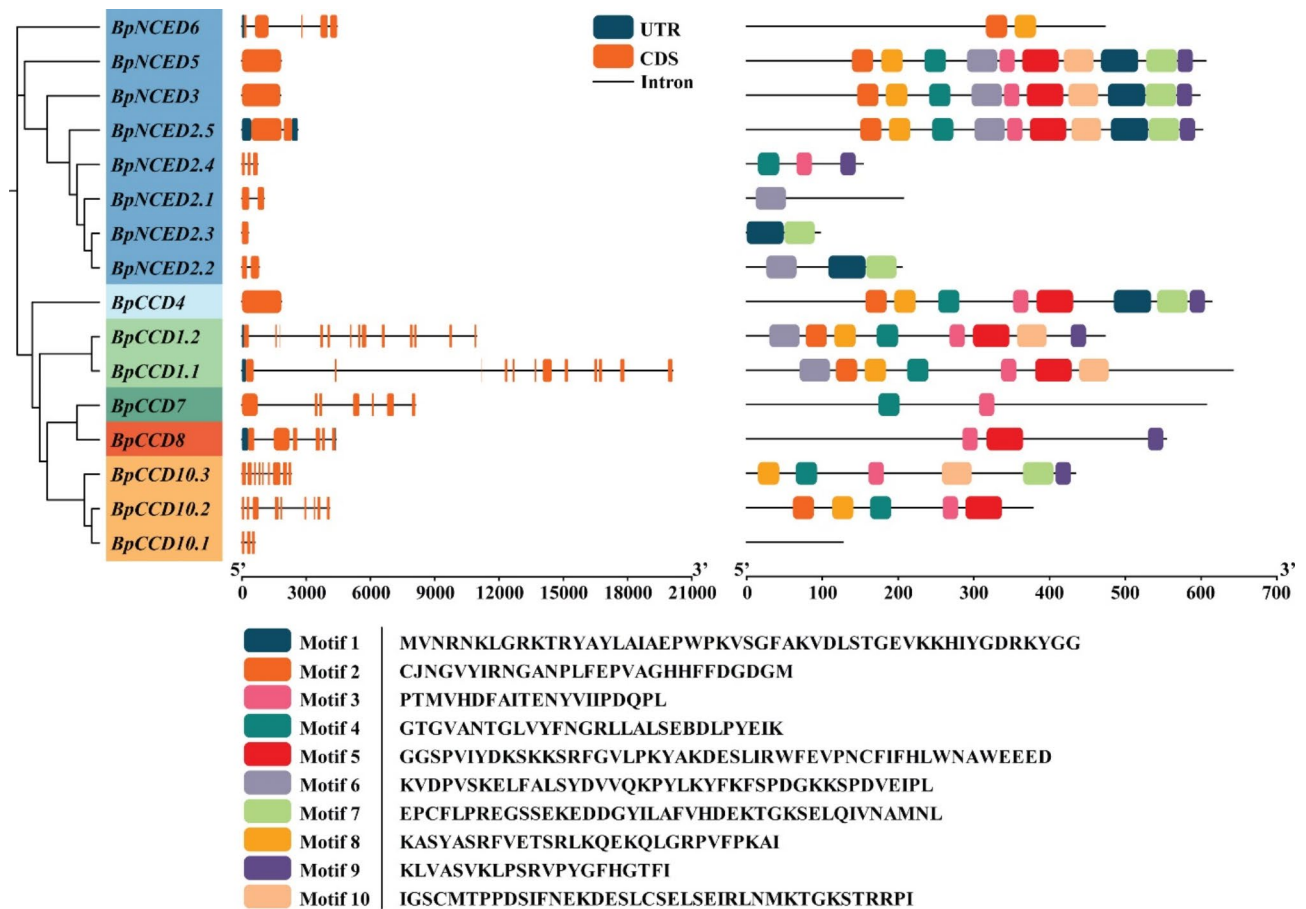
Fig. 2 Exon-intron and protein motif of the BpCCO gene family members. In the left part, boxes symbolize exon. Thin lines symbolize. In the right part, colored boxes symbolize different motifs. The clustering was performed according to the phylogenetic analysis. Scale bar indicates the length of gene (bp)

branch exhibited structural resemblances. The results showed that the number of exons in members of the NCED subfamily ranged from 1 to 5, with an average of 2.1. The number of exons in members of the CCD subfamily ranged from 1 to 13, with an average of 7.5. Except few genes, the number of exons was relatively conserved in most genes of the different subfamilies.

Chromosomal distribution and collinearity analysis of *BpCCO*s

By searching the latest version of *Betula platyphylla* v1.1 database, the position information of *BpCCO*s on chromosomes was obtained and visualized as shown in Fig. 3 and Supplementary Table 1. The results showed that *BpCCO*s were unevenly distributed on eight chromosomes. Among them, three *BpCCO*s were distributed on Chr01, Chr06 and Chr13 each, two *BpCCO*s were distributed on Chr02 and Chr04 each, and one *BpCCO* was distributed on Chr03, Chr07 and Chr08 each. To probe the gene duplication events of the BpCCO family members, we calculated the homology ratios and the total length of homologous fragments among the members.

A total of five pairs that evolved into homologous genes due to gene duplication events were identified, of which four were segmental duplications and one was a tandem duplication (*BpCCD1.1/BpCCD1.2*) (Fig. 3). We also found that the Ka/Ks values of these gene pairs were less than 1, indicating a strong purifying selection (Table 2).

The syntenic relationships between CCOs from birch and other four plant species (*Salix purpurea, Populus trichocarpa, Arabidopsis thaliana, Gossypium hirsutum*) were investigated. It is shown in Fig. 4 and Supplementary Table 2 that there were 9, 10, 6, 11, and 4 homologous pairs between *B. platyphylla* with *S. purpurea, P. trichocarpa, A. thaliana, G. hirsutum*, and *O. sativa*, respectively. *BpNCED2.5* had the highest number of homologous genes (nine) in the other species. *BpCCD8, BpCCD1.1, BpCCD10.1* and *BpCCD7* also had more than four homologous genes.

Expression patterns of *BpCCO*s in different birch tissues

The expression patterns of *BpCCO*s in different birch tissues were investigated by using qRT-PCR. As shown in Fig. 5 and Supplementary Table 3, a total of nine
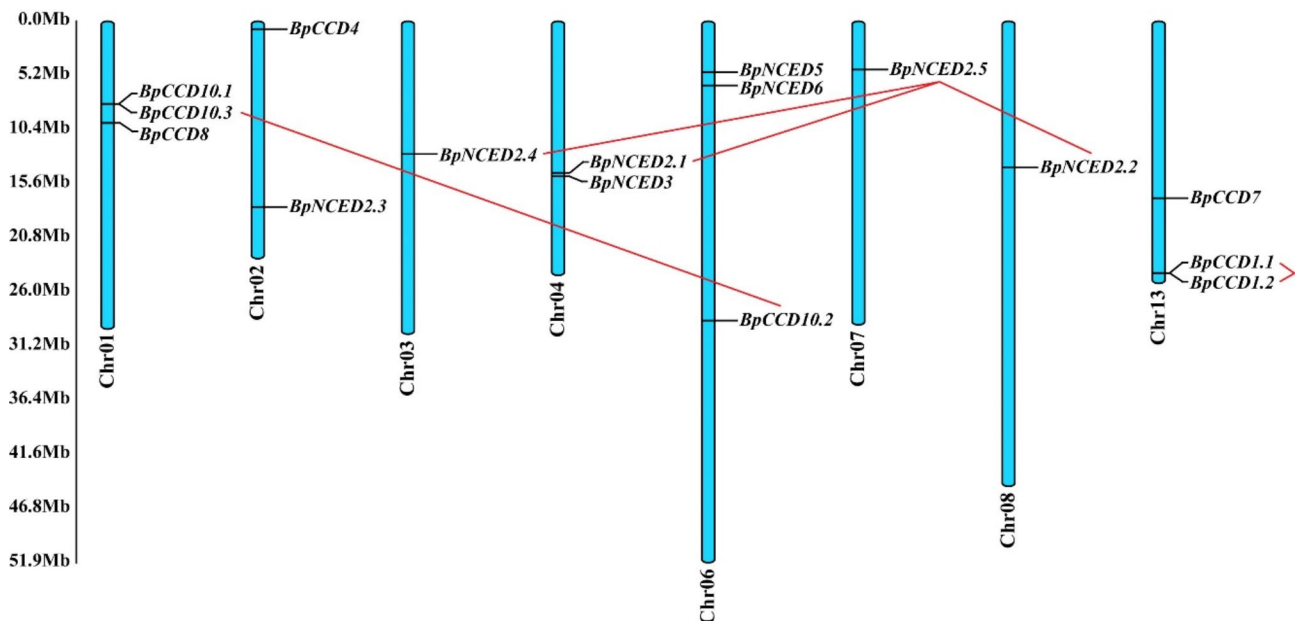
Fig. 3 Chromosomal localization and collinearity analysis of the BpCCO family. Blue rectangles symbolize chromosomes. Red lines indicate duplication gene pairs of *BpCCO*s

Table 2 The Ka/Ks ratios of duplication for *BpCCO*s

Duplicated gene pairs	Ka	Ks	Ka/Ks	The length of homologous fragment (bp)	Homology/%
BpCCD1.1/BpCCD1.2	0.113365629	0.165798838	0.683754063	1245	95.36
BpCCD10.2/BpCCD10.3	0.164266791	0.52110026	0.315230684	582	86.43
BpNCED2.1/BpNCED2.5	0.710159655	1.224909883	0.579764817	311	94.34
BpNCED2.2/BpNCED2.5	0.172012872	0.322564992	0.533265779	450	98.09
BpNCED2.4/BpNCED2.5	0.086501291	0.168350384	0.513817008	352	95.22

*BpCCO*s were highly expressed in leaves, namely *BpCCD4*, *BpNCED2.2*, *BpCCD1.2*, *BpNCED2.3*, *BpCCD10.1*, *BpNCED2.5*, *BpNCED2.1*, *BpCCD10.3* and *BpNCED3*. Two *BpCCO*s were highly expressed in roots, *BpCCD8* and *BpNCED2.4*. Five *BpCCO*s were highly expressed in stems, *BpCCD7*, *BpCCD1.1*, *BpCCD10.2*, *BpNCED5* and *BpNCED6*. Overall, individual members of the *BpCCO*s were expressed in roots, stems and leaves, and most members were the highest expressed in leaves.

Cis-elements analysis of *BpCCO*s promoters

The *cis*-elements in the promoter regions of *BpC-CO*s were analyzed to explore the biological processes in which *BpCCO*s may be involved. It is shown in Fig. 6 that the different family members contain a number of elements associated with stress and hormonal responses. A large number of *cis*-elements associated with abiotic stresses were found in the promoter regions of most *BpCCO*s, such as the response to dehydration (CBFHV, MYBCORE, DRECRTCOREAT, MYCATRD22, etc.), ammonium (AMMORESIIUDCRNIA1 and AMMORESIVDCRNIA1), anaerobic-related (ANAERO1/2/3CONSENSUS),

water stress (MYCATERD1 and MYBATRD22), drought (ABREZMRAB28, DRECRTCOREAT, and DRE1COREZMRAB17, etc.), low temperature (LTRE-COREATCOR15), hypoxic (CURECORECR), etc. Similarly, a large number of *cis*-elements associated with biotic stress were found in the promoter regions of most members, for example, pathogen-response (SEBFCONS-STPR10A and GCCCORE) and pathogenesis-Related (MYB1LEPR). Furthermore, a large number of phyto-hormone-responsive elements have been identified in the promoter regions of *BpCCO*s, for instance, abscisic acid response elements (SBOXATRBCS, MYB1AT, 2SSEEDPROTBANAPA, etc.), auxin response elements (CATATGGMSAUR, ARFAT, AUXREPSIAA4, etc.), gibberellin response element (GADOWNAT, CAREOS-REP1, GAREAT, etc.), jasmonic acid response elements (T/GBOXATPIN2 and GCCCORE), and ethylene response elements (ERELEE4 and LECPLEACS2), suggesting that *BpCCO*s may be involved in multiple biological processes.

Quantitatively, the most frequently abiotic stress response elements were ABA response elements, with a total of 593. It was followed by anaerobic-related

Fig. 4 Collinearity analysis of the *CCO* genes from *Betula platyphylla* and five other species. The *CCO* collinear genes are connected with a red line, while other collinear genes are connected with gray line

elements (400) and JA-responsive elements (354). It suggests that *BpCCO*s are likely to show strong responses to these three stresses. On the other hand, *BpCCO*s also contain a large number of biotic stress response elements, 502 in total, which suggests that *BpCCO*s may also play important roles in the biotic stress response.

GO enrichment of BpCCOs

The results of the enrichment analysis showed *BpCCO*s were classified into three main categories: biological process, cellular component, and molecular function (Fig. 7). In the biological process category, the most abundant GO term was carotene catabolic process, which included the genes *BpCCD1.1*, *BpNCED5.1*, *BpCCD8.1*, *BpCCD10.1* and *BpCCD7.1*. In the Molecular Function category, oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two oxygen atoms, metal ion binding, and carotenoid

dioxygenase activity were the top three GO terms in terms of abundance. In the Cellular Component category, the GO term chloroplast was enriched with the highest abundance.

miRNAs that target *BpCCOs*

A total of 114 miRNA-*BpCCO* target pairs were obtained, and subsequently a network of miRNA-*BpCCO*s was generated (Fig. 8). It was found that among all the predicted pairs, miR156-related pairs took the largest proportion (number is 11). The related genes were *BpNCED3* and *BpCCD10.1*. The prediction results showed that the binding sites of these miR156s for *BpCCD10.1* were the same (Supplementary Table 9), suggesting that the miR156s were able to specifically recognize specific region in the sequence, which implies that this prediction model is of high confidence level.
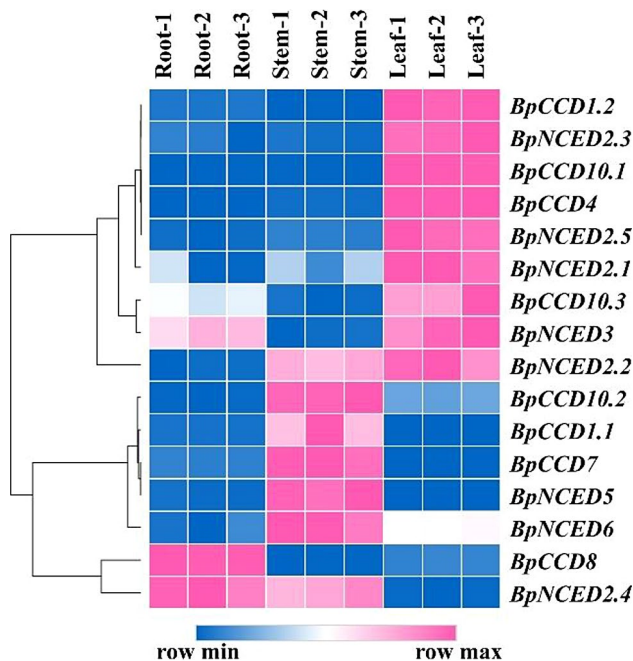
Fig. 5 Tissue-specific expression analysis of *BpCCOs*. The analysis was performed using roots, stems and leaves of birch as plant materials. Relative expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method with root expression of *BpCCD10.3* as a control. Primers are shown in Supplementary Table 8. The *18 S* rRNA gene was used as an internal reference

Responses of BpCCOs upon abiotic stresses

The results of *cis*-element analysis indicated that the promoters of *BpCCOs* were enriched with a variety of stress-responsive *cis*-acting elements, and these genes may respond to a wide range of abiotic stresses. Therefore, we subjected birch to ABA, PEG, salt and cold treatments and determined the relative expression levels of each member at different treatment times using qRT-PCR (Supplementary Tables 4, 5, 6, 7).

Upon ABA treatment (Fig. 9), the expression levels of all *BpCCOs* were significantly changed except *BpCCD8*. A total of five *BpCCOs* were consistently up-regulated with the ABA treatment, namely *BpCCD1.1*, *BpCCD10.3*, *BpNCED2.2*, *BpNCED5* and *BpNCED6*. Six *BpCCOs* were consistently down-regulated in expression, namely *BpCCD1.2*, *BpCCD4*, *BpCCD10.1*, *BpCCD10.2*, *BpNCED2.3* and *BpNCED2.5*. In addition, three *BpCCOs* (*BpCCD7*, *BpNCED2.1* and *BpNCED3*) showed up-regulation from 0 to 24 h and down-regulation at 48 h. *BpNCED2.4*, on the other hand, showed down-regulation from 0 to 12 h followed by increased expression. Overall, most of the *BpCCOs* showed significant response to ABA, however, the response pattern was different for different members.

Upon PEG treatment (Figs. 10), 12 *BpCCOs* showed significantly up-regulated expression at one time point, except *BpCCD8*, *BpCCD10.3*, *BpNCED2.4* and *BpNCED6*, which showed a tendency to first significantly up-regulate and then significantly down-regulate their expression, which suggests that PEG treatment activates most *BpCCOs*.

The results of salt treatment assay (Fig. 11) showed that a total of eight *BpCCOs* were activated and consistently up-regulated in expression by salt treatment, namely *BpCCD7*, *BpCCD10.3*, *BpNCED2.1*, *BpNCED2.3*, *BpNCED2.4*, *BpNCED2.5*, *BpNCED3* and *BpNCED5*. Six *BpCCOs* firstly showed significant activation and then significant decrease in expression, namely *BpCCD1.1*, *BpCCD8*, *BpCCD10.1*, *BpCCD10.2*, *BpNCED2.2* and *BpNCED6*. In addition, two *BpCCDs* (*BpCCD1.2* and *BpCCD4*) were significantly down-regulated after salt treatment. Overall, most of the *BpCCOs* were activated in expression 0–6 h after salt stress.

The results of cold treatment (Fig. 12) showed that a total of eight *BpCCOs* were significantly up-regulated 3 h after cold treatment, namely *BpCCD1.2*, *BpCCD4*, *BpCCD10.3*, *BpNCED2.1*, *BpNCED2.2*, *BpNCED2.4*, *BpNCED5* and *BpNCED6*. Four *BpCCOs* were significantly down-regulated, namely *BpCCD7*, *BpNCED2.3*, *BpNCED2.5* and *BpNCED3*. The expression levels of other genes showed no significant changes. Collectively, it indicates that cold stress can cause significant changes in the expression of most *BpCCOs*.

Protein-protein interaction network of BpCCOs

To better understand the putative functions and interactions of BpCCOs, we first constructed an interaction network model of BpCCO proteins using STRING (Fig. 13). Subsequent functional annotation of all genes in the network was performed using the homologous gene identification method, which showed that only a few genes could be annotated, including *HEC1* (*bHLH*), *BPChr06G29579* (*bHLH*), *ATABA1* (*BPChr06G16528*), *ATRPB15.9* (*BPChr03G15114*), *ATTBP2* (*BPChr07G24868*), *ATVAMP714* (*BPChr10G15896*), *CYP90D1* (*BPChr06G19200*), *MAX2* (*BPChr11G15589*), *RACK1A* (*BPChr12G02842*) and *ZDS* (*BPChr05G31567*). The proteins encoded by these genes may interact with CCOs and jointly play important roles in certain biological processes.

.

Transcriptional activation activity of BpCCD4 and BpNCED5

The transcription activation activity of BpCCD4 and BpNCED5 was investigated by yeast transformation. As shown in Fig. 14, the yeast transformed with positive control can grow both on the SD/-Trp and SD/-Trp/-His/-Ade medium. The yeast transformed with BpCCD4, BpNCED5, respectively, can grow on the SD/-Trp medium while cannot grow on SD/-Trp/-His/-Ade
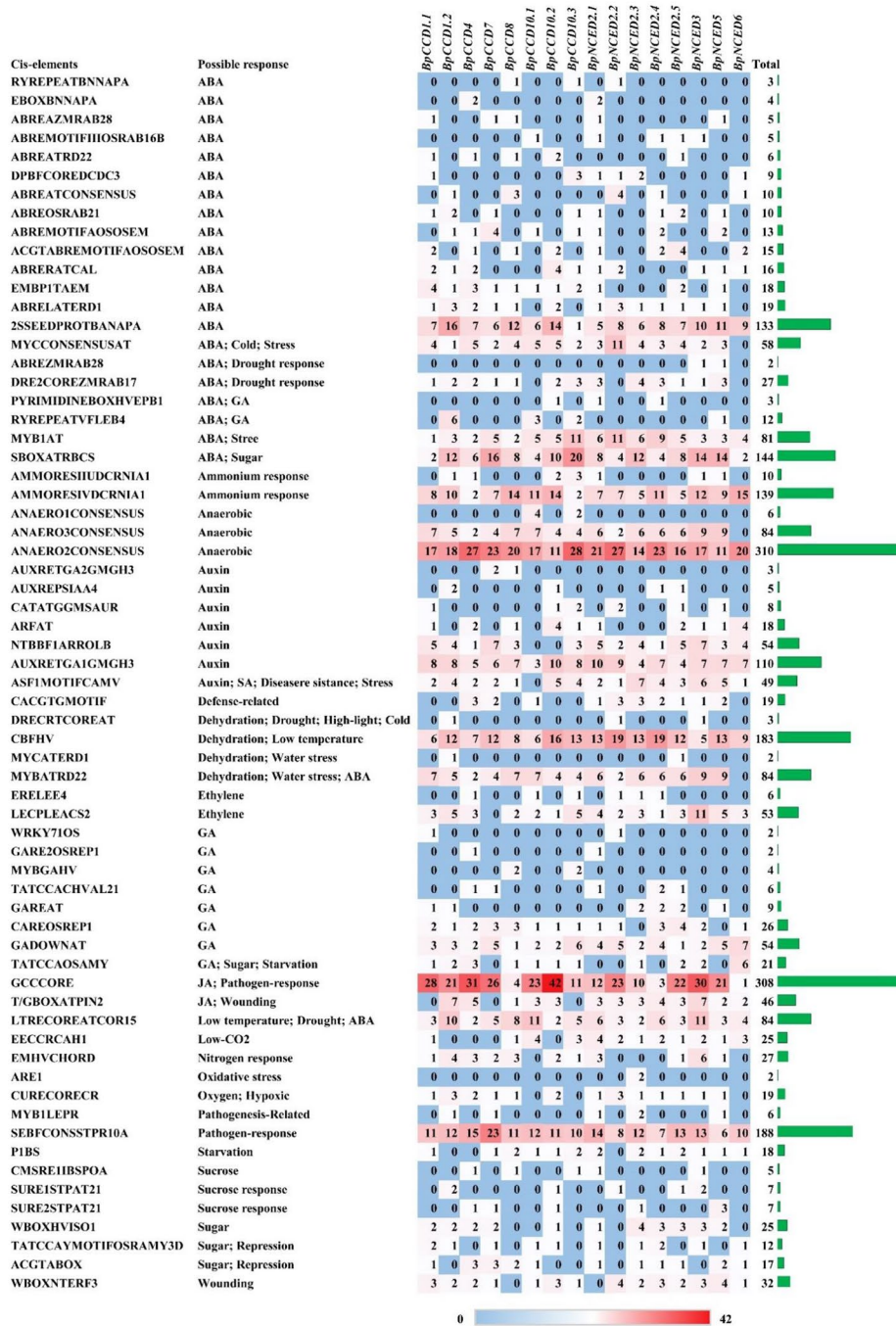
Fig. 6 *Cis*-elements analysis of birch *CCO* genes promoters. The PLACE software was used to analyze the 2000 bp DNA sequence upstream the transcription starts site (TSS) of *BpCCO*s. The heat map represents the number of elements, and the bar graph represents the total number of elements

medium, which indicates the successful transformation and no transcription activation activity.

Interaction between BpCCD4 and BpABA1, BpNCED5 and BpVAMP714

To further verify the protein-protein interaction network, yeast two hybrid assay was conducted with two protein combinations, BpCCD4/BpABA1 and BpNCED5/BpVAMP714. As shown in Fig. 15, the results showed that the yeast transformed with two vector combinations as well as positive control and negative control, respectively, can grow on SD/-Trp/-Leu medium, which indicates successful co-transformation. The yeast transformed with the two vector combinations, respectively and positive control, can grow on the SD/-Trp/-Leu/-His/-Ade medium, while negative control cannot. It
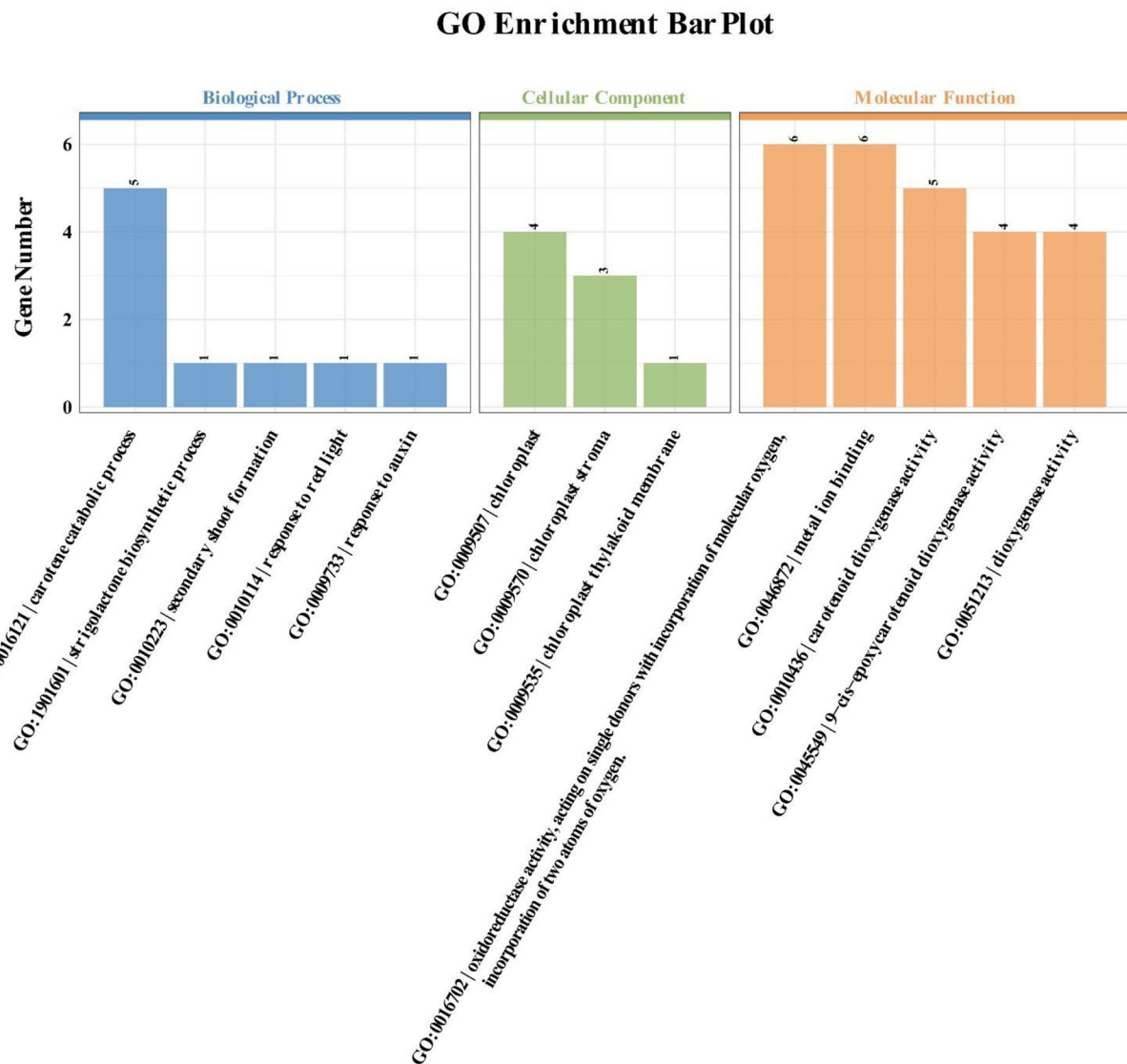
GO Enrichment Bar Plot

Fig. 7 GO enrichment analysis of *BpCCOs*. The enrichment results were classified into three categories, Biological Process, Cellular Component, and Molecular Function

proved the protein-protein interaction between BpCCD4 and BpABA1, as well as BpNCED5 and BpVAMP714.

Function of *BpCCD7, BpNCED2.5* and *BpNCED5* in salt tolerance

According to the expressional responses of *BpCCOs* upon salt treatment, *BpCCD7*, *BpNCED2.5* and *BpNCED5* were selected for further verification of their function in salt tolerance. As shown in Fig. 16, the growth traces of all experimental groups and control weakened with increasing salt concentration. The yeast transformed with pYES2-NTB-BpCCD7, pYES2-NTB-BpNCED2.5 and pYES2-NTB-BpNCED5 can grow on SD/-Ura media with 2.0 M Nacl while negative control cannot. It proved the

positive effects of *BpCCD7*, *BpNCED2.5* and *BpNCED7* in salt tolerance.

Discussion

To date, CCO gene family members have been identified in a genome-wide range in several plant species, with 13 in banana [47], 10 in *Liriodendron chinese* [48], 10 in cucumber [49], 11 in pepper [50] and 47 in *Saccharum spontaneum* [18]. Here, the BpCCO gene family members were identified from a genome-wide range, and their expressional responses upon different abiotic stresses were explored.

In this study, a total of 16 *BpCCO* genes were identified from birch genome, including 8 *BpCCD*s and 8 *BpNCED*s. *BpCCO*s can be classified into two
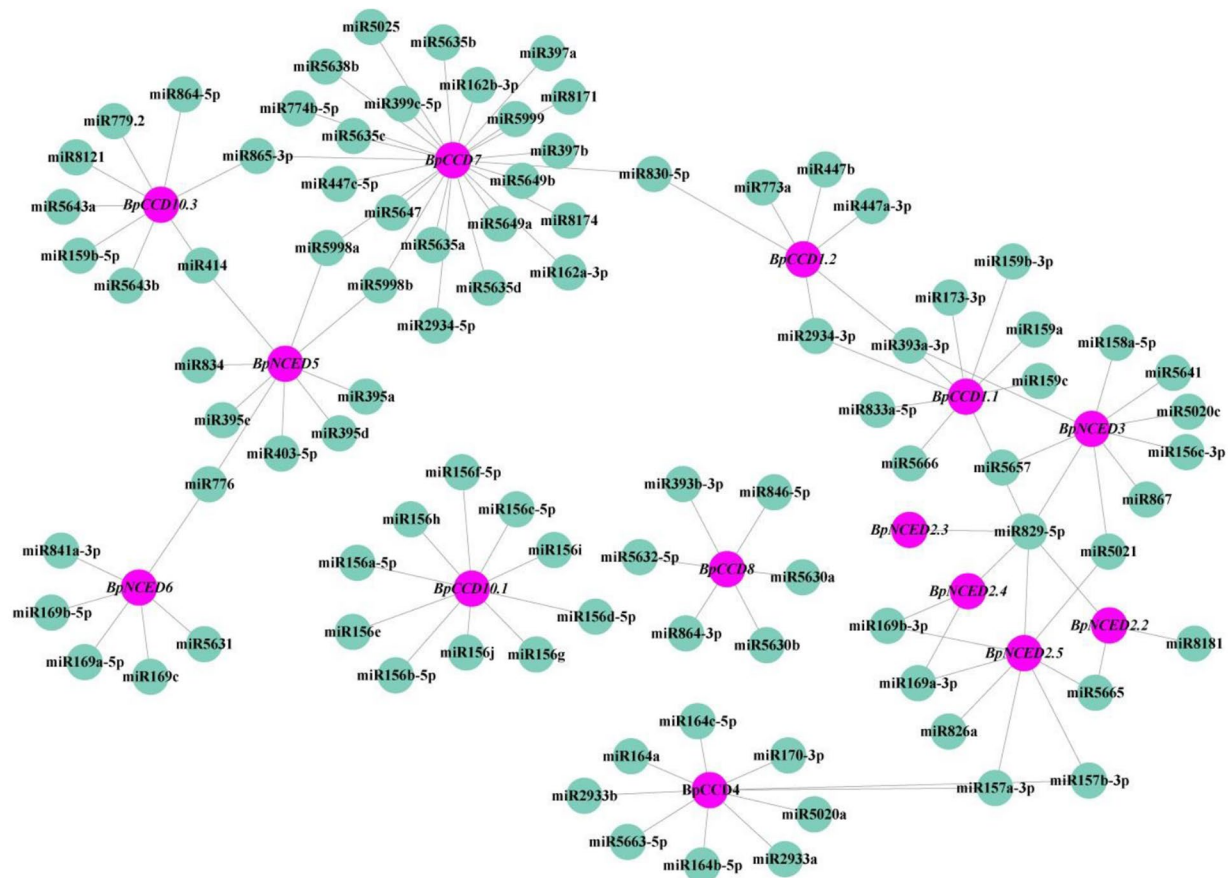
Fig. 8 miRNAs potentially targeting *BpCCO*s. Purple dots represent *BpCCO*s. Green dots represent miRNAs. The targeting relationships were predicted using psRNATarget online software

subfamilies, namely *BpCCD*s (II-VI) and *BpNCED*s (I). The phylogenetic analysis showed that the CCO family members from birch, poplar and *Arabidopsis* can be classified into six subgroups. It was found that all the subgroups, except subgroup VI, contained *CCO*s from all the three species, suggesting possible similarity in the evolutionary patterns of *CCO*s in herbaceous and woody plants. Furthermore, the phylogenetic tree also showed that all *BpCCO*s except *BpCCD1.1* and *BpCCD1.2* were on the same branch with *PtrCCO*s, suggesting that the *CCO*s in birch and poplar may share closer evolutionary relationships. Consequently, it can be inferred that *CCO*s have diverged with the evolution of the herbaceous and woody plant. In addition, the exon-intron patterns of *BpCCO*s exhibited similarity within the same subfamily. The motif analysis showed that only few genes in the adjacent branches had similar motifs, such as *BpNCED2.5*, *BpNCED3* and *BpNCED5*; and *BpCCD1.1* and *BpCCD1.2*. The other genes differed greatly in the type of motifs, suggesting that there are large functional differences between CCO members which may have emerged during evolution. *BpCCO*s had multiple colinear genes across species, which reached the most in

Gossypium hirsutum (*Gh*), suggesting that *BpCCO*s have evolved to be more closely related to *GhCCO*s. In addition, *BpCCO*s had the lowest number of colinear genes in *O. sativa* (4 pairs), which was similar to the results of the previously reported study on poplar *CCO*s (1 pair) [14], probably due to the fact that rice is a monocotyledon and distantly related to other dicotyledons. According to the results of GO enrichment analysis, BpCCOs can be enriched in three categories, biological process, cellular component, and molecular function, in which carotenoid is an importantpart. Carotenoid is known to be one of the most important antioxidants in living organisms [14, 51]. These *BpCCO*s were enriched in relevant GO terms involved in carotenoid synthesis or metabolism, suggesting that these genes have potential ROS scavenging functions, implying that they may play an important role in plants under biotic/abiotic stress. miRNAs play important roles in plant growth and development and are involved in the post-transcriptional regulation of genes by binding to specific sequences. In recent years, miRNA has been proved to be involved in plant abiotic stress response [52, 53]. The results showed that miR156 participates the most in the miRNA-gene regulation. miR156
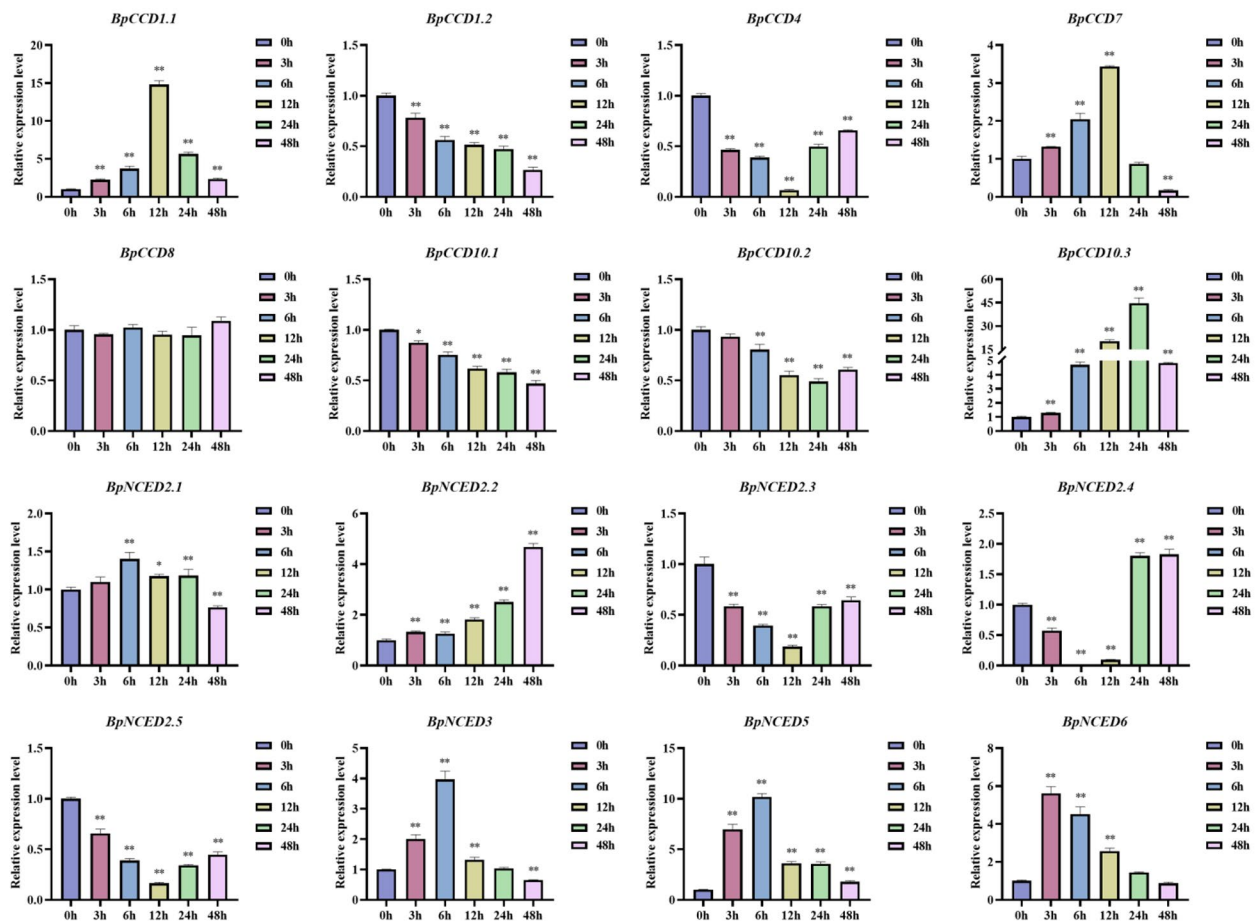
Fig. 9 Expression patterns of *BpCCO*s in response to 100 μM ABA treatment. X-axis shows treatment time point, Y-axis represents relative expression level. The data was processed using the $2^{-\Delta\Delta Ct}$ method. Gene expression at 0 h was set to 1 and expression in the other time points was relative to it. (*t* test, * $p < 0.05$, ** $p < 0.01$)

is one of the largest gene families in plants and plays important roles in plant defense against biotic/abiotic stresses. For example, MicroRNA156 is able to improve drought tolerance in *Medicago sativa* [54]. miR156 is able to influence salt tolerance in apple by regulating downstream gene expression [55]. It is therefore hypothesized that miR156s are likely to be involved in plant stress tolerance processes through the regulation of *BpNCED3* and *BpCCD10.1*, and thus in plant stress tolerance. In addition, five miR169 family members were predicted to potentially regulate the expression of *BpNCED2.5*, *BpNCED2.4* and *BpNCED6*. miR169s also possess important functions in plant stress tolerance, such as most of the poplar miR169s can respond to ABA and salt stress. In *Arabidopsis*, it was found that plants were more sensitive to drought stress after overexpression of miR169a [56]. zma-miR169 was reported to respond positively to salt stress in maize leaves [57]. In this study, miR169s targets were found to be *BpNCED2.5*, *BpNCED2.4* and *BpNCED6*, suggesting that these genes may be regulated

by miR169s and thus involved in abiotic stresses such as drought and salt. In addition, there are a number of miRNAs predicted to bind *BpCCO*s, such as miR159s, miR393s and so on, all of which may be involved in plant stress tolerance processes by regulating the expression of *BpCCO*s.

The expression patterns of *BpCCD*s and *BpNCED*s in different plant tissues, including root, stem and leaf, were investigated by using qRT-PCR. The results indicated a clear divergence between the expression levels of these genes, even those with closer evolutionary relationships. For instance, *BpCCD1.1* and *BpCCD1.2* shared close evolutionary relationship. However, *BpCCD1.2* had high expression levels over all tissues, especially leaf, while *BpCCD1.1* was not so abundantly expressed over all tissues and the highest in stem. According to Auldridge et al. and Simkin et al. [58, 59], *AtCCD1* and *PhCCD1* were both expressed in high levels over all tissues. *BpCCD1.2* shared a similar expression pattern with these two genes, which indicates *BpCCD1.2* may play an important role
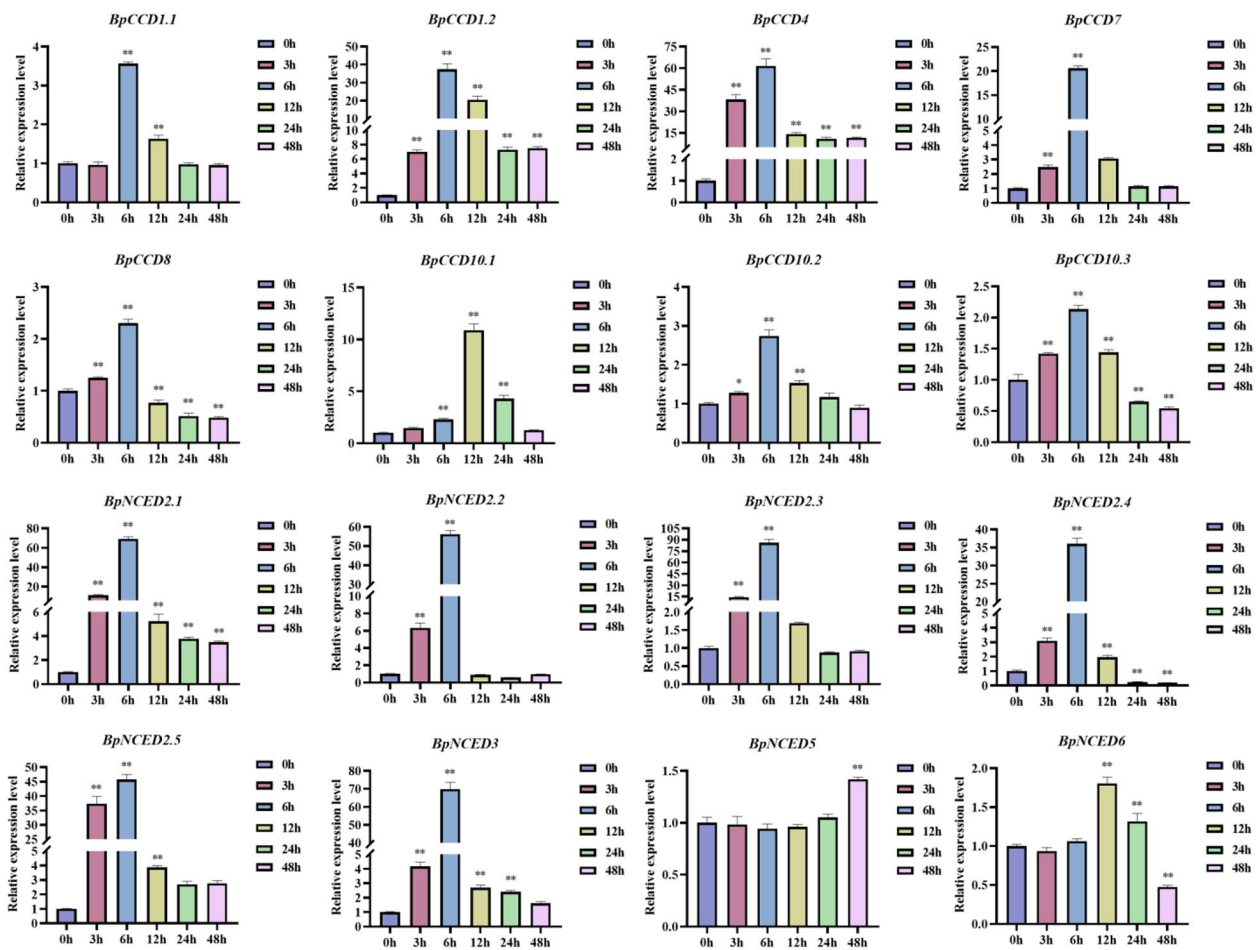
Yu *et al. BMC Genomics*   (2024) 25:872

Page 15 of 21

Fig. 10 Expression patterns of *BpCCO*s upon PEG6000 treatment. X-axis shows treatment time point, Y-axis represents relative expression level. The data was processed using the $2^{-\Delta\Delta Ct}$ method. Gene expression at 0 h was set to 1 and expression in the other time points was relative to it. (*t* test, * $p < 0.05$, ** $p < 0.01$)

in certain metabolisms. *CCD1* gene in poplar was also the most expressed in leaf [14], which implies the similar function between *BpCCD1* and *PtCCD1*. *BpCCD4* and *BpCCD1.2* shared similar expression patterns with approximately half of *BpNCED*s. These genes were the highest expressed in leaf, which indicated the important functions of them in plant leaf. Previous evidence has proven that *AtCCD7* had the highest expression level in root, and the expression levels of *PtCCD7* were low over all tissues [14]. As showed in Fig. 1, *BpCCD7* had close evolutionary relationship with *PtCCD7*. *BpCCD7* was also expressed at low levels over all tissues, which may indicate certain functional similarity between these two genes. It may also imply the different functions of *CCD7* in herbaceous and woody plants. *NCED*s are the rate limiting enzymes for the biosynthesis of ABA, a key signal molecule in the growth and development and stress response of plants [60, 61]. *NCED*s participate in the regulation of endogenous ABA and thereby function in

stress response. In *Arabidopsis*, *AtNCED3* was induced by drought stress and regulated the drought stress response of plants by altering the transpiration rate of leaf and controlling the level of endogenous ABA [62]. In this study, *BpNCED3* was also induced by ABA and drought treatment, as well as salt and cold stress, which can also cause osmotic harm. These evidences indicate the potential important roles this gene may play in the stress response of plants. *AtNCED6* and *AtNCED9* participated in the biosynthesis of ABA in seed germination [63]. Upon abiotic stress, *NCED*s in *Crocus sativus* were closely related to the content of ABA [64]. In this study, *BpNCED3, -5, -6* and *AtNCED3, -6, -9* had similar motif and gene structures, which indicates that *BpNCED3, -5, -6* may function in the accumulation of ABA and abiotic stress response of plants. *BpCCD4* shared similar tissue-specific expression pattern with most of *BpNCED*s. However, their expression patterns upon abiotic stress were different. It indicates that the genes with
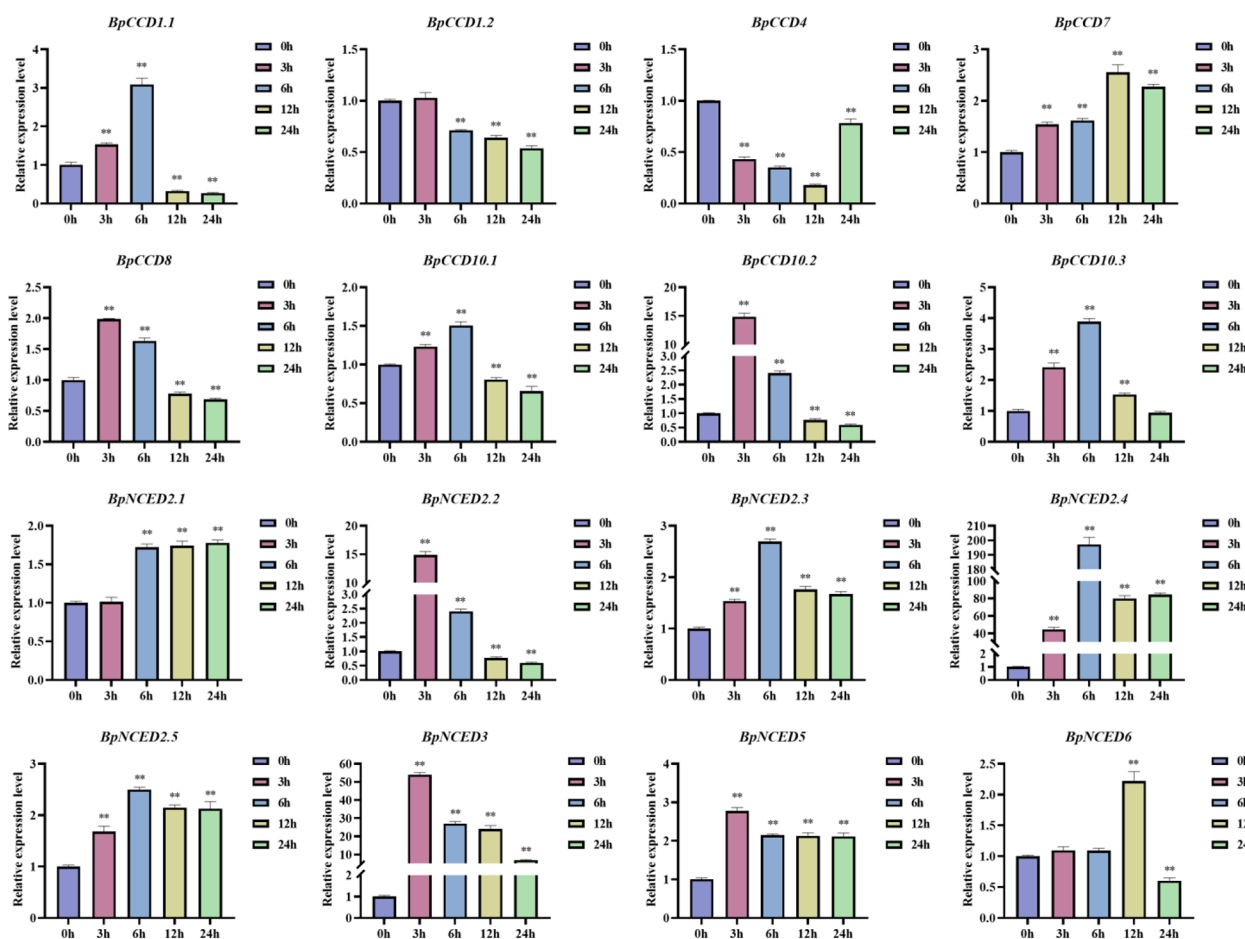
Fig. 11 Expression patterns of *BpCCO*s in response to salt treatment. X-axis shows treatment time point, Y-axis represents relative expression level. The data was processed using the $2^{-\Delta\Delta Ct}$ method. Gene expression in 0 h was set to 1 and expression in the other time points was relative to it. (*t* test, * $p < 0.05$, ** $p < 0.01$)

Fig. 12 Expression patterns of *BpCCO*s in response to cold treatment. X-axis shows treatment time point, Y-axis represents relative expression level. The data was processed using the $2^{-\Delta\Delta Ct}$ method. Gene expression in 0 h was set to 1 and expression in the other time points was relative to it. (*t* test, * $p < 0.05$, ** $p < 0.01$)
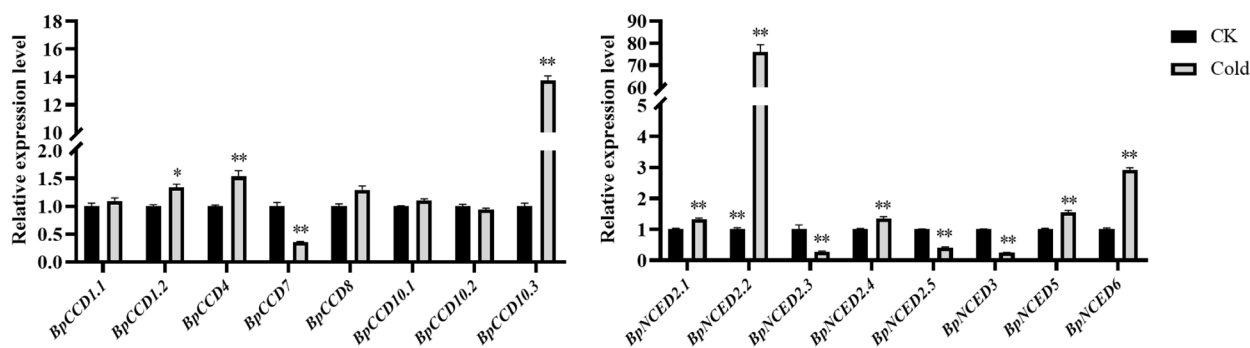
similar tissue-specific expression patterns probably do not response to abiotic stress synchronously. *BpNCED*s seem to play more significant roles in the abiotic stress of plants than *BpCCD*s according to the expression

analysis under abiotic stresses. Although *CCD4* had high expression levels in various plant species, the evidence of its functions in stress response remains insufficient. More studies on *CCD4* focused mainly on fruit and floral
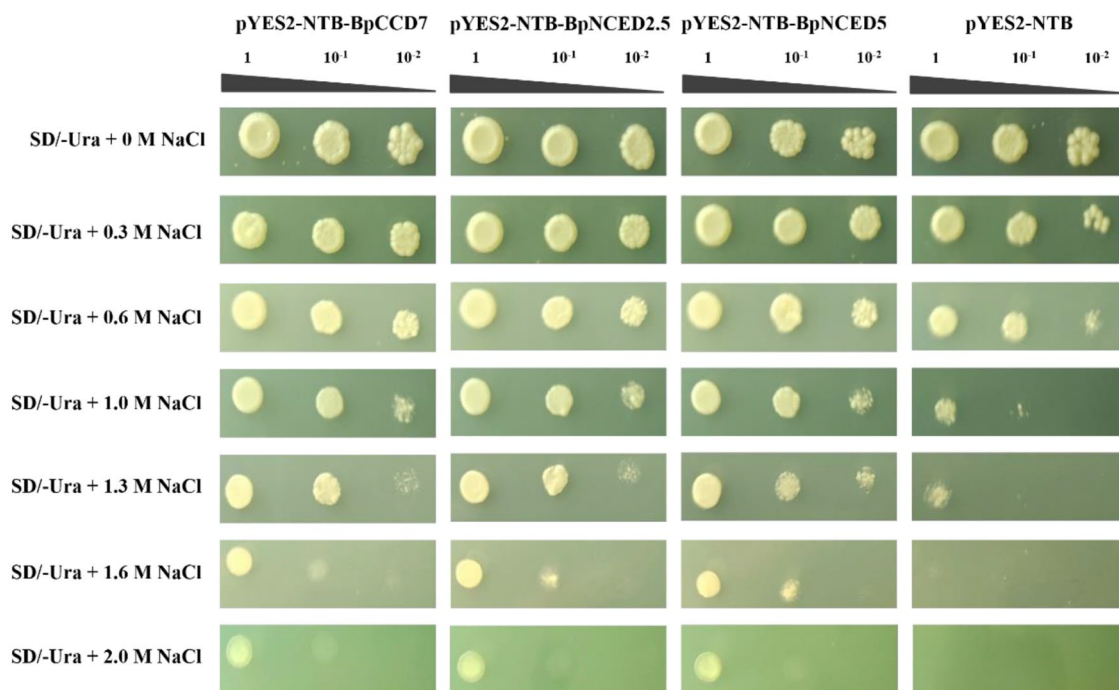
Fig. 13 Protein interaction network analysis of BpCCOs. Interaction networks were constructed by using STRING database (https://string-db.org/, accessed on 9 Sep. 2023). The purple node represents BpCCOs and the blue-green node represents proteins that can interact with BpCCOs. Node size represents the connectivity of the protein

Fig. 14 Transcriptional activation activity of BpCCD4 and BpNCED5. Positive control is pGBKT7-53/pGADT7-T. Negative control is pGBKT7. SD means nutrition-deprived yeast medium

Fig. 15 Protein-protein interaction between BpCCD4 and BpABA1, as well as BpNCED5 and BpVAMP714. Positive control is pGBKT7-53/pGADT7-T. Negative control is pGBKT7-LAM/pGBKT7-T. SD means nutrition-deprived yeast medium

organ. For instance, the *Arabidopsis* plant overexpressing *AtCCD4* exhibited lower contents of β-carotene and lutein and more β-violone accumulation [65]. Function loss of *CCD4* changed the petal color of azalea from yellow to white [66], as well as that of *Eustoma grandiflorum*

Fig. 16 Functions of *BpCCD7*, *BpNCED2.5* and *BpNCED5* in salt tolerance. Negative control is pYES2-NTB. SD means nutrition-deprived yeast medium. 1, 10^{-1}, 10^{-2} indicates different dilution of yeast fluids

from light yellow to white [67]. These evidences indicate that *BpCCD4* may function in the biological process of fruit and floral organ. Previous studies proved that *CCD*s play a major role in the biosynthesis and signal transduction of ABA upon stress condition [68, 69]. The functions of *CCD* and *NCED* genes in birch remain unclear. In this case, the *cis*-acting elements in the promoters of *BpCCD*s and *BpNCED*s were analyzed. The expression patterns of *BpCCD*s and *BpNCED*s upon abiotic stress were determined by using qRT-PCR. The results that these genes were responsive to different abiotic stresses to different degrees indicate their important functions in the abiotic stress response of birch. To further investigate their function in abiotic stress response, *BpCCD7*, *BpNCED2.5* and *BpNCED5* was subjected to salt-resistant yeast transformation. The boosted salt tolerance of the yeast transformed with these genes indicates their positive roles in salt response, which provides an insight into the important functions of *BpCCOs* in abiotic stress response.

Protein-protein interaction is an important way of molecular signal transduction. In this case, the protein interaction networks of BpCCOs were constructed. The results of yeast two hybrid assay proved the interaction between BpCCD4 and BpABA1, BpNCED5 and BpVAMP714, which validates the reliability of the prediction result. ATABA1 has the highest connectivity in the network and its function has been reported. The *aba1* mutant has a defect in the biosynthesis of ABA. The *aba1* mutation results in a reduced pool of ABA precursors,

such as violaxanthin and neoxanthin, and affects the oxidative cleavage of epoxy-carotenoids, resulting in reduced ABA synthesis [70, 71]. In our constructed network, the ATABA1 protein may interact with five CCDs (BpCCD1.1, BpCCD1.2, BpCCD4, BpCCD7 and BpCCD8), which may act synergistically with ATABA1 to affect the process of oxidative cleavage of epoxy-carotenoids. ATVAMP714 is predicted to interact with seven CCO proteins (BpCCD4, BpNCED2.2, BpNCED2.3, BpNCED2.4, BpNCED2. 5, BpNCED3 and BpNCED5), and previous studies have shown that *ATVAMP71* and *ATVAMP712* [72, 73], members of the *ATVAMP71* family, are able to participate in the ABA-mediated drought response process in plants, while ATVAMP714 [74], which is structurally related to ATVAMP71 and ATVAMP712, is likely to have a similar function, and this process is likely to involve CCO proteins. Two members of the basic helix-loop-helix (bHLH) transcription factor (TF) gene family, BPChr06G29579 (bHLH) and HEC1 (bHLH), are predicted to potentially interact with CCOs. bHLH is one of the largest TF families in plants, with several members in different species involved in abiotic stress response and phytohormone (ABA, JA, IAA, etc.) signaling pathways [75, 76], biological processes in which CCOs are also likely to be indirectly involved. In addition to the above proteins, CYP90D1 (BPChr06G19200) [77], ZDS (BPChr05G31567) [78, 79], RACK1A (BPChr12G02842) [80], MAX2 (BPChr11G15589) [81, 82], and ATTBP2 (BPChr07G24868) [83, 84] have also

Yu *et al. BMC Genomics*   (2024) 25:872

Page 19 of 21

been reported in other species to be involved in a variety of growth and developmental processes as well as in the stress response of plants, all of which also have the possibility of indirect involvement of CCOs.

The functions of most of *BpCCO*s still await further verification. With the development of molecular biology, bioinformatics and botany, the specific functions of these genes are expected to be determined. Overall, these results provide a reference for the future study on *BpCCO* genes and the exploration of gene resources on stress response of plants.

Conclusions

A total of 16 *CCO* genes were identified from birch genome, and phylogenetic analysis showed that these genes could be classified into two subfamilies or six subgroups. Structural analysis and motif analysis showed that members of the subgroups had similar gene structures, but the motifs were highly variable. Chromosomal localisation results showed that the 16 *BpCCO*s genes were unevenly distributed on eight chromosomes. Colinearity analysis showed that *BpCCO*s had more colinear genes in the dicotyledonous plant *Gossypium hirsutum*. Promoter *cis*-element analysis showed that the promoters of *BpCCO*s contained a large number of abiotic stress response elements. Most of the *BpCCO*s were able to respond to ABA, salt, PEG and cold treatment. In addition, the protein interaction networks were constructed. Several proteins that may interact with BpCCO proteins were identified, which may function together with BpCCOs during plant resistance. Overall, this study provides an important reference for the study of the *CCO*s family and enriches the information related to the birch gene family.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10777-2.

Supplementary Material 1

Author contributions
R.W. designed the research, acquired the funding, and revised the manuscript. J.Y., Y.W. and X.Z. conducted the experiments, analyzed the data. H.B., R.W. and X.Z. drafted the manuscript and provided the plant materials.

Data availability
All data generated/analyzed are included in this manuscript (and supplementary files).

Declarations

References
1. Weedon BCL, Moss GP. ChemInform Abstract: structure and nomenclature. Cheminform 1995, 26(32).
2. Britton G. Structure and nomenclature of carotenoids. Springer Netherlands; 1993.
3. Woitsch S, Römer S. Expression of xanthophyll biosynthetic genes during light-dependent chloroplast differentiation. Plant Physiol. 2003;132(3):1508–17.
4. Bartley GE, Scolnik PA. Plant carotenoids: pigments for photoprotection, visual attraction, and human health. Plant Cell. 1995;7(7):1027–38.
5. Bouvier F, Isner JC, Dogbo O, Camara B. Oxidative tailoring of carotenoids: a prospect towards novel functions in plants. Trends Plant Sci. 2005;10(4):187–94.
6. Bouvier F, Isner JC, Dogbo O, Camara B. Oxidative tailoring of carotenoids: a prospect towards novel functions in plants. Trends Plant ence. 2005;10(4):187–94.
7. Heo J, Kim SH, Lee PC. New insight into the cleavage reaction of Nostoc sp. strain PCC 7120 carotenoid cleavage dioxygenase in natural and nonnatural carotenoids. Appl Environ Microbiol. 2013;79(11):3336–45.
8. Akemi O, Katsuhiko S, Ryutaro A. Yellow Jimba: suppression of carotenoid cleavage dioxygenase (CmCCD4a) expression turns White Chrysanthemum petals Yellow. J Japanese Soc Hortic Sci. 2009;78(4):450–5.
9. Wei Y, Wan H, Wu Z, Wang R, Ruan M, Ye Q, Li Z, Zhou G, Yao Z, Yang Y. A comprehensive analysis of carotenoid cleavage dioxygenases genes in Solanum Lycopersicum. Plant Mol Biology Report. 2016;34(2):512–23.
10. Schwartz SH, Tan BC, Gage DA, Zeevaart JA, McCarty DR. Specific oxidative cleavage of carotenoids by VP14 of maize. Sci (New York NY). 1997;276(5320):1872–4.
11. Tan BC, Schwartz SH, Zeevaart JA, McCarty DR. Genetic control of abscisic acid biosynthesis in maize. Proc Natl Acad Sci U S A. 1997;94(22):12235–40.
12. Mccarty D. Characterization of three members of the Arabidopsis carotenoid cleavage dioxygenase family demonstrates the divergent roles of this multifunctional enzyme family. Plant J 2006, 45.
13. Kloer DP, Schulz GE. Structural and biological aspects of carotenoid cleavage. Cell Mol Life Ences. 2006;63(19–20):2291–303.
14. Wei H, Movahedi A, Liu G, Li Y, Liu S, Yu C, Chen Y, Zhong F, Zhang J. Comprehensive analysis of carotenoid cleavage dioxygenases Gene Family and its expression in response to Abiotic Stress in Poplar. Int J Mol Sci 2022, 23(3).
15. Zhou Q, Li Q, Li P, Zhang S, Liu C, Jin J, Cao P, Yang Y. Carotenoid cleavage dioxygenases: identification, expression, and evolutionary analysis of this Gene Family in Tobacco. Int J Mol Sci 2019, 20(22).
16. Zhao XL, Yang YL, Xia HX, Li Y. Genome-wide analysis of the carotenoid cleavage dioxygenases gene family in Forsythia suspensa: expression profile and cold and drought stress responses. Front Plant Sci. 2022;13:998911.
17. Lashbrooke JG, Young PR, Dockrall SJ, Vasanth K, Vivier MA. Functional characterisation of three members of the Vitis vinifera L. carotenoid cleavage dioxygenase gene family. BMC Plant Biol. 2013;13:156.
18. Su W, Zhang C, Feng J, Feng A, You C, Ren Y, Wang D, Sun T, Su Y, Xu L, et al. Genome-wide identification, characterization and expression analysis of the carotenoid cleavage oxygenase (CCO) gene family in Saccharum. Plant Physiol Biochem. 2021;162:196–210.
19. Yao Y, Jia L, Cheng Y, Ruan M, Ye Q, Wang R, Yao Z, Zhou G, Liu J, Yu J, et al. Evolutionary origin of the carotenoid cleavage oxygenase family in plants and expression of Pepper genes in response to Abiotic stresses. Front Plant Sci. 2021;12:792832.

20. Alder A, Jamil M, Marzorati M, Bruno M, Vermathen M, Bigler P, Ghisla S, Bouwmeester H, Beyer P, Al-Babili S. The path from β-carotene to carlactone, a strigolactone-like plant hormone. Sci (New York NY). 2012;335(6074):1348–51.

21. Booker J, Sieberer T, Wright W, Williamson L, Leyser O. MAX1 encodes a cyto-chrome P450 Family Member that acts downstream of MAX3/4 to produce a carotenoid-derived branch-inhibiting hormone. Dev Cell. 2005;8(3):443–9.

22. Schwartz SH, Qin X, Loewen MC. The biochemical characterization of two carotenoid cleavage enzymes from Arabidopsis indicates that a carotenoid-derived compound inhibits lateral branching. J Biol Chem. 2004;279(45):46940.

23. Seo M. Complex regulation of ABA biosynthesis in plants. Trends Plant Sci. 2002;7(1):41–8.

24. Tan BC, Schwartz SH, Zeevaart JAD, Mccarty DR. Genetic control of abscisic acid biosynthesis in maize. Proc Natl Acad Sci USA 1997.

25. Rui-Kai W, Wang C-E, Fei Y-Y. Jun-Yi, Gai, Tuan-Jie, Zhao: genome-wide iden-tification and transcription analysis of soybean carotenoid oxygenase genes during abiotic stress treatments. Mol Biol Rep 2013.

26. Zhang J, He L, Dong J, Zhao C, Tang R, Jia X. Overexpression of Sweet Potato Carotenoid cleavage dioxygenase 4 (IbCCD4) decreased Salt Tolerance in Arabidopsis thaliana. Int J Mol Sci 2022, 23(17).

27. Cai X, Jiang Z, Tang L, Zhang S, Li X, Wang H, Liu C, Chi J, Zhang X, Zhang J. Genome-wide characterization of carotenoid oxygenase gene family in three cotton species and functional identification of GaNCED3 in drought and salt stress. J Appl Genet. 2021;62(4):527–43.

28. Baba SA, Jain D, Abbas N, Ashraf N. Overexpression of Crocus carotenoid cleavage dioxygenase, CsCCD4b, in Arabidopsis imparts tolerance to dehy-dration, salt and oxidative stresses by modulating ROS machinery. J Plant Physiol. 2015;189:114–25.

29. Kim Y, Hwang I, Jung HJ, Park JI, Kang JG, Nou IS. Genome-wide classification and abiotic stress-responsive expression profiling of Carotenoid Oxygenase genes in Brassica rapa and Brassica oleracea. J Plant Growth Regul 2016.

30. Chen H, Zuo X, Shao H, Fan S, Ma J, Zhang D, Zhao C, Yan X, Liu X, Han M. Genome-wide analysis of carotenoid cleavage oxygenase genes and their responses to various phytohormones and abiotic stresses in apple (Malus domestica). Plant Physiol Biochem Ppb 2017;81.

31. Hu P, Zhang K, Yang C. BpNAC012 positively regulates abiotic stress responses and secondary Wall Biosynthesis. Plant Physiol. 2019;179(2):700–17.

32. Guo H, Wang Y, Wang L, Hu P, Wang Y, Jia Y, Zhang C, Zhang Y, Zhang Y, Wang C, et al. Expression of the MYB transcription factor gene *BplMYB46* affects abi-otic stress tolerance and secondary cell wall deposition in *Betula platyphylla*. Plant Biotechnol J. 2016;15(1):107–21.

33. Wang Z, He Z, Gao C, Wang C, Song X, Wang Y. Phosphorylation of birch BpNAC90 improves the activation of gene expression to confer drought tolerance. Hortic Res 2024, 11(4).

34. Chen S, Wang Y, Yu L, Zheng T, Wang S, Yue Z, Jiang J, Kumari S, Zheng C, Tang H, et al. Genome sequence and evolution of Betula platyphylla. Hortic Res. 2021;8(1):37.

35. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. Nucleic Acids Res. 2018;46(W1):W200–4.

36. Wang Y, Wang R, Yu Y, Gu W, Wang S, Liao S, Xu X, Jiang T, Yao W. Genome-wide analysis of SIMILAR TO RCD ONE (SRO) family revealed their roles in Abiotic Stress in Poplar. 2023, 24(4):4146.

37. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evo-lutionary Genetics Analysis across Computing platforms. Mol Biol Evol. 2018;35(6):1547–9.

38. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. Mol Biol Evol. 2016;33(7):1870–4.

39. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: an integrative Toolkit developed for interactive analyses of big Biological Data. Mol Plant. 2020;13(8):1194–202.

40. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49.

41. Higo K, Ugawa Y, Iwamoto M, Korenaga T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res. 1999;27(1):297–300.

42. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51(D1):D638–46.

43. Majeed A, Mukhtar S. Protein-protein Interaction Network Exploration using Cytoscape. Methods Mol Biology (Clifton NJ). 2023;2690:419–27.

44. Chen S, Wang Y, Yu L, Zheng T, Wang S, Yue Z, Jiang J, Kumari S, Zheng C, Tang H et al. Genome sequence and evolution of *Betula platyphylla*. Hortic Res 2021, 8(37).

45. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods (San Diego Calif). 2001;25(4):402–8.

46. Tan BC, Joseph LM, Deng WT, Liu L, Li QB, Cline K, McCarty DR. Molecular characterization of the Arabidopsis 9-cis epoxycarotenoid dioxygenase gene family. Plant J. 2003;35(1):44–56.

47. Zeng L, Zeng L, Wang Y, Xie Z, Zhao M, Chen J, Ye X, Tie W, Li M, Shang S, et al. Identification and expression of the CCO family during development, ripen-ing and stress response in banana. Genetica. 2023;151(2):87–96.

48. Xue G, Hu L, Zhu L, Chen Y, Qiu C, Fan R, Ma X, Cao Z, Chen J, Shi J et al. Genome-wide identification and expression analysis of CCO Gene Family in Liriodendron chinense. Plants 2023, 12(10).

49. Akram J, Siddique R, Shafiq M, Tabassum B, Manzoor MT, Javed MA, Anwar S, Nisa BU, Saleem MH, Javed B et al. Genome-wide identification of CCO gene family in cucumber (Cucumis sativus) and its comparative analysis with A. Thaliana. BMC Plant Biol 2023, 23(1).

50. Zhang XH, Liu HQ, Guo QW, Zheng CF, Li CS, Xiang XM, Zhao DF, Liu J, Luo J, Zhao DK et al. Genome-wide identification, phylogenetic relationships, and expression analysis of the carotenoid cleavage oxygenase gene family in pepper. Genet Mol Res 2016, 15(4).

51. Rao Y, Peng T, Xue S. Mechanisms of plant saline-alkaline tolerance. J Plant Physiol. 2023;281:153916.

52. Huang JH, Zhang LY, Lin XJ, Gao Y, Zhang J, Huang WL, Zhao D, Ferrarezi RS, Fan GC, Chen LS. *CsiLAC4* modulates boron flow in *Arabidopsis* and *Citrus* via high-boron-dependent lignification of cell walls. New Phytol. 2021;233(3):1257–73.

53. Jiang Y, Wu X, Shi M, Yu J, Guo C. The miR159-MYB33-ABI5 module regulates seed germination in *Arabidopsis*. Physiol Plant 2022, 174(2).

54. Arshad M, Feyissa BA, Amyot L, Aung B, Hannoufa A. MicroRNA156 improves drought stress tolerance in alfalfa (Medicago sativa) by silencing SPL13. Plant Sci. 2017;258:122–36.

55. Ma Y, Xue H, Zhang F, Jiang Q, Yang S, Yue P, Wang F, Zhang Y, Li L, He P, et al. The miR156/SPL module regulates apple salt stress tolerance by activating MdWRKY100 expression. Plant Biotechnol J. 2021;19(2):311–23.

56. Li WX, Oono Y, Zhu J, He XJ, Wu JM, Iida K, Lu XY, Cui X, Jin H, Zhu JK. The Arabidopsis NFYA5 transcription factor is regulated transcription-ally and posttranscriptionally to promote drought resistance. Plant Cell. 2008;20(8):2238–51.

57. Luan M, Xu M, Lu Y, Zhang L, Fan Y, Wang L. Expression of zma-miR169 miR-NAs and their target ZmNF-YA genes in response to abiotic stress in maize leaves. Gene. 2015;555(2):178–85.

58. Auldridge ME, Block A, Vogel JT, Dabney-Smith C, Mila I, Bouzayen M, Magallanes-Lundback M, DellaPenna D, McCarty DR, Klee HJ. Characterization of three members of the Arabidopsis carotenoid cleavage dioxygenase fam-ily demonstrates the divergent roles of this multifunctional enzyme family. Plant J. 2006;45(6):982–93.

59. Simkin AJ, Underwood BA, Auldridge M, Loucas HM, Shibuya K, Schmelz E, Clark DG, Klee HJ. Circadian regulation of the PhCCD1 carotenoid cleavage dioxygenase controls Emission of β-Ionone, a Fragrance Volatile of Petunia flowers. Plant Physiol. 2004;136(3):3504–14.

60. Pei X, Wang X, Fu G, Chen B, Nazir MF, Pan Z, He S, Du X. Identification and functional analysis of 9-cis-epoxy carotenoid dioxygenase (NCED) homologs in G. Hirsutum. Int J Biol Macromol. 2021;182:298–310.

61. Zhang J, Zhang P, Huo X, Gao Y, Chen Y, Song Z, Wang F, Zhang J. Compara-tive phenotypic and transcriptomic analysis reveals key responses of Upland cotton to salinity stress during postgermination. Front Plant Sci 2021, 12.

62. Iuchi S, Kobayashi M, Taji T, Naramoto M, Seki M, Kato T, Tabata S, Kakubari Y, Yamaguchi-Shinozaki K, Shinozaki K. Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in Arabidopsis. Plant J. 2001;27(4):325–33.

63. Lefebvre V, North H, Frey A, Sotta B, Seo M, Okamoto M, Nambara E, Marion-Poll A. Functional analysis of Arabidopsis NCED6 and NCED9 genes indicates that ABA synthesized in the endosperm is involved in the induction of seed dormancy. Plant J. 2006;45(3):309–19.

64. Ahrazem O, Rubio-Moraga A, Trapero A, Gomez-Gomez L. Developmen-tal and stress regulation of gene expression for a 9-cis-epoxycarotenoid dioxygenase, CstNCED, isolated from Crocus sativus stigmas. J Exp Bot. 2011;63(2):681–94.

65. Song M-H, Lim S-H, Kim JK, Jung ES, John KMM, You M-K, Ahn S-N, Lee CH, Ha S-H. In planta cleavage of carotenoids by Arabidopsis carotenoid cleavage dioxygenase 4 in transgenic rice plants. Plant Biotechnol Rep. 2016;10(5):291–300.

66. Ureshino K, Nakayama M, Miyajima I. Contribution made by the carotenoid cleavage dioxygenase 4 gene to yellow colour fade in azalea petals. Euphytica. 2015;207(2):401–17.

67. Liu H, Kishimoto S, Yamamizo C, Fukuta N, Ohmiya A, Debener T. Carotenoid accumulations and carotenogenic gene expressions in the petals of Eustoma grandiflorum. Plant Breeding. 2013;132(4):417–22.

68. Alder A, Jamil M, Marzorati M, Bruno M, Vermathen M, Bigler P, Ghisla S, Bouwmeester H, Beyer P, Al-Babili S. The path from β-Carotene to Carlactone, a Strigolactone-Like Plant hormone. Science. 2012;335(6074):1348–51.

69. Kim Y, Hwang I, Jung H-J, Park J-I, Kang J-G, Nou I-S. Genome-wide classification and abiotic stress-responsive expression profiling of Carotenoid Oxygenase genes in Brassica rapa and Brassica oleracea. J Plant Growth Regul. 2015;35(1):202–14.

70. Rock CD, Zeevaart JA. The aba mutant of Arabidopsis thaliana is impaired in epoxy-carotenoid biosynthesis. Proc Natl Acad Sci U S A. 1991;88(17):7496–9.

71. Choi J, Kim H, Suh MC. Disruption of the ABA1 encoding zeaxanthin epoxidase caused defective suberin layers in Arabidopsis seed coats. Front Plant Sci. 2023;14:1156356.

72. Leshem Y, Golani Y, Kaye Y, Levine A. Reduced expression of the v-SNAREs AtVAMP71/AtVAMP7C gene family in Arabidopsis reduces drought tolerance by suppression of abscisic acid-dependent stomatal closure. J Exp Bot. 2010;61(10):2615–22.

73. Xue Y, Yang Y, Yang Z, Wang X, Guo Y. VAMP711 is required for abscisic acid-mediated inhibition of plasma membrane H(+)-ATPase activity. Plant Physiol. 2018;178(3):1332–43.

74. Gu X, Fonseka K, Agneessens J, Casson SA, Smertenko A, Guo G, Topping JF, Hussey PJ, Lindsey K. The Arabidopsis R-SNARE VAMP714 is essential for polarisation of PIN proteins and auxin responses. New Phytol. 2021;230(2):550–66.

75. Hao Y, Zong X, Ren P, Qian Y, Fu A. Basic Helix-Loop-Helix (bHLH) transcription factors regulate a wide range of functions in Arabidopsis. Int J Mol Sci 2021, 22(13).

76. Zhao K, Li S, Yao W, Zhou B, Li R, Jiang T. Characterization of the basic helix-loop-helix gene family and its tissue-differential expression in response to salt stress in poplar. PeerJ. 2018;6:e4502.

77. Enoki S, Tanaka K, Moriyama A, Hanya N, Mikami N, Suzuki S. Grape cytochrome P450 CYP90D1 regulates brassinosteroid biosynthesis and increases vegetative growth. Plant Physiol Biochem. 2023;196:993–1001.

78. Domonkos I, Kis M, Gombos Z, Ughy B. Carotenoids, versatile components of oxygenic photosynthesis. Prog Lipid Res. 2013;52(4):539–61.

79. Yang MH, Lu YS, Ho TC, Shen DH, Huang YF, Chuang KP, Yuan CH, Tyan YC. Utilizing Proteomic Approach to analyze potential antioxidant proteins in plant against Irradiation. Antioxid (Basel Switzerland) 2022, 11(12).

80. Rahman MA, Ullah H. Receptor for activated C Kinase1B (RACK1B) delays Salinity-Induced Senescence in Rice leaves by regulating chlorophyll degradation. Plants (Basel Switzerland) 2023, 12(12).

81. Zheng X, Liu F, Yang X, Li W, Chen S, Yue X, Jia Q, Sun X. The MAX2-KAI2 module promotes salicylic acid-mediated immune responses in Arabidopsis. J Integr Plant Biol. 2023;65(6):1566–84.

82. Temmerman A, Marquez-Garcia B, Depuydt S, Bruznican S, De Cuyper C, De Keyser A, Boyer FD, Vereecke D, Struk S, Goormachtig S. MAX2-dependent competence for callus formation and shoot regeneration from Arabidopsis thaliana root explants. J Exp Bot. 2022;73(18):6272–91.

83. Hwang MG, Kim K, Lee WK, Cho MH. AtTBP2 and AtTRP2 in Arabidopsis encode proteins that bind plant telomeric DNA and induce DNA bending in vitro. Mol Genet Genomics. 2005;273(1):66–75.

84. Schrumpfová P, Kuchar M, Miková G, Skrísovská L, Kubicárová T, Fajkus J. Characterization of two Arabidopsis thaliana myb-like proteins showing affinity to telomeric DNA sequence. Genome. 2004;47(2):316–24.

Publisher's note