Article

Genome-Wide Identification and Analysis of Plant Cysteine Oxidase (PCO) Gene Family in *Brassicaceae* and Their Roles in Abiotic Stress Response

Xiaohua Bian [1], Yifan Cao [1], Ximin Zhi [1] and Ni Ma [1,*]

[1]   Oil Crops Research Institute, Chinese Academy of Agricultural Sciences, Wuhan 430062, China

*   Correspondence: mani@caas.cn

Abstract: Plant Cysteine Oxidase (PCO) is a plant O_2-sensing enzyme of catalyzing oxidation of cysteine to Cys-sulfinic acid at the N-termini of target proteins. To better understand the *Brassica napus PCO* gene family, *PCO* genes in *B. napus* and related species were analyzed. In this study, 20, 7 and 8 *PCO* genes were identified in *Brassica napus*, *Brassica rapa* and *Brassica oleracea*, respectively. According to phylogenetic analysis, the *PCOs* were divided into five groups: *PCO1*, *PCO2*, *PCO3*, *PCO4* and *PCO5*. Gene organization and motif distribution analysis suggested that *PCO* gene family was relatively conserved during evolution. Expression analysis showed that *PCO* genes were expressed in different tissues at different developmental stages. Interestingly, most of the *Bna/Br/BoPCO5* members were expressed in leaves, roots, flowers, buds and siliques, suggesting an important role in both vegetative and reproductive development. Expression of *BnaPCO* was induced by various abiotic stress, especially waterlogging stress, which was consistent with the result of cis-element analysis. In this study, the *PCO* gene family of *Brassicaceae* was analyzed for the first time, which contributes to a comprehensive understanding of the origin and evolution of *PCO* genes in *Brassicaceae* and the function of *BnaPCO* in regulating plant responses to abiotic stresses.

Keywords: *Brassica napus*; Plant Cysteine Oxidase (PCO); gene expression; abiotic stress

1. Introduction

Enzymes are the fundamental to all metabolic, catabolic and adaptive biological processes. They have evolved so that their functions meet biological needs, fine-tuning their activities to achieve optimal survival and replication fitness of the organism. Plant Cysteine Oxidases (PCOs) use molecular oxygen atoms to catalyze dioxygenation of Nt-Cys to Cys-sulfinic acid in ERF-VIIs (ETHYLENE RESPONSE FACTOR group VIIs, as well as other substrates VRN2 and ZPR2) [1, 2] for subsequent Nt-arginylation [3, 4]. Meanwhile, ERF-VIIs are known as the regulators of hypoxia-regulated transcriptional reprogramming to adapt the environmental change [5-7]. PCO activity is sensitive to physiologically relevant fluctuations in O_2 availability [8], thus these enzymes can act as plant O_2 sensors with a key role in regulating ERF-VII stability. In other words, as an enzyme, Plant Cysteine Oxidases (PCOs) are a direct link between environmental stimuli and molecular physiological outcomes [9].

In Arabidopsis, AtPCOs catalyze oxidation of cysteine to Cys-sulfinic acid at the N-termini of target proteins, a reaction that co-translational methionine cleavage exposes the N-terminal Cys for oxidation, leading to the degradation of the target protein via the Cys/Arg branch of the N-degron pathway [3, 4, 8]. Oxidized N-terminal Cys residues are substrates for arginyl transferase enzymes, with the arising arginylated N-termini recognized by ubiquitin ligases [10]. Ubiquitination is the signal for protein degradation by the proteasome. It is reported that PCOs isolated from *Marchantia polymorpha* (MpPCO) and *Klebsormidium nitens* (KnPCO) exhibit cysteine dioxygenase activity, indicating that PCO enzymes are conserved in early land plants and algae [11]. The MpPCO was incubated with the N-termini of MpERF-like (a 14-mer peptide representing the Cys-initiating N-terminus of MpERF-like, CRMNKRLGKGETGL), and MpPCO-catalyzed MpERF-like oxidation reached 89.3% after 1h. Meanwhile, homologs of arginyl-tRNA transferase (ATE) and E3 N-recognin,

PROTEOLYSIS (PRT) 6 can be found in *Marchantia polymorpha*, suggesting the conserved way of PCO in catalyzing oxidation of cysteine to Cys-sulfinic acid at the N-termini of target proteins.

As PCOs regulate the target protein levels in N-degron pathway, it is critical to uncover PCOs structures to manipulate their enzyme activities for crop improvement [12]. The structures of AtPCO4_1, AtPCO4_2 (two different structures of AtPCO4 from independent crystallization conditions) and AtPCO5 are resolved to 1.82-, 1.24- and 1.91-Å resolutions [9]. There is a core double-stranded beta-helix (DSBH) supporting three histidine residues to coordinate the active site metal ions involved in catalysis. To recognize the active site, the Tyr_{182}-Ser_{183}-Ser_{184}-Glu_{185}-His_{186}-Asp_{187}-Arg_{188}-His_{189}-Cys_{190} fragment is characterized by targeted mutagenesis. Furthermore, AtPCO4 (H164D) or AtPCO4 (D176N) fails to reconstitute the wild-type phenotype in *4pco* (quadruple *pco1/2/4/5* mutant plants) mutant, while AtPCO4 rescues the phenotype of *4pco* mutant. Therefore, the sites (amino acid 160-190 position) of PCO4 play an important role in the enzyme activity [9, 13, 14].

The enzyme activities of PCOs are involved in plant development and stress responses. Quadruple *pco1/2/4/5* mutant plants, but not triple *pco1/2/4* mutants, in *A. thaliana* generated by T-DNA insertion have previously been shown to cause developmental defects, including retarded growth, leaf serration, and male and female sterility, along with constitutive expression of low-oxygen response genes [2, 4, 9]. Overexpression of *AtPCO1* or *AtPCO2* in Arabidopsis show smaller leaves and decreased biomass [4]. It suggests that appropriate PCOs level or enzyme activity is the fundamental to maintain plant growth and development. The PCOs catalyze the reaction of oxygen with the conserved N-terminal cysteine of ERF-VIIs to form cysteine sulfinic acid triggering degradation under normal condition. PCO activity decreases under hypoxia condition and then the stabilized ERF-VIIs active the expression of anaerobic genes in response to stress [3, 15]. Previous studies have shown that aerobic or hypoxia related genes are down-regulated in either *AtPCO1* or *AtPCO2* over-expressed plants, whereas anaerobic genes are constitutively up-regulated in quadruple *pco1/2/4/5* mutants [10]. Compared with wide type, plants over-expressing *AtPCO1* or *AtPCO2* show sensitive to prolonged submergence stress [4]. Meanwhile, overexpression of the ERF-VII RAP2.12 in *Arabidopsis* results in improved tolerance to submergence and up-regulation of genes associated with the hypoxic response [7, 12]. Therefore, PCOs, as an O_2 sensor, play a critical role in stress response, especially hypoxia stress.

About 7,500 years ago, a natural heterotetraploid *Brassica napus* (AACC, 2n = 38) was formed by the hybridization of *Brassica rapa* (2n = 20, AA) and *Brassica oleracea* (2n = 18, CC) [16-18]. *Brassica napus* is an important oil crop, with the largest planting area as well as total oil production in China. With the changes of the global climate, the cultivation of rapeseed has been effected by various abiotic stress, such as flooding, waterlogging, drought and cold. Research on genetics, evolution and stress response of rapeseed is a good way to improve the quality and yield of rapeseed [19-23]. While *PCOs* are expected to play important and diverse roles in stress responses, detailed genome-wide analysis of *PCO* gene family has not been performed. In this study, we investigated the significant role of *PCO* genes in *Brassicaceae* plants. As a result, we identified 35 *PCO* genes of *Brassica napus*, *Brassica rapa* and *Brassica oleracea* and compared their gene structures, chromosomal locations, evolutionary relationships and expression patterns in different tissues and under different abiotic stress treatments. Therefore, the genome-wide analysis in this study provides a comprehensive understanding of PCOs in development and stress responses in *Brassicaceae* plants.

2. Results

2.1. Identification and Classification of PCO Genes in B. napus, B. oleracea and B. rapa

To identify PCO proteins in B. napus, B. oleracea and B. rapa, we performed a BLASTp search against the annotated proteins of B. napus, B. oleracea and B. rapa in BRAD (http://www.brassicadb.cn/#/BLAST/) using Arabidopsis AtPCO protein (AtPCO1, AtPCO2, AtPCO3, AtPCO4 and AtPCO5) sequences as queries. Sets of 20 (4 BnaPCO1, 2 BnaPCO2, 6 BnaPCO3, 4 BnaPCO4 and 4 BnaPCO5), 8 (2 BoPCO1, 1 BoPCO2, 1 BoPCO3, 2 BoPCO4 and 2 BoPCO5) and 7 (2 BraPCO1, 1

BraPCO2, 1 BraPCO3, 1 BraPCO4 and 2 BraPCO5) PCO proteins were identified in B. napus, B. oleracea and B. rapa, respectively (Table 1). All 35 PCOs in B. napus, B. oleracea and B. rapa encoded the amino acids ranging in length from 82 to 308 with the molecular weight from 8,650 to 34,180 Dalton. Isoelectric point (pI) of these amino acids ranged from 4.39 to 8.78 (Table 1).

Table 1. List of PCO genes identified in B. rape, B. oleracea and B. napus.

	Gene ID	Nucleotide length (bp)	Amino acid	Molecular weight (KD)	PI	Genome location	Number of introns	Number of exons
PCO1	BnaA03G0061400ZS	897	299	33.22	7.5	ChrA03: 2,917,715-2,919,518	4	5
	BnaA10G0211400ZS	924	308	34.18	8.01	ChrA10: 22,378,464-22,380,486	4	5
	BnaC03G0070600ZS	879	293	32.75	7.77	ChrC03: 3,619,461-3,621,089	4	5
	BnaC09G0512300ZS	771	257	28.66	5.91	ChrC09: 61,478,635-61,483,755	4	5
	Bo3G010080	876	292	32.62	7.52	3,923,348-3,924,992	4	5
	Bo9G165400	918	306	33.96	8.78	60,734,007-60,728,419	4	5
	BraA03G006660	897	299	33.3	7.5	2,874,949-2,876,803	4	5
	BraA10G024330	717	239	26.63	5.14	16,636,583-16,634,714	4	5
PCO2	BnaA04G0111100ZS	846	282	31.35	8.21	ChrA04: 12,952,829-12,954,006	4	5
	BnaC04G0395100ZS	858	286	31.75	8.21	ChrC04: 52,079,647-52,080,830	4	5
	Bo4G139190	858	286	31.75	8.21	46,652,407-46,653,590	4	5
	BraA04G013090	846	282	31.34	8.02	9,741,294-9,742,752	4	5
PCO3	BnaA06G0127100ZS	831	277	30.61	5	ChrA06: 7,448,257-7,449,546	4	5
	BnaC04G0074400ZS	246	82	8.65	4.39	ChrC04: 6,512,883-6,514,191	2	3
	BnaC05G0155300ZS	831	277	30.62	5.01	ChrC05: 9,956,060-9,957,351	4	5
	BnaC05G0283800ZS	318	106	11.58	4.17	ChrC05: 24,420,297-24,421,163	1	2
	BnaC05G0471900ZS	387	129	14.3	8.6	ChrC05: 52,043,598-52,044,332	1	2
	BnaC09G0521600ZS	387	129	14.23	8.37	ChrC09: 62,173,896-62,174,630	1	2
	Bo5G025500	834	278	30.67	5.01	9,490,230-9,491,521	4	5
	BraA06G014250	831	277	30.55	4.89	7,467,637-7,469,108	4	5
PCO4	BnaA03G0207500ZS	624	208	23.51	6.5	ChrA03: 10,842,878-10,844,151	3	4
	BnaA05G0034400ZS	729	243	27.11	6.03	ChrA05: 1,925,460-1,927,173	4	5
	BnaC03G0244400ZS	636	212	23.54	8.04	ChrC03: 14,957,859-14,959,210	4	5
	BnaC04G0037200ZS	717	239	26.81	6.42	ChrC04: 3,448,037-3,452,967	4	5
	Bo3G036180	636	212	23.53	8.04	15,839,677-15,841,028	4	5
	Bo4G017390	717	239	26.75	6.23	3,496,211-3,497,884	4	5

Gene ID	Nucleo-tide length (bp)	Amino acid	Molec-ular weight (KD)	PI	Genome location	Number of in-trons	Number of exons
BraA05G003220	717	239	26.72	6.42	1,719,860-1,722,377	4	5
BnaA04G0023400ZS	729	243	27.23	6.84	ChrA04: 1,525,382-1,527,222	4	5
BnaA09G0537000ZS	729	243	27.13	6.59	ChrA09: 56,193,470-56,194,841	4	5
BnaC04G0289700ZS	729	243	27.22	6.78	ChrC04: 39,729,996-39,732,404	4	5
BnaC08G0384200ZS	729	243	27.2	6.5	ChrC08: 44,790,558-44,792,029	4	5
Bo4G108280	729	243	27.22	6.78	34,730,028-34,731,729	4	5
Bo8G092800	729	243	27.2	6.5	41,444,747-41,446,184	4	5
BraA04G002890	729	243	27.22	6.99	1,612,409-1,614,775	4	5
BraA09G050770	729	243	27.12	6.59	36,671,235-36,668,985	4	5

(PCO5 is the row-group label spanning the B. napus / B. oleracea / B. rapa rows)

To explore the classification and evolutionary characteristics of the PCO proteins, an unrooted phylogenetic tree based on the 40 protein sequences of B. napus (20), B. oleracea (8) and B. rapa (7) and Arabidopsis (5) PCO genes was constructed in MEGA 11 (Figure 1). According to the phylogenetic analysis, the PCO genes were divided into five groups: PCO1 (homologous to AT5G15120.1/AtPCO1), PCO2 (homologous to AT5G39890.1/AtPCO2), PCO3 (homologous to AT1G18490.1/AtPCO3), PCO4 (homologous to AT2G42670.1/AtPCO4) and PCO5 (homologous to AT3G58670.1/AtPCO5). In addition, according to the phylogenetic tree, PCOs can be clearly divided into three clades: PCO1s and PCO2s are clade A, PCO3s are clade B, and PCO4 and PCO5 are clade C, respectively.
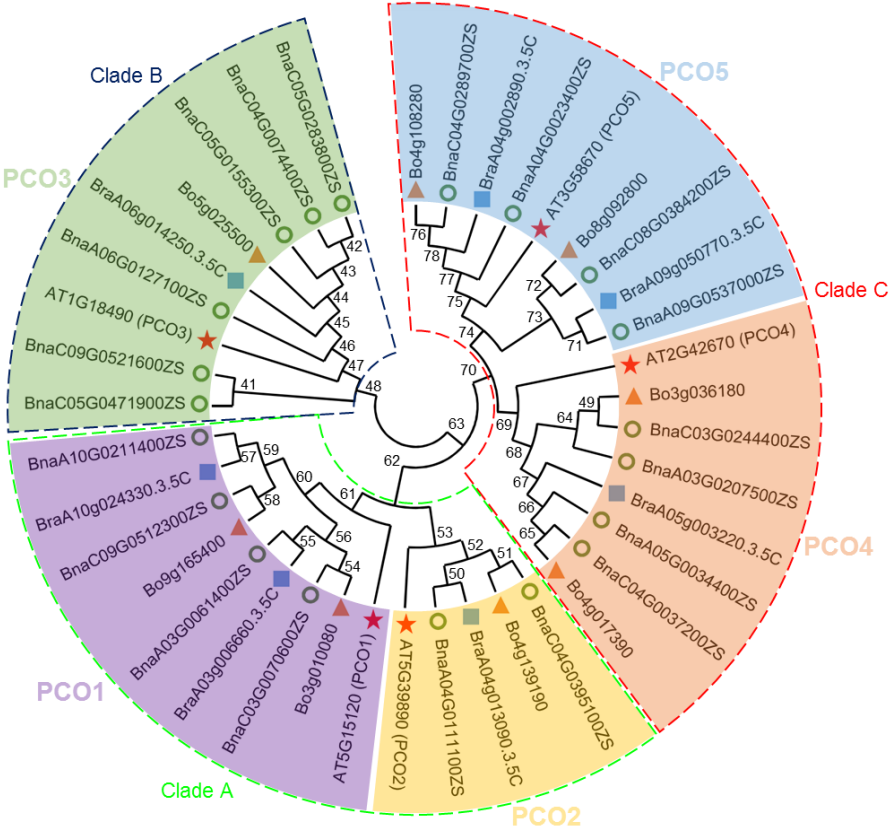

Figure 1. Phylogenetic analysis of 40 PCO proteins from B.napus (20), B. rapa (7), B. oleracea (8) and Arabidopsis (5).

2.2. Chromosomal Distribution and Duplication of BnaPCOs

As shown in Figure 2, the 20 BnaPCOs were scattered on 11 of the 19 B. napus chromosomes. The numbers of BnaPCOs on An-subgenome (8 genes) and Cn-subgenome (12 genes) showed a bi-ased trend, with more genes on Cn-subgenome. B. napus (AnAnCnCn, 2n = 38) originated from the hybridization event between Brassica rapa (AnAn, n = 10) and Brassica oleracea (CnCn, n = 9) about 7500 years ago, and Brassicaceae species underwent a whole-genome triplication event [16]. There-fore, theoretically, B. rapa, B. oleracea, and B. napus genomes may have 15, 15, and 30 PCOs expanded from the 5 AtPCOs. In fact, only 7, 8 and 20 genes were identified in B. rapa, B. oleracea and B. napus, respectively (Table S1), indicating that some PCOs may have been lost during evolution.
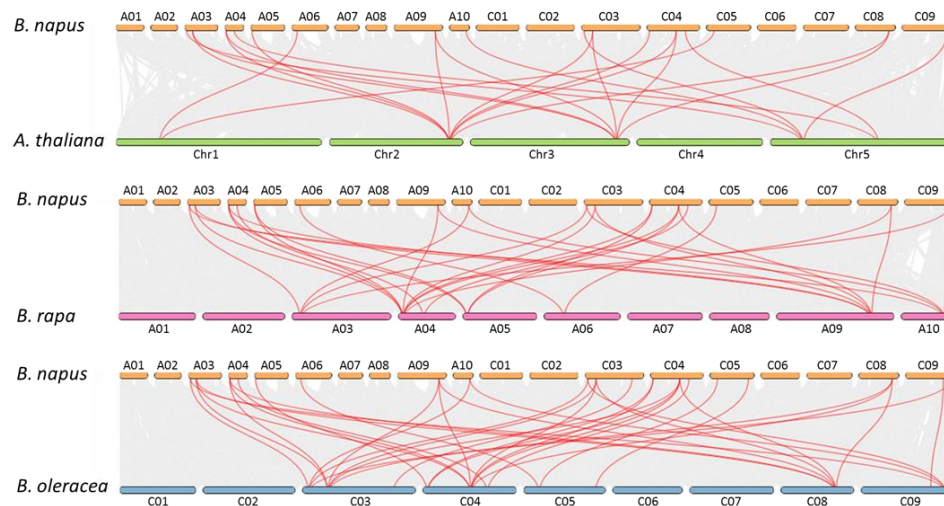

Figure 2. Syntenic relationship of PCO genes in B.napus and three ancestral plant species. Grey lines in the background showed the collinear blocks within rapeseed and other plant genomes, while the red lines highlighted the syntenic PCO gene pairs.

The An and Cn subgenomes of B. napus were collinear with the corresponding diploid Ar and Co genomes, most of the An-Ar and Cn-Co homologous pairs showed similar chromosomal locations [16]. However, the C genome of B. napus had more homologous genes corresponding to the B. oleracea PCO3 gene (Bo5G025500), and the A genome had more homologous genes corresponding to the B. rapa PCO4 genes (BraA05G003220), while the homologous genes in the PCO1, PCO2 and PCO5 subgroups were relatively conserved. It was inferred that the tandem genome was mutated or duplicated in the homologous evolution process.

2.3. Gene Structures and Motif Analysis of PCOs in B. napus

In order to study the homology domain and conservation degree of the BnaPCOs, MEME [24] and TBtools [25] were used to predict and visualize their conserved domain and gene organization, respectively. 8 motifs were predicted by MEME (<u>https://meme-suite.org/meme/tools/meme</u>) (Figure S1). As shown in Figure 3, Motif 2 and 5 were the most conserved, with 38 of the 40 genes containing these two motifs. All of the PCO4 and PCO5 proteins contained Motif 8, whereas none of the other genes contained Motif 8. All PCO5 proteins contained eight motifs, of which the Motif 7 was located at the C terminal. However, all PCO2 proteins contained seven motifs in addition to the Motif 8, with the Motif 7 located at the N terminal. Two of the four short abnormal genes, BnaC05G0471900ZS and BnaC09G0521600ZS, lacked Motif 1-3-4-6-7-8, while BnaC04G0074400ZS and BnaC05G0283800ZS lacked Motif 1-2-4-5-6-8.
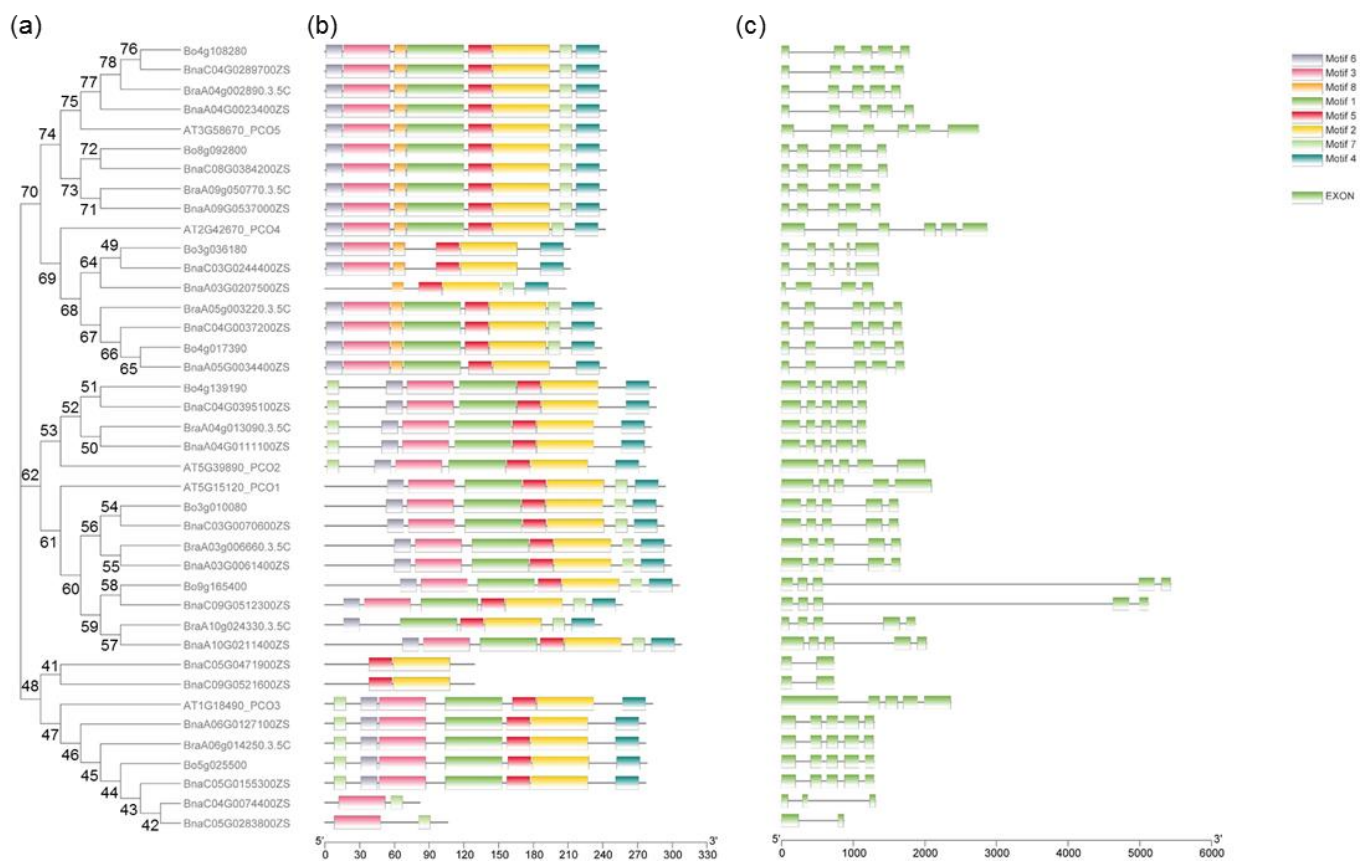
Figure 3. Phylogenetic tree (a), gene motif (b) and gene structure (c) of PCO of Arabidopsis, B.napus, B. rapa and B. oleracea. (a) Neighbor-joining phylogenetic tree showed the relationship among 5 Arabidopsis, 20 B.napus, 7 B.rapa and 8 B. oleracea PCO proteins. (b) Eight motifs in PCO proteins were identified by MEME tools. (c) Green box indicated the exon regions on PCO genes.

It is reported that some introns play an essential role in transcriptional regulation [26]. We also investigated the distribution of introns and exons to study the diversity of gene structure. Five exons and four introns were distributed on most of PCOs genes. And the intron phases of BnaPCOs were highly conserved in the same group, implicating the evolutionary similarity between these members.

2.4. Cis-Element Analysis of BnaPCOs

Cis-elements regulate the initiation and efficiency of gene transcription by binding to transcription factors [27]. We analyzed the cis-elements of 20 PCOs promoters with PlantCARE [28]. Most cis-elements of BnaPCOs promoters were associated with plant growth and development, hormone response and abiotic stress response (Table S2). Some cis-elements, such as, anaerobic induction, circadian control, defense and stress responsive, drought induction, light response, low temperature response and meristem expression, were the main members in regulating environmental stress. Among them, the most common elements were associated with light response and anaerobic induction, indicating that the growth and development of plants regulated by BnaPCOs was affected by light and oxygen. PCOs was known as a sensor of oxygen, which was consistent with the cis-elements result. However, the light response of PCOs needed further study.

In addition, there were more anaerobic-induction elements on PCO2s promoter region compared with other genes and there was no anaerobic-induction elements on the promoter of BnaC05G0471900ZS (belongs to BnaPCO3) (Figure 4). The cis-elements in the upstream promoter region of genes are closely related to the expression and function of downstream genes [29]. Therefore, according to the anaerobic-induction elements result, it indicated that the expression of PCO2s may be highly induced by hypoxia and it also implies the diversification of biological functions of PCOs genes in Brassica napus.
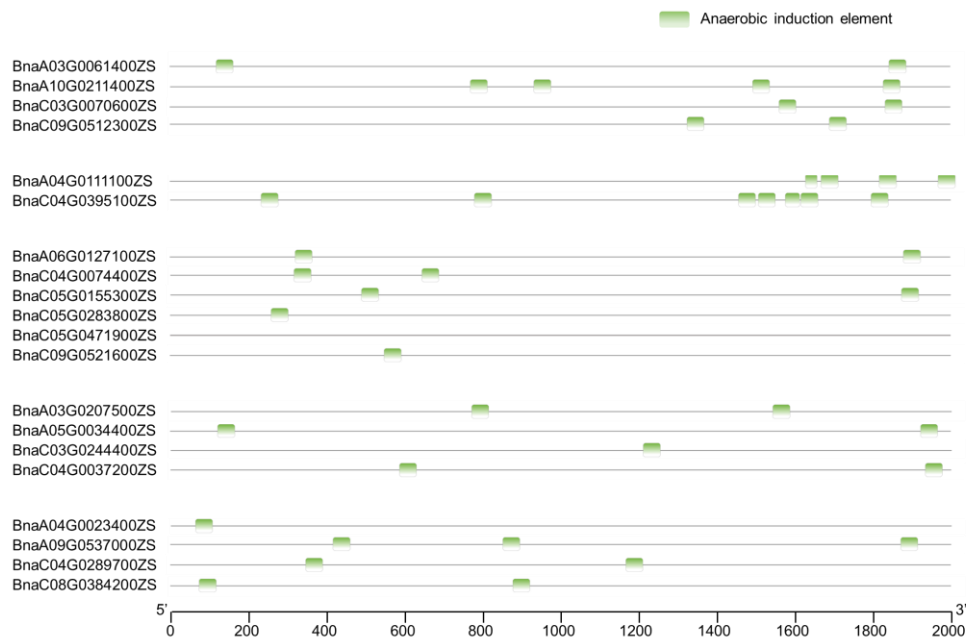
Figure 4. Anaerobic-induction cis-elements distributed on the BnPCO promoter regions. Cis-elements were identified by PlantCARE.

2.5. Expression Profiling of PCO Genes in Different Tissues

Based on Arabidopsis eFP Browser data (http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi) and RNA-Seq data (B. rapa: GSE43245, B. oleateaca: GSE42891 and B. napus: BnIR database) (Table S3), the PCOs gene were expressed in different vegetative and reproductive organs of the four species at different developmental stages. In general, the expression pattern of PCO differed between groups (Figure 5). Almost all PCO were weakly expressed in pollen, indicating that PCO expression was down-regulated in sperm cells, presumably due to chromosomal structure or histone modifications. On the other hand, PCO5s were expressed at similar levels in all tissues, indicating that it appeared to be a housekeeping gene involved in plant growth and development.
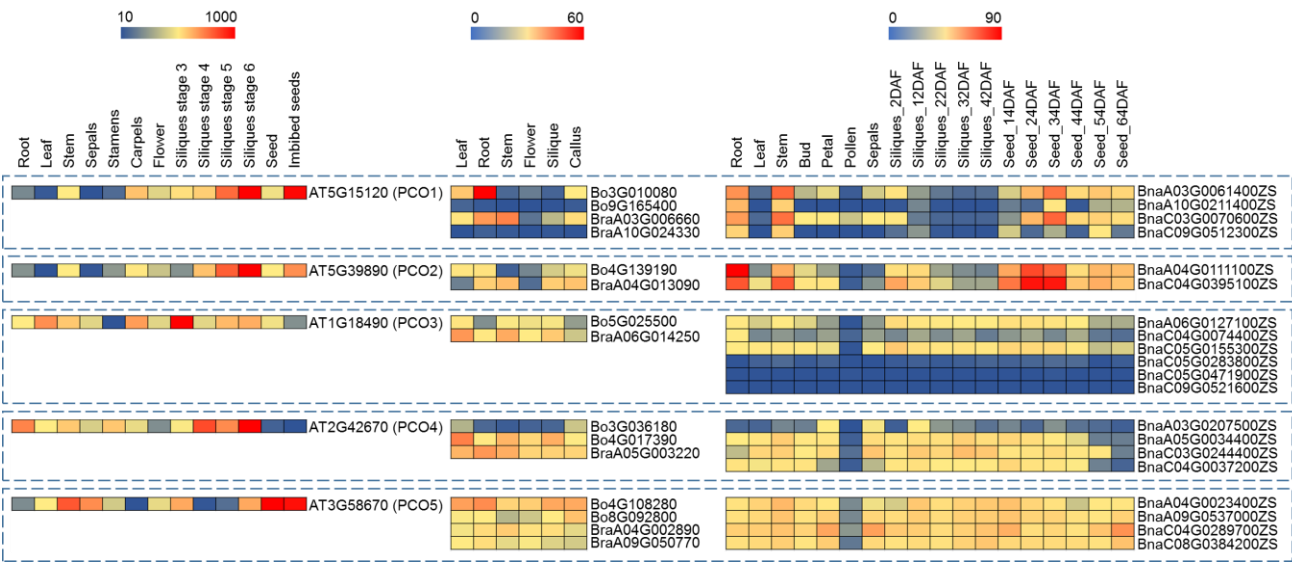

Figure 5. Expression of AtPCO, BraPCO, BoPCO and BnaPCO in different tissues. The expression levels of PCO genes were indicated by differently colored rectangles. DAF represented day after flower.

There were six PCO3 genes in B. napus, three of which were expressed at high levels and the other three at low levels. The genes with higher expression levels were generated from B. oleateaca and B. rapa, and the expression pattern of these genes was also similar to their original genes. However, the genes with lower expression levels were extra mutated. This suggested that the previous derived gene may play a major role, while the new genes may be pseudogenes that have been amplified during evolution or may be silent under normal conditions and be as a backup for their homolog genes under special conditions.

2.6. Expression Profiling of PCO Genes under Abiotic Stress Treatment

To reveal the roles of PCOs in stress response in B. napus, the expression patterns upon various abiotic treatment were investigated (Figure 6, Table S4). In general, the expression of most PCO genes in leaves did not change significantly under various abiotic stress treatments (Figure 6a). In leaves, PCO3 gene expression was significantly induced under freezing stress, and PCO5 gene expression was increased under salt and osmotic stress as well. However, the response of PCO genes to stress was much stronger in roots than in leaves (Figure 6b). In roots, the expression of PCO1 and PCO2 were strongly induced by drought, and the expression of PCO1, PCO2, PCO3 and PCO4 were extremely down-regulated by freezing and cold stress. But PCO5 gene expression in roots did not change significantly under the abiotic stress treatments nonetheless, which was consistent with the hypothesis that PCO5 worked as a housekeeping gene.
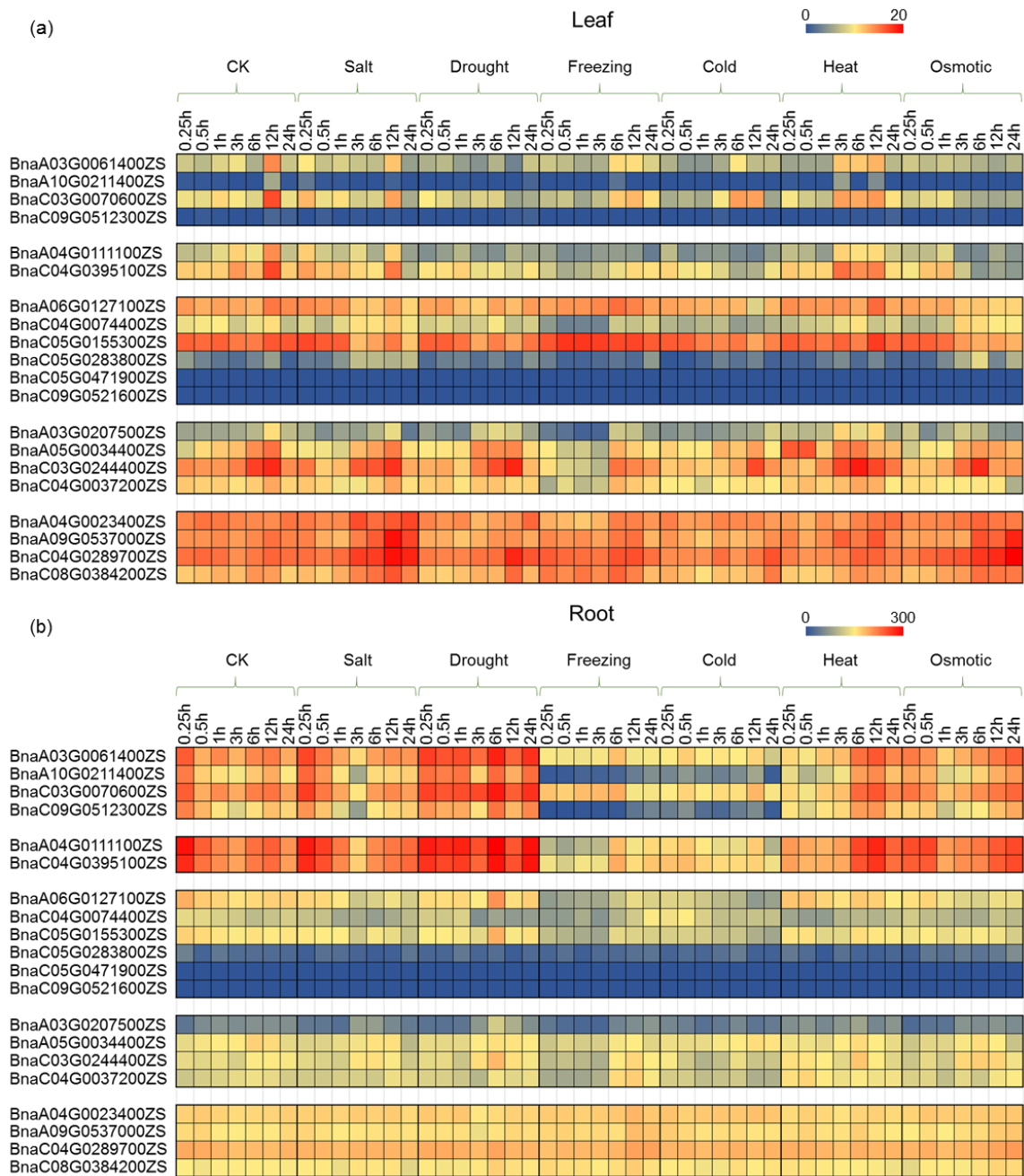
Figure 6. Expression of BnaPCO genes under different abiotic stress treatments. The expression levels (TPM values) of BnaPCO genes were indicated by differently colored rectangles. (a) represented the expression in leaf. (b) represented the expression in root.

Though there were similar motifs of drought-inducible and low-temperature responsive on PCOs promoter region, the gene expression pattern was various under drought, freezing or cold stress treatment. It implied that there were other regulators along with cis-elements to regulate PCO gene expression. It was interesting that PCO3 was up-regulated in leaves and down-regulated in roots after freezing treatment. In other words, PCO3 showed opposite response patterns in leaves and roots under freezing stress treatment. This suggested that PCO3 may play different roles in leaves and roots under freezing stress, and further studies were needed.

2.7. Expression Profiling of PCO Genes under Waterlogging Stress

Waterlogging removes air from soil leading to a blockage of gas exchange between the soil and the atmosphere [30, 31]. Meanwhile, the diffusion rate of oxygen in water is only one tenth of that in

air. As a result, oxygen availability in flooded soils is greatly limited, leading to suppression of root respiration. As mentioned before, PCO is an oxygen sensor in plant [8]. Therefore, it is speculated that PCO plays an important role in the hypoxic response induced by waterlogging stress. In order to elucidate the potential function of BnaPCO in response to waterlogging stress, the transcriptional patterns of BnaPCO in B. napus were studied after waterlogging stress. Transcriptome data (Figure 7, Table S5) showed that compared with CK, the expression of PCOs was significantly induced by waterlogging stress. Compared to other genes, the expression of PCO2 gene pairs (BnaA04G0111100ZS and BnaC04G0395100ZS) were strongly induced both in leaves and roots. While, the expression of PCO4 gene pairs (BnaA05G0034400ZS, BnaC03G0244400ZS and BnaC04G0037200ZS) were particularly strongly expressed in leaves. In addition, gene expression of PCO3 (BnaC05G0471900ZS and BnaC09G0521600ZS) were barely detectable in the transcriptome data, which was consistent with the previous result (Figure 5 and Figure 6). According to the expression results, the function of PCO was conserved under hypoxic stress. It suggested that BnaPCO was a vital component, connecting environmental stimulus with cellular and physiological response and BnaPCO could be a potential target for improving waterlogging stress tolerance.
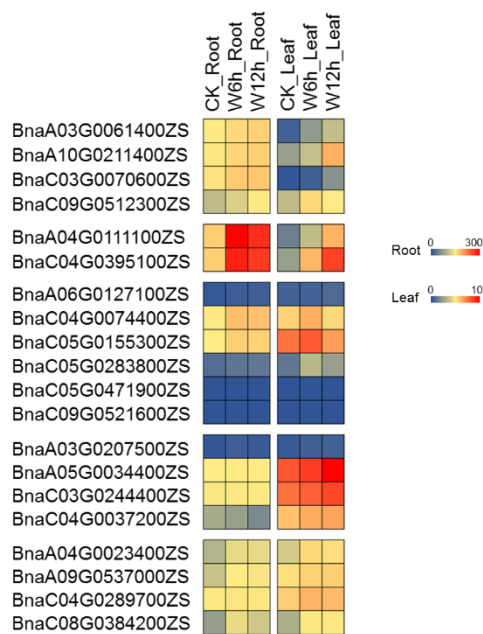

Figure 7. Expression of BnaPCO genes (FPKM) under waterlogging stress treatment for 6h and 12h. CK represented normal condition. W represented waterlogging stress treatment.

3. Discussion

Plant cysteine oxidase family (PCO) is a set of plant O_2-sensing enzymes, which catalyze the O_2-dependent step [8]. In *Arabidopsis*, PCO has five members, PCO1, PCO2, PCO3, PCO4 and PCO5. In this study, we found 20, 8 and 7 *PCO* genes in *B. napus*, *B. oleracea* and *B. rapa*, respectively (Figure 1). In *B. napus*, the number of *PCO* genes in the An subgenome (8) was almost the same as that in their diploid ancestors *B. rapa* (7) (Table 1). This showed that all An subgenome *PCO* genes were preserved after the whole genome duplication event in *B. napus*. However, *PCO3* genes in the Cn subgenome (*BnaC04G0074400ZS*, *BnaC05G0155300ZS*, *BnaC05G0283800ZS*, *BnaC05G0471900ZS* and *BnaC09G0521600ZS*) were mutated from the diploid ancestor *B. oleracea* (*Bo5G025500*). It suggested that Cn subgenome was much flexible during the evolution, compared with An subgenome [32]. *PCO* sequence alignment (Figure S2-S6) revealed that most *PCOs* were conserved in *Brassicaceae*, indicating that these duplicated *PCO* genes can still retain the function of their ancestors in these species.

The gene expression pattern of duplicated genes with similar functions may change during the formation of allopolyploids, which takes several typical patterns, including transgressive up–/down-regulation, unequal parental contributions, and silencing [32, 33]. Although *PCO3* was expanded in

B. napus, half of them were slightly expressed in different tissues or abiotic stress treatment. According to the results of *PCO3* alignment (Figure S4), *BnaA06G0127100ZS* was derived from *BraA06G014250*, *BnaC05G0155300ZS* was derived from *Bo5G025500*, and the remaining four genes may be generated by mutation or amplification of DNA fragments during the evolution of polyploid. The expression patterns of *BnaA06G0127100ZS* and *BnaC05G0155300ZS* maintained their expression patterns in two diploid progenitors. In addition, the expression of *BnaC04G0074400ZS* and *BnaC05G0283800ZS* responded to abiotic stress (Figure 6). It indicated that the newly generated genes may contribute to phenotypic differences between allopolyploids and their parental species under abiotic stress conditions.

The binding between transcriptional factors (TFs) and cis-regulatory DNA sequences in the 5'-upstream of genes either activates or represses gene expression [34, 35]. In general, genes with related expression patterns contain stress response elements in their promoter region and are likely to be regulated by stress related TFs [20, 26, 31, 36, 37]. It is reported that HSFB2b directly binds to *GmC4H* and *GmCHS3* to regulate the gene expression in response to salt stress, since there are HSEs (Heat Shock Elements) in the promoter regions of the *GmC4H* and *GmCHS3* [38]. It was showed that hormone-responsive elements and environmental stress-related elements were distributed on the *BnaPCO* promoters (Table S2). Combined with the expression data with abiotic stress treatment, *BnaPCO* expression was regulated by various stress responses, especially drought, freezing, cold and waterlogging stress (Figure 6 and 7). However, the expression level was different under different stress conditions. For instance, there were 7 anaerobic-induction cis-elements on the promoter region of *BnaC04G0395100ZS*, but the degree of waterlogging induced expression was much different between leaves and roots (Figure 7). As chromosome structure, histone modification, DNA methylation, transcriptional factors, cis-elements and other regulators work together to regulate gene expression [35], more studies are needed to explore the regulation mechanism of *PCO* expression under abiotic stress.

Oxygen homeostasis is critical for crop development, and hypoxia in plants is typically a consequence of reduced O_2 diffusion under conditions of waterlogging or submergence [19, 39-41]. The quality and yield of rapeseed are seriously affected by waterlogging stress in China [23, 41]. It is reported that the response to hypoxia in rice, Arabidopsis and barley is mediated by the group VII ETHYLENE RESPONSE FACTORs (ERF-VIIs) [1, 5-7, 15]. Moreover, the PCOs directly link O_2 availability to ERF-VII stability and anaerobic adaptation, leading to the suggestion that they act as plant O_2 sensors [3, 4, 8]. There were multiple anaerobic-induction cis-elements distributed on the *BnaPCO* promoters, and *BnaPCO* expression were induced by waterlogging stress after 6h treatment both in leaves and roots. Furthermore, H164 and D176 were in the AtPCO4 active site [9] and it was conserved in BnaPCO4 as shown in Figure S7. It implied that the function of BnaPCO in catalyzing ERF-VIIs could be conserved and targeting PCOs will be an effective way to improve the rapeseed tolerance to waterlogging stress by manipulating their O_2 sensitivity and/or substrate specificity.

4. Materials and Methods

4.1. Identification of the PCO Gene Family

The protein and nucleotide sequences of *AtPCOs* were obtained with TAIR (https://www.arabidopsis.org/). AtPCO proteins were used as query sequences to search for the PCO proteins of *B. napus*, *B. rapa* and *B. oleracea* using BLASTp (E-value < 1 × 10−5) in BARD (http://www.brassicadb.cn/#/BLAST/; protein databases were Brana ZS V2.0 pep, Brara Chiifu V3.5 pep and Braol JZS V2.0 pep, respectively). ExPASy (Gasteiger et al., 2003) was used to investigate the physical and chemical properties of these PCO proteins.

4.2. Phylogenetic Analysis and Chromosomal Locations

ClustalW was used to align the multiple sequences of all PCO proteins (from *Arabidopsis*, *B. napus*, *B. rapa* and *B. oleracea*), and MEGA 11 was used to build a phylogenetic tree using the neighbor-joining (NJ) phylogenetic technique with 1000 bootstrap replicates. The genome data and annotation

information (gff3 files) were downloaded from public databases. TBtools version 1.116 was used to examine the gene chromosomal localization with the genome data and gff3 files [25].

4.3. Analysis of Gene Structure, Motif Composition and Cis-element

The motifs of BnaPCO proteins were predicted using the MEME v5.5.2 [24]. The number of motifs should not exceed 8. The distribution of motifs occurs zero or one time in each sequence. In order to investigate the structural characteristics of the *BnaPCOs*, the gff3 file was downloaded from the *B. napus* database (http://cbi.hzau.edu.cn/cgi-bin/rape/download_ext), which has the annotation information of the *B. napus* genome. The position information of introns and exons is obtained from the gff3 file. Meanwhile, the motif information was submitted to TBtools to graphically display gene structures and motif distributions.

To identify the cis-element of *BnaPCOs*, TBtools was used to obtain the 2,000 bp sequences in front of the genomic CDS. Then, the PlantCARE [28] was used to predict the cis- elements on these promoters. Thus, the number and types of different cis-acting elements in *BnaPCOs* were classified and visualized with TBtools.

4.4. Plant Materials and Treatments, Heat Map Analysis of the PCO Transcriptome Data

Zhongshuang 11 (the semi-winter cultivar of *Brassica napus*) seeds were germinated on filter paper, and the seedlings were then transplanted into pots with soil and nurtured in a growth chamber for four weeks (23 °C, 16 h light/8 h dark cycle). The waterlogging treatment was performed as previously described [22]. The pots of 4-week-old seedlings were placed in a 28cm x 14cm x 14cm container filled with water and the water level was maintained at about 2cm above the soil surface. Control plants (CK) remained well-watered throughout the experiment. Leaf and root samples of CK, 6h-treatment and 12h-treatment were collected for RNA isolation and RNA Sequencing was performed by the Illumina NovaSeq 6000 (Novogene). Hisat2 v2.0.5 was used to map the reads to reference genome and gene expression level was determined by FPKM calculation.

The expression patterns of *BnaPCOs* in different tissues and other abiotic stress treatment were obtained from the BnTIR (Brassica napus transcriptome information resource) database. The expression data of *AtPCOs*, *BraPCOs* and *BoPCOs* were obtained from Arabidopsis eFP Browser data, RNA-Seq data of *B.rapa* (GSE43245) and RNA-Seq data of *B.oleareaca* (GSE42891). All the expression data was standardized based on log2 scale, and clustered and visualized with TBtools.

5. Conclusions

In this study, we identified 20, 7 and 8 PCO (Plant Cysteine Oxidase) proteins in *B. napus*, *B. rapa* and *B. oleracea*, respectively, by exploring the important role of *PCO* genes in *Brassicaceae* plants. Collinearity analysis shows that the *PCO* gene family was relatively conserved in evolution of *B. rapa*, *B. oleracea*, *Arabidopsis*, and *B. napus*. However, there were mutations or duplications of *PCO3* and *PCO4* during homologous evolution process. The cis-elements that regulate plant growth and development and response to biotic and abiotic stresses were found in the *BnaPCO* promoters. In addition, we found that the *Bna/Bra/BoPCO* genes were expressed differently in different tissues at different developmental stages. Remarkably, some of *BnaPCOs* were significantly induced after waterlogging treatment, which was consistent with the cis-element analysis and previous studies. This study provides a foundation for further understanding the biology and stress response functions of PCO family genes in *B. napus*.

Supplementary Materials: Figure S1: Sequence logos of conserved motifs in BnPCOs; Figure S2: *PCO1* CDS sequence alignment; Figure S3: *PCO2* CDS sequence alignment; Figure S4: *PCO3* CDS sequence alignment; Figure S5: *PCO4* CDS sequence alignment; Figure S6: *PCO5* CDS sequence alignment; Figure S7: PCO4 protein sequence alignment; Table S1: Sequence information of PCO; Table S2: The cis-acting elements on the *BnaPCOs* promoters; Table S3. *PCO* gene expression in different tissues; Table S4. *BnaPCO* gene expression under abiotic stress; Table S5. *BnaPCO* gene expression under waterlogging stress.

References

1.　Gibbs, D.J., H.M. Tedds, A.M. Labandera, M. Bailey, M.D. White, S. Hartman, C. Sprigg, S.L. Mogg, R. Osborne, C. Dambire, T. Boeckx, Z. Paling, L. Voesenek, E. Flashman, and M.J. Holdsworth, Oxygen-dependent proteolysis regulates the stability of angiosperm polycomb repressive complex 2 subunit VERNALIZATION 2. *Nat Commun*, 2018. **9**(1): p. 5438.

2.　Weits, D.A., A.B. Kunkowska, N.C.W. Kamps, K.M.S. Portz, N.K. Packbier, Z. Nemec Venza, C. Gaillochet, J.U. Lohmann, O. Pedersen, J.T. van Dongen, and F. Licausi, An apical hypoxic niche sets the pace of shoot meristem activity. *Nature*, 2019. **569**(7758): p. 714-717.

3.　White, M.D., M. Klecker, R.J. Hopkinson, D.A. Weits, C. Mueller, C. Naumann, R. O'Neill, J. Wickens, J. Yang, J.C. Brooks-Bartlett, E.F. Garman, T.N. Grossmann, N. Dissmeyer, and E. Flashman, Plant cysteine oxidases are dioxygenases that directly enable arginyl transferase-catalysed arginylation of N-end rule targets. *Nat Commun*, 2017. **8**: p. 14690.

4.　Weits, D.A., B. Giuntoli, M. Kosmacz, S. Parlanti, H.M. Hubberten, H. Riegler, R. Hoefgen, P. Perata, J.T. van Dongen, and F. Licausi, Plant cysteine oxidases control the oxygen-dependent branch of the N-end-rule pathway. *Nat Commun*, 2014. **5**: p. 3425.

5.　Licausi, F., J.T. van Dongen, B. Giuntoli, G. Novi, A. Santaniello, P. Geigenberger, and P. Perata, *HRE1 and HRE2,* two hypoxia-inducible ethylene response factors, affect anaerobic responses in *Arabidopsis thaliana. Plant J*, 2010. **62**(2): p. 302-15.

6.　Hattori, Y., K. Nagai, S. Furukawa, X.J. Song, R. Kawano, H. Sakakibara, J. Wu, T. Matsumoto, A. Yoshimura, H. Kitano, M. Matsuoka, H. Mori, and M. Ashikari, The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water. *Nature,* 2009. **460**(7258): p. 1026-30.

7.　Hinz, M., I.W. Wilson, J. Yang, K. Buerstenbinder, D. Llewellyn, E.S. Dennis, M. Sauter, and R. Dolferus, *Arabidopsis RAP2.2:* an ethylene response transcription factor that is important for hypoxia survival. *Plant Physiol*, 2010. **153**(2): p. 757-72.

8.　White, M.D., J. Kamps, S. East, L.J. Taylor Kearney, and E. Flashman, The plant cysteine oxidases from *Arabidopsis thaliana* are kinetically tailored to act as oxygen sensors. *J Biol Chem*, 2018. **293**(30): p. 11786-11795.

9.　White, M.D., L. Dalle Carbonare, M. Lavilla Puerta, S. Iacopino, M. Edwards, K. Dunne, E. Pires, C. Levy, M.A. McDonough, F. Licausi, and E. Flashman, Structures of *Arabidopsis thaliana* oxygen-sensing plant cysteine oxidases 4 and 5 enable targeted manipulation of their activity. *Proc Natl Acad Sci U S A*, 2020. **117**(37): p. 23140-23147.

10.　Masson, N., T.P. Keeley, B. Giuntoli, M.D. White, M.L. Puerta, P. Perata, R.J. Hopkinson, E. Flashman, F. Licausi, and P.J. Ratcliffe, Conserved N-terminal cysteine dioxygenases transduce responses to hypoxia in animals and plants. *Science*, 2019. **365**(6448): p. 65-69.

11.　Taylor-Kearney, L.J., S. Madden, J. Wilson, W.K. Myers, D.M. Gunawardana, E. Pires, P. Holdship, A. Tumber, R.E.M. Rickaby, and E. Flashman, Plant cysteine oxidase oxygen-sensing function is conserved in early land plants and algae. *ACS Bio Med Chem Au*, 2022. **2**(5): p. 521-528.

12.　Taylor-Kearney, L.J. and E. Flashman, Targeting plant cysteine oxidase activity for improved submergence tolerance. *Plant J*, 2022. **109**(4): p. 779-788.

13.　Dirr, A., D.M. Gunawardana, and E. Flashman, Kinetic measurements to investigate the oxygen-sensing properties of plant cysteine oxidases. *Methods Mol Biol*, 2023. **2648**: p. 207-230.

14.　Chen, Z., Q. Guo, G. Wu, J. Wen, S. Liao, and C. Xu, Molecular basis for cysteine oxidation by plant cysteine oxidases from *Arabidopsis thaliana. J Struct Biol*, 2021. **213**(1): p. 107663.

15.   Xu, K., X. Xu, T. Fukao, P. Canlas, R. Maghirang-Rodriguez, S. Heuer, A.M. Ismail, J. Bailey-Serres, P.C. Ronald, and D.J. Mackill, *Sub1A* is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature*, 2006. **442**(7103): p. 705-8.

16.   Cheng, F., J. Wu, and X. Wang, Genome triplication drove the diversification of Brassica plants. *Hortic Res*, 2014. **1**: p. 14024.

17.   Hu, J., B. Chen, J. Zhao, F. Zhang, T. Xie, K. Xu, G. Gao, G. Yan, H. Li, L. Li, G. Ji, H. An, H. Li, Q. Huang, M. Zhang, J. Wu, W. Song, X. Zhang, Y. Luo, J. Chris Pires, J. Batley, S. Tian, and X. Wu, Genomic selection and genetic architecture of agronomic traits during modern rapeseed breeding. *Nat Genet*, 2022. **54**(5): p. 694-704.

18.   Chalhoub, B., F. Denoeud, S. Liu, I.A. Parkin, H. Tang, X. Wang, J. Chiquet, H. Belcram, C. Tong, B. Samans, M. Correa, C. Da Silva, J. Just, C. Falentin, C.S. Koh, I. Le Clainche, M. Bernard, P. Bento, B. Noel, K. Labadie, A. Alberti, M. Charles, D. Arnaud, H. Guo, C. Daviaud, S. Alamery, K. Jabbari, M. Zhao, P.P. Edger, H. Chelaifa, D. Tack, G. Lassalle, I. Mestiri, N. Schnel, M.C. Le Paslier, G. Fan, V. Renault, P.E. Bayer, A.A. Golicz, S. Manoli, T.H. Lee, V.H. Thi, S. Chalabi, Q. Hu, C. Fan, R. Tollenaere, Y. Lu, C. Battail, J. Shen, C.H. Sidebottom, X. Wang, A. Canaguier, A. Chauveau, A. Berard, G. Deniot, M. Guan, Z. Liu, F. Sun, Y.P. Lim, E. Lyons, C.D. Town, I. Bancroft, X. Wang, J. Meng, J. Ma, J.C. Pires, G.J. King, D. Brunel, R. Delourme, M. Renard, J.M. Aury, K.L. Adams, J. Batley, R.J. Snowdon, J. Tost, D. Edwards, Y. Zhou, W. Hua, A.G. Sharpe, A.H. Paterson, C. Guan, and P. Wincker, Plant genetics. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science*, 2014. **345**(6199): p. 950-3.

19.   Kayum, M.A., J.I. Park, U.K. Nath, M.K. Biswas, H.T. Kim, and I.S. Nou, Genome-wide expression profiling of aquaporin genes confer responses to abiotic and biotic stresses in *Brassica rapa. BMC Plant Biol*, 2017. **17**(1): p. 23.

20.   Ma, L., J. Wu, W. Qi, J.A. Coulter, Y. Fang, X. Li, L. Liu, J. Jin, Z. Niu, J. Yue, and W. Sun, Screening and verification of reference genes for analysis of gene expression in winter rapeseed *(Brassica rapa L.)* under abiotic stress. *PLoS One*, 2020. **15**(9): p. e0236577.

21.   Tan, X., W. Long, L. Zeng, X. Ding, Y. Cheng, X. Zhang, and X. Zou, Melatonin-induced transcriptome variation of rapeseed seedlings under salt stress. *Int J Mol Sci*, 2019. **20**(21).

22.   Tong, J., T.C. Walk, P. Han, L. Chen, X. Shen, Y. Li, C. Gu, L. Xie, X. Hu, X. Liao, and L. Qin, Genome-wide identification and analysis of high-affinity nitrate transporter 2 (NRT2) family genes in rapeseed *(Brassica napus L.)* and their responses to various stresses. *BMC Plant Biol*, 2020. **20**(1): p. 464.

23.   Wang, Z., Y. Han, S. Luo, X. Rong, H. Song, N. Jiang, C. Li, and L. Yang, Calcium peroxide alleviates the waterlogging stress of rapeseed by improving root growth status in a rice-rape rotation field. *Front Plant Sci*, 2022. **13**: p. 1048227.

24.   Bailey, T.L., J. Johnson, C.E. Grant, and W.S. Noble, The MEME suite. *Nucleic Acids Res*, 2015. **43**(W1): p. W39-49.

25.   Chen, C., H. Chen, Y. Zhang, H.R. Thomas, M.H. Frank, Y. He, and R. Xia, TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant*, 2020. **13**(8): p. 1194-1202.

26.   Zhao, J., D.S. Favero, J. Qiu, E.H. Roalson, and M.M. Neff, Insights into the evolution and diversification of the AT-hook Motif Nuclear Localized gene family in land plants. *BMC Plant Biol*, 2014. **14**: p. 266.

27.   Maire, P., J. Wuarin, and U. Schibler, The role of cis-acting promoter elements in tissue-specific albumin gene expression. *Science*, 1989. **244**(4902): p. 343-6.

28.   Lescot, M., P. Dehais, G. Thijs, K. Marchal, Y. Moreau, Y. Van de Peer, P. Rouze, and S. Rombauts, PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res*, 2002. **30**(1): p. 325-7.

29.   Chow, C.N., Y.F. Chiang-Hsieh, C.H. Chien, H.Q. Zheng, T.Y. Lee, N.Y. Wu, K.C. Tseng, P.F. Hou, and W.C. Chang, Delineation of condition specific Cis- and Trans-acting elements in plant promoters under various Endo- and exogenous stimuli. *BMC Genomics*, 2018. **19**(Suppl 2): p. 85.

30.   Mendiondo, G.M., D.J. Gibbs, M. Szurman-Zubrzycka, A. Korn, J. Marquez, I. Szarejko, M. Maluszyn-ski, J. King, B. Axcell, K. Smart, F. Corbineau, and M.J. Holdsworth, Enhanced waterlogging tolerance in barley by manipulation of expression of the N-end rule pathway E3 ligase PROTEOLYSIS6. *Plant Biotechnol J*, 2016. **14**(1): p. 40-50.

31.   Klok, E.J., I.W. Wilson, D. Wilson, S.C. Chapman, R.M. Ewing, S.C. Somerville, W.J. Peacock, R. Dolferus, and E.S. Dennis, Expression profile analysis of the low-oxygen response in Arabidopsis root cultures. Plant Cell, 2002. **14**(10): p. 2481-94.

32.   Jackson, S. and Z.J. Chen, Genomic and expression plasticity of polyploidy. *Curr Opin Plant Biol*, 2010. **13**(2): p. 153-9.

33.   Li, M., R. Wang, X. Wu, and J. Wang, Homoeolog expression bias and expression level dominance (ELD) in four tissues of natural allotetraploid *Brassica napus. BMC Genomics*, 2020. **21**(1): p. 330.

34.   Hernandez-Garcia, C.M. and J.J. Finer, Identification and validation of promoters and cis-acting regulatory elements. *Plant Sci*, 2014. **217-218**: p. 109-19.

35.   Lu, L., W. Wei, J.J. Tao, X. Lu, X.H. Bian, Y. Hu, T. Cheng, C.C. Yin, W.K. Zhang, S.Y. Chen, and J.S. Zhang, Nuclear factor Y subunit GmNFYA competes with GmHDA13 for interaction with GmFVE to positively regulate salt tolerance in soybean. *Plant Biotechnol J*, 2021. **19**(11): p. 2362-2379.

36.   Kumari, A., V. Dogra, R. Joshi, and S. Kumar, Stress-responsive cis-regulatory elements underline podophyllotoxin biosynthesis and better performance of *Sinopodophyllum hexandrum* under water deficit conditions. *Front Plant Sci*, 2021. **12**: p. 751846.

37.   Raza, A., W. Su, A. Gao, S.S. Mehmood, M.A. Hussain, W. Nie, Y. Lv, X. Zou, and X. Zhang, Catalase (CAT) Gene Family in Rapeseed (*Brassica napus L.*): Genome-wide analysis, identification, and expression pattern in response to multiple hormones and abiotic stress conditions. *Int J Mol Sci*, 2021. **22**(8).

38.   Bian, X.H., W. Li, C.F. Niu, W. Wei, Y. Hu, J.Q. Han, X. Lu, J.J. Tao, M. Jin, H. Qin, B. Zhou, W.K. Zhang, B. Ma, G.D. Wang, D.Y. Yu, Y.C. Lai, S.Y. Chen, and J.S. Zhang, A class B heat shock factor selected for during soybean domestication contributes to salt tolerance by promoting flavonoid biosynthesis. *New Phytol*, 2020. **225**(1): p. 268-283.

39.   Bailey-Serres, J., T. Fukao, D.J. Gibbs, M.J. Holdsworth, S.C. Lee, F. Licausi, P. Perata, L.A. Voesenek, and J.T. van Dongen, Making sense of low oxygen sensing. *Trends Plant Sci*, 2012. **17**(3): p. 129-38.

40.   Ambros, S., M. Kotewitsch, P.R. Wittig, B. Bammer, and A. Mustroph, Transcriptional response of two *Brassica napus* cultivars to short-term hypoxia in the root zone. *Front Plant Sci*, 2022. **13**: p. 897673.

41.   Hong, B., B. Zhou, Z. Peng, M. Yao, J. Wu, X. Wu, C. Guan, and M. Guan, Tissue-specific transcriptome and metabolome analysis reveals the response mechanism of *Brassica napus* to waterlogging stress. *Int J Mol Sci*, 2023. **24**(7).

42.   Chen, T., X. Chen, S. Zhang, J. Zhu, B. Tang, A. Wang, L. Dong, Z. Zhang, C. Yu, Y. Sun, L. Chi, H. Chen, S. Zhai, Y. Sun, L. Lan, X. Zhang, J. Xiao, Y. Bao, Y. Wang, Z. Zhang, and W. Zhao, The genome sequence archive family: toward explosive data growth and diverse data types. *Genomics Proteomics Bioinformatics*, 2021. **19**(4): p. 578-583.

43.   Members, C.-N. and Partners, Database resources of the National Genomics Data Center, China National Center for Bioinformation in 2022. *Nucleic Acids Res*, 2022. **50**(D1): p. D27-D38.