

Genome Identification and *CtCP1* Gene expression of Cysteine Protease in *Carthamus Tinctorius* under Abiotic Stress

Yingqi Hong

Jilin Agricultural University <https://orcid.org/0000-0001-8450-1571>

Lv Yanxi

Jilin Agricultural University

Zhang Jianyi

Jilin Agricultural University

Nguyen Quoc Viet Hoang

Jilin Agricultural University

Li Youbao

Jilin Agricultural University

Wang Nan

Jilin Agricultural University

Liu Xiuming

Jilin Agricultural University

Yao Na (✉ yaona1103@126.com)

Jilin Agricultural University <https://orcid.org/0000-0002-0205-7882>

Research

Keywords: CP gene, *Carthamus tinctorius*, abiotic stress, overexpression and inhibition of expression

Posted Date: June 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-653292/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

Cysteine protease (CP) plays an important role in plant senescence. In this study, the whole genome evolution information of CPs was analyzed by using safflower, and the function of CtCP1 under abiotic stress was analyzed.

Results

25 CPs members were identified in the safflower genome and divided into 9 subfamilies. Gene structure analysis showed that the possible evolutionary conservatism and functional similarity of the same family members. qRT-PCR at different florescence showed that the expression of CPs gene was the highest in the decline period, and CtCP1 gene changed significantly under abiotic stress. We cloned the qRT-PCR of CtCP1, at different florescence and stress, which showed that the expression of CtCP1 was the highest in the decline stage and low temperature. In order to study the function of CtCP1 gene, we obtained the overexpression CtCP1 line (OE) and the inhibition CtCP1 expression line (Anti) in *Arabidopsis thaliana*. The results of transgenic lines under low temperature stress showed that inhibition of CtCP1 expression enhanced the resistance of *Carthamus tinctorius* to low temperature, and overexpression of CtCP1 weakened the resistance of *Carthamus tinctorius* to low temperature.

Conclusion

We have identified the cysteine protease genome of safflower and CtCP1 gene expression under abiotic stress

Introduction

Safflower (*Carthamus tinctorius* L.), belongs to Compositae, is an annual or biennial herb, also known as safflower, safflower or sub-safflower and so on [1]. It is found that up to now, more than 200 compounds have been isolated from *Carthamus tinctorius*, among which more than 104 compounds have been identified. these compounds are: flavonoids, fatty acids, alkaloids, polyacetylene, spermidine, lignans, sesquiterpenes, sterols, polysaccharides and other components [2]. Modern pharmacological experiments also show that *Carthamus tinctorius* and its active compounds have a wide range of biological activities, which can dilate coronary arteries and improve myocardial ischemia. Regulate the immune system, anticoagulation and antithrombotic, antioxidant, hypoxia tolerance, anti-fatigue, anti-liver fibrosis, anti-tumor, anti-inflammation, analgesia and so on [3, 4]. Because safflower seed oil is rich in unsaturated fatty acids, it can be used as edible oil, especially linoleic acid, with a content of about 80% in safflower seeds, which is called "king of linoleic acid" [1]. Therefore, in many countries, such as India, Mexico, the United States, Ethiopia, Spain and Australia, safflower seed oil has become a common plant edible oil [5].

Cysteine protease (EC3.4.22), also known as mercaptan protease, is a large class of proteolytic enzymes. The international classification of proteases is based on the active center of the enzyme. The active site of cysteine protease contains nucleophilic Cys residues, which together with the other two amino acids form

Cys-His-Asn/Glu catalytic triad to catalyze the hydrolysis of corresponding proteins[6, 7]. Cysteine protease exists in animals, plants and microorganisms, and is widely involved in a variety of life activities. Plant cysteine proteases are divided into 5 families (CA, CD, CE, CF, PC) and 15 families[7].With the sequencing of the genomes of a large number of species, cysteine proteases in plant genomes have been reported in Arabidopsis, citrus, papaya, rubber and other dicotyledonous plants and rice. Cysteine proteases are very conservative among different species. in the process of evolution, cysteine proteases are divided into several subfamilies. As early as 2004, Beers and his colleagues identified and classified 138 cysteine proteases in plants. According to the clustering characteristics, these proteases were divided into 8 subfamilies [8]. Richau et al divided 723 species of cysteine proteases into 9 subfamilies in 2012. The results of the two classifications were similar. The latter subdivided the subfamily RD21 into subfamilies RD21 and XBCP3 [9] according to sequence characteristics. At present, the cysteine protease gene family is generally divided into 9 subfamilies, named RD21, CEP, XCP, XBCP3, THI, SAG12, RD19, ALP and CTB9 respectively. Its Central Asian family RD21, CEP, XCP, XBCP3, THI, SAG12 is named as the cath L-like subfamily, while the subfamilies RD19, ALP and CTB9 are named as the cath Fmurike subfamily cath Hliklippe cath B-like subfamily. The subfamily cath Lmurike in plants, cathepsin L, F, H, B, correspond to the subfamilies cathepsin L, F, H, B [10] in animals, respectively.

Similar to caspases, which regulates apoptosis in animals, cysteine proteases in plant cells are involved in programmed cell death (programmed cell death, PCD) [11, 12]. Arabidopsis thaliana XCP1 and XCP2 are specifically expressed in xylem and control catheter programmed death [13]. In the process of pollen development, the tapetum needs to be degraded during pollen maturation, and the tapetum needs to undergo programmed death to release nutrients and ensure the synthesis of the pollen wall. CEP1 is a cysteine protease involved in programmed tapetum cell death and pollen development. Cep1 mutant tapetum cells can not be degraded, pollen fertility decreased, pollen exine abnormal [14]. Song Liping cloned two cysteine proteases Bna C.CP20.1 and Bna C. CP13.4 in Brassica napus and used tapetum-specific promoters to initiate the expression of these two genes. Transgenic plants did not pollinate due to premature tapetum degradation at tetrad stage. Therefore, it is very important for cysteine protease to participate in tapetum PCD during tapetum degradation during pollen development. In tobacco, programmed cell death of embryo sac stalk is regulated by CP14 and its suppressor gene CYS. Knockout of CP14 leads to delay of PCD of embryo sac stalk, while inhibition of CYS expression leads to advance of [15] embryo sac stalk PCD. It is reported that tobacco Cathepsin B is necessary for allergic reaction and resistance caused by non-host pathogens [16–18]. Arabidopsis ctb mutants can reduce programmed cell death induced by biological stress [19]. Further studies have shown that silencing Cathepsin B reduces ROS accumulation and endoplasmic reticulum stress-induced programmed death (ER-stress-induced PCD, ERSID)[20].

Plant senescence is a physiological phenomenon caused by internal and environmental factors. Senescence will be accompanied by the degradation of chlorophyll, the decline of photosynthesis, the degradation of macromolecules, the redistribution of nutrients, and the degradation of cells until the final death. Protein degradation, especially the degradation of target proteins in chloroplasts, is an important hydrolysis process in plant leaf senescence. A member of the SAG12,cath L-like subfamily in Arabidopsis thaliana is currently a marker gene for leaf senescence, [21]. Under normal conditions, the phenotype of sag12 mutant did not change significantly, but the yield of sag12 mutant decreased under low nitrogen condition, indicating that

sag12 was involved in the reuse of nitrogen in seeds [22]. SAG12 and RD21A in rape were reported to be related to protein degradation and nitrogen limitation in senescent leaves [23, 24]. There are 17 Arabidopsis SAG12 homologous genes [25] in rubber, in which Hb SAG12H1 is used as a molecular marker of senescence. Six genes Hb RD21B, Hb RD21E, Hb RD21F, Hb CEP1, Hb XBCP3L, Hb RD19B [26] which play an important role in leaf senescence were identified by deep sequencing in papaya. In Arabidopsis thaliana, XCP1 and XCP2 are associated with senescence, and XCP1 is specifically expressed in xylem, which leads to a decrease in chlorophyll [27] when it is ectopic in leaves. Barley is a model plant for the study of leaf senescence in monocotyledons. Hv Pap-1 accumulates [28] in senescent flag leaves and primary leaves induced by darkness. Van Wyk et al identified a large number of papaya cysteine protease [29] during the senescence of soybean nodules. Eggplant SmCP expressed [30] in senescent leaves, fruits, xylem and anthers.

Plants always encounter abiotic stress in the process of growth, and the recovery and utilization of protein through protein hydrolysis is the main line of defense for plant survival. A large number of studies have shown that cysteine protease up-regulates the expression of cysteine proteases in plants under environmental stress. AtRD21A and AtRD19A are markers of dehydration response and up-regulate the expression of [31] after drought and salt stress. Under PEG, salt and cold stress, TaCP expression in wheat was up-regulated by [32]. Arabidopsis thaliana overexpressing Ta CP was more resistant to drought than wild type under drought stress, suggesting that TaCP plays an important role in response to drought. Sweet potato SPCP2 is expressed in naturally senescent leaves, and darkness, ABA, JA and ethephon can also induce the expression of SPCP2 [33]. Arabidopsis thaliana transformed with SPCP2 is more resistant to salt and drought than wild type [33]. Sweet pepper CaCP expression was induced in senescent leaves, and was significantly up-regulated under biotic and abiotic stress, which inhibited CaCP expression and enhanced plant resistance to salt and osmotic stress [34].

Based on the data mining and processing of the whole genome of *Carthamus tinctorius*, 25 cysteine proteases were identified in *Carthamus tinctorius*. According to the evolutionary relationship, these members were divided into 9 subfamilies. We further analyzed the basic properties of the protein, the conserved domain, the gene family structure, the cis-acting element on the promoter, the expression pattern of the gene at different flowering stages and the changes of gene expression under stress, and cloned a gene *CtCP1* that responded to low temperature. Inhibition of *CtCP1* gene expression enhanced the resistance of transgenic Arabidopsis to low temperature stress, while overexpression of *CtCP1* gene reduced the resistance of transgenic Arabidopsis thaliana to low temperature stress. As a hydrolase, cysteine protease may play a broad-spectrum resistance role in plant abiotic stress. In this study, the use of genome-wide biological analysis and transgenic materials are helpful to analyze the molecular mechanism of the effect of cysteine protease on plant freezing resistance, and provide theoretical guidance and genetic resources for the breeding of freeze-resistant safflower varieties.

Materials And Methods

Plant Materials

In this study, the safflower variety "Jihong 1" was used as material. The seeds were planted in the experimental field of Jilin Agricultural University in Changchun, China in late March 2019 until they blossomed. The petals of *Carthamus tinctorius* are aseptically picked at different stages of flower development, placed in tin foil, immediately placed in liquid nitrogen, and then stored at -80°C for next use. The carrier pCAMBIA3301 comes from our laboratory.

For different stress treatments of *Carthamus tinctorius*

H₂O₂ induced expression treatment: 100mm H₂O₂ solution was irrigated into wild-type *Carthamus tinctorius*, and samples were taken after 24 hours. Low temperature induced expression treatment: the wild-type safflower was treated in a constant temperature incubator at 4 °C, and the sampling was the same as above. Drought-induced expression treatment: the roots of wild-type tomatoes were soaked in PEG 6000 solution with a mass concentration of 20%, and the samples were the same as above. NaCl induced expression treatment: 100mm NaCl solution was irrigated into wild-type safflower, and the samples were the same as above.

Whole Genome Identification of *Carthamus tinctorius* CPs Family

The whole genome sequence of *Carthamus tinctorius* is derived from NCBI (<https://www.ncbi.nlm.nih.gov>), accession number: (PRJNA399628; submitted on Aug.23,2017). The whole genome data of *Arabidopsis thaliana* come from <http://www.arabidopsis.org/>. Based on the 36 amino acid sequences reported in *Arabidopsis thaliana* CPs, the candidate sequence of safflower CPs was obtained by BLASTP alignment in *Carthamus tinctorius*, and the CPs protein sequence was further screened by the conserved domain of cysteine protease.

Analysis of phylogenetic Tree, Gene structure, Promoter, Conservative sequence and protein characteristics of *Carthamus tinctorius* CPs Gene

The CPs sequences obtained by Clustal X1.83 were multi-aligned, and the following three kinds of sequences were deleted: (1) sequences without CWF, (2) sequences with less than 50 amino acids from the N-terminal of CWF, and (3) sequences with no more than 150 amino acids from the C-terminal of CWF. The rest of the sequences are constructed by MEGA6 software and based on multiple sequence alignment using adjacency method (Neighbor-Joining). The parameters are repeated for bootstrap 1000 times, and the other parameters are the default values of the system. According to the genome information, the gene sequence and coding sequence of *Carthamus tinctorius* CPs were found. The conservative sequence was predicted by software GSDS2.0 (<https://gsds.cbi.pku.edu.cn/>) and online software MEME (<http://meme-suite.org/>). The output parameter of motif was set to 15. The basic message of egg white comes from Ex PASy-Prot Param tool (<https://web.expasy.org/protparam>). PlantCare (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) was used to analyze the distribution of conserved cis regulatory elements in the flanking sequence of the 5' non-coding region upstream of the 2kb promoter of each CtCP gene.

Prediction of protein-protein interaction Network of CtCPs Gene

The selected safflower CPs protein sequence has been added to the String Version 11 online network tool (<https://string-db.org>) to identify possible protein-protein interaction network predictions. Then, we chose *Arabidopsis thaliana* genome as the comparative model organism. Based on the highest bit score during the analysis, a collective hierarchical protein interaction network of CtCP gene was generated.

RNA extraction, reverse transcription and real-time quantitative PCR analysis (q RT-PCR)

According to the purpose of the experiment, *Carthamus tinctorius* materials with different flowering stages and different stress treatments were frozen in liquid nitrogen and stored at -70 °C. When extracting RNA, we first grind it fully with liquid nitrogen, and then use the RNA extraction kit of *Radix Scutellariae* to operate. The specific steps of RNA reverse transcription are as follows: 1) take 3 µg RNA into 0.5 ml RNA-free centrifuge tube, add 1 µl oligo (dT), then replenish the volume to 15 µl with DEPC water and mix well. 2) the mixture mentioned above was placed in a PCR instrument, thermally denatured at 70 °C for 10 min, to open the secondary structure of RNA, and then quickly placed on the ice to place 5min. 3) the following reagents were added to the mixture: 5 µl 5 × MLV buffer, 1 µl Rnasin, 1 µl M-MLV RTase, 1.25 µl 10 mmol/L d NTP and 1.75 µl DEPC water. After fully mixing, the c-DNA was reacted at 42 °C for 60–90 min, and then at 70 °C for 10 min,. After the reaction, the reversed c-DNA was stored at -20 °C. qRT-PCR analysis: 1) the prepared cDNA was diluted 100x as a template for qRT-PCR analysis. 2) the reaction system was configured according to the following proportion (15 µl): 7 µl c DNA template diluent, 7.5 µl SYBR Green-Mix (Bio-Rad company), forward and reverse primers 0.25 µl, mixed and centrifuged briefly at low speed. 3) the instrument used for qRT-PCR analysis in our laboratory is set up as follows: stage1 95°C 30 sec, 1 cycles, stage2 95°C 30 sec/60°C 35 sec, 40 cycles, and read the fluorescence intensity value when the step is completed in each round. 4) after the completion of the reaction, 18S RNA was used as the internal reference gene to analyze the results. The primers for qRT-PCR analysis used in this study are shown in Table S1.

Extraction and cDNA Cloning of CtCP1

The flower petals from the JH1 cultivar of *C. tinctorius* were used as a source of mRNA extraction in order to amplify the full-length cDNA sequence of the candidate *CtCP1*. The total RNA content was collected from the preserved flower petals using RNAiso Plus reagent (TaKaRa Bio Co., Beijing, China). The Reverse transcription PCR system was exploited to synthesize the first strand cDNA templates using the reagent kit of reverse transcriptase superscript IV (Thermo Fisher Scientific). The PCR amplification was performed with gene specific oligos of forward: CP-1F:5'ATGAAGCTCCTCCTTTCCATGG3' and reverse: CP-1R:5'TCAAACACTGCTCCTCTTTCGC' primers. The target PCR fragment of the full length cDNA sequence of *CtCP1* was successfully cloned using Pfu DNA polymerase (Takara, Beijing, China). The amplified product of *CtCP1* was initially confirmed on gel electrophoresis and then transformed into the cloning sites of pEASY-T1 vector (Takara, Dalian, China), for further confirmation through sequencing to detect possible base mutation.

Floral dip Transformation and expression analysis of CtCP1

The overexpression vector of pCAMBIA3301 containing the selectable marker gene (BAR) under the control of the 35S (CaMV) promoter was constructed following the sticky end ligation of the BamHI/EcoRV digested

CtCP1 full-length fragment into the expression cassettes of pCAMBIA3301 vector using T4 ligase following the instructions given by 87. The recombinant plasmid produced after ligation (pCAMBIA3301-*CtCP1*) was primarily confirmed with double restriction digestion and then transformed into DH5α cells of *E. coli* using the traditional heat-shock protocol. The half colony PCR positive bacterial cells were then confirmed with Sanger sequencing for sequence integrity. The recombinant plasmid generating (pCAMBIA3301-*CtCP1*) construct under the control of 35S viral promoter and the empty plasmid of pCAMBIA3301 alone was inserted into *A. tumefaciens* (EHA105) competent cells. The positive colonies of EHA105 were selected on YEP + antibiotic growth medium and further confirmed by half colony PCR method. About 6-week-old Arabidopsis was transformed by EHA105 containing the binary vectors over-expressing *CtCP1* following the floral dip method. In the qRT-PCR assay, three different types of experimental lines were selected where wild type, *CtCP1*-Anti and one *CtCP1* independent transgenic overexpressed (OE) line of Arabidopsis were used. The total RNA content from the wild type, *CtCP1*-Anti line and *CtCP1*-OE line was simultaneously extracted using the previously described protocol. The cDNA templates were prepared using the PrimeScript RT reverse transcription kit (Takara, Beijing) following the given protocol of the manufacture. The qRT-PCR analyses for *AtCP1*(At1g47128.1), was conducted in wild-type, *CtCP1*-Anti line, and *CtMYB63*-OE transgenic line under the same conditions. The relative fold expression change was standardized using the reference gene of 18s ribosomal RNA gene (AT5G38720.1). The relative expression of each transcript was measured using the $2^{-\Delta\Delta Ct}$ formula. The primers used in this study are listed in Table S1.

Cold stress Treatment

The three Arabidopsis seedlings of WT, *CtCP1*-Anti and *CtCP1*-OE lines were uniformly elicited with cold stress simultaneously according to the previously described method 89. The cold stress induction was introduced in 7-day-young Arabidopsis seedlings by gradient cooling provided in the incubator for a period of 3 days. The survival rate of these selected Arabidopsis seedlings after the elicitation of the cold stress treatment was observed thereafter.

Antioxidant Enzymatic Activity

For the antioxidant enzyme assay, the superoxide dismutase (SOD), catalase (CAT) and peroxidase (POD) activity were measured by using Beijing Soleibao Technology Co., Ltd. kit. For SOD, CAT and POD activity, BC0170, BC0200, and BC0090 kits were purchased, respectively. Five-hundred milligrams of leaves were crushed into powder in liquid nitrogen, and the following procedure was conducted, according to manufacturer's instructions for SOD, CAT and POD determination.

Determination of malondialdehyde, proline and soluble sugar

Malondialdehyde (MDA), proline (Pro) and soluble sugar content (TSS) were determined by BC0020, BC0290 and BC0030 kits of Beijing Soleibao Technology Co., Ltd. Homogenize 100 mg of leaves according to the producer's instructions.

Statistical analysis

All our findings were presented as mean \pm S.D. with three independent biological replicates. Differences between means of each group were assessed by one-way analysis of variance using Statistics 8.1 software. P-values equal to 0.05 was kept statistically significant.

Results

Identification and Bioinformatics Analysis of Safflower CP Family Gene

Using 36 known Arabidopsis CP family gene sequences as templates, the safflower CP family candidate genes were searched in the safflower genome, and 25 safflower CP genes were identified by domain analysis and transcriptional level determination (Table 1). Through the prediction and analysis of various online websites, the length of the protein encoded by CtCP is 85-660bp, the molecular weight MW is about 9.15kDa-71.66kDa, and the isoelectric point pI is between 4.33kDa and 7.87kDa. The prediction of subcellular localization shows that except for CtCP04,CtCP07,CtCP09,*CtCP16* and CtCP25 located in endoplasmic reticulum, most of the other proteins are located in extracellular space, while *CtCP13* is located in the cytoplasm.

Table 1
Analyses of *CtCP* gene family members

Gene Name	Gene ID	Protein Length	PI	MW(kDa)	Subcellular Localization	Arabidopsis Homology	GRAVY
CtCP01	CCG012229.1	433	4.90	47.51	Extracellular	AT1G47128.1	-0.397
CtCP02	CCG020600.1	341	4.60	37.62	Extracellular	AT3G49340.1	-0.389
CtCP03	CCG027222.1	366	5.21	39.94	Extracellular	AT4G39090.1	-0.260
CtCP04	CCG028374.1	364	6.58	40.49	Endoplasmic reticulum	AT1G47128.1	-0.466
CtCP05	CCG002592.1	660	5.29	71.66	Extracellular	AT2G34080.1	-0.456
CtCP06	CCG007611.1	467	4.71	51.34	Extracellular	AT1G47128.1	-0.476
CtCP07	CCG010912.1	509	4.98	55.70	Endoplasmic reticulum	AT1G47128.1	-0.457
CtCP08	CCG011149.1	85	5.04	9.15	Extracellular	AT4G11310.1	-0.173
CtCP09	CCG012686.1	505	6.15	55.39	Endoplasmic reticulum	AT2G34080.1	-0.588
CtCP10	CCG016643.1	343	4.59	37.61	Extracellular	AT2G34080.1	-0.387
CtCP11	CCG016644.1	343	4.84	37.76	Extracellular	AT2G34080.1	-0.419
CtCP12	CCG016645.1	325	4.63	35.69	Extracellular	AT3G49340.1	-0.484
CtCP13	CCG022433.1	464	5.21	51.23	Cytoplasmic	None	-0.376
CtCP14	CCG023081.1	422	6.27	46.81	Extracellular	AT1G47128.1	-0.253
CtCP15	CCG023118.1	353	6.10	37.77	Extracellular	AT4G39090.1	-0.137
CtCP16	CCG025047.1	303	4.95	33.40	Endoplasmic reticulum	AT2G34080.1	-0.359
CtCP17	CCG025048.1	94	7.87	10.66	Extracellular	AT2G27420.1	0.199
CtCP18	CCG027053.1	354	6.29	39.71	Extracellular	AT1G47128.1	-0.346
CtCP19	CCG030007.1	90	4.85	9.89	Extracellular	AT3G12490.2	-0.447
CtCP20	CCG030009.1	91	5.04	10.35	Extracellular	AT2G40880.1	-0.281
CtCP21	CCG000752.1	340	4.51	37.04	Extracellular	AT2G27420.1	-0.307
CtCP22	CCG000753.1	296	4.33	32.37	Extracellular	AT2G34080.1	-0.281
CtCP23	CCG001395.1	340	4.51	37.04	Extracellular	AT2G27420.1	-0.307
CtCP24	CCG001396.1	307	4.46	33.39	Extracellular	AT3G49340.1	-0.418
CtCP25	CCG015538.1	364	6.58	40.49	Endoplasmic reticulum	AT1G47128.1	-0.466

Analysis of conserved motif, structure and Promoter of *Carthamus tinctorius* CPs Family

CPs is a large group of gene families in plants. In order to understand the classification of CPs in safflower, a phylogenetic tree was constructed based on the amino acid sequences of *Arabidopsis* and safflower (Fig. 1). The phylogenetic tree showed that the CPs of *Arabidopsis thaliana* and *Carthamus tinctorius* could be divided into 9 types. According to the nomenclature principle of CPs protein in *Arabidopsis thaliana*, safflower CPs protein was named as RD21, CEP, XCP, XBCP3, THI, SAG12, RD19, ALP and CTB. The specific gene naming is shown in Table 1. There is an imbalance in the number of members among different families. There are 5 members of CEP subfamily, 2 members of XCP and CTB subfamily, 1 member of ALP and XBCP3 subfamily, and 12 members of SAG12 subfamily. However, there are no safflower members in RD21 and RD19 subfamily. We speculate that the subfamily gene is lost in the process of species evolution. We further analyzed the conserved motif of safflower CPs protein and predicted 15 conserved motif. Motif13, 6, 7 and 12 constitute the self-inhibitory domain (PF08246) of papain. The sequence of motif6 in RD21, CEP, XCP, XBCP3, THI, SAG12 and ALP7 subfamily is ERFNI/VN. The sequence of motif6 in RD19 subfamily is ERFNAQ, but the sequence of CTB subfamily does not contain motif6-like sequence. Motif1, 3, 4, 10, 5, 14, 8, 2 and 9 form a Peptidase_C1 domain (PF00112). Motif1, 8 and 2 contain three hydrolysis sites of Cys, His, Asn, respectively. Most family members contain these three hydrolysis sites, but among the two members of CtCP04 and CtCP25, Gln and Arg replace His respectively. Motif15 is unique to the RD19 subfamily. Motif11 forms granule protein domain (PF00396). The granulin domain exists only in XBCP3 and 8 members of the RD21 subfamily (Fig. 2A).

We used the software GSDS2.0 (<https://gsds.cbi.pku.edu.cn/>) to analyze the gene structure of CP. There are differences in gene length among different members, but the number of introns and exons is similar among members of the same subfamily. The CEP subfamily has 3–4 introns, the XCP subfamily has 3–6 introns, the XBCP3 subfamily has 1 intron, the SAG12 subfamily has 1 intron, most of the ALP subfamilies have 7 introns, and the CTB subfamilies have 1–2 introns (Fig. 2B).

The promoter regulates the transcriptional level of genes and controls the expression patterns of genes in different tissues and environments. We downloaded the upstream 2K sequence of safflower CPs gene and analyzed these sequences by promoter prediction software PlantCARE. It was found that there were many elements involved in abscisic acid response, light response, anaerobic induction, methyl jasmonate response, endosperm expression, gibberellin response and translation initiation in the promoter region of safflower CPs gene (Fig. 2C). This indicates that safflower CPs gene is involved in the regulation of safflower hormone signal and abiotic response.

Protein interaction Networks of CtCP Genes

In addition, to identify the predicted protein interaction network of CtCP genes and their associated proteins, an in silico prediction system was generated by exploiting the online web-server of STRING using the *Arabidopsis* orthologous proteins (Fig. 3). In the CtCP derived network, it was found that a total of 20 identified proteins were interacted actively. Among these, the ubiquitin-like superfamily proteins (APG8A,

APG8H, ATG8H, APG7, ATG8F, ATG8C, ATG8G, ATG8B, ATG12A and APG12B); autophagy-related proteins (PEUP1,ATG8E,ATG3,ATG13,AT3G18770,ATG18B,ATG6 and AT2G37840); Disease resistance protein (RRS1) and senescence associated protein (SAG13) were most abundantly interacted with the putative CtCP orthologous. These proteins are necessary for autophagy, which plays an important role in the plant nutritional cycle. The interaction of CtCPs gene with these proteins participates in the early stage of autophagy biogenesis, participates in the negative regulation of plant defense response to biological nutritional pathogens, participates in a negative feedback loop, regulates NPR1-dependent salicylic acid (SA) signal, restricts plant senescence and immunity-related programmed cell death (PCD); and participates in the degradation of damaged peroxisome. These findings suggest that they may be related to autophagy and senescence of plant cells, combined with biotic and abiotic stress sensitivity, supporting the positive role of CtCPs in the pathway of plant senescence.

Expression pattern of CPs gene family in *Carthamus tinctorius* under different flowering stages and abiotic stress.

Based on the tissue expression profile data of safflower CPs gene, it was detected that all 25 safflower CP genes were expressed at different flowering stages (Fig. 4). This shows that most of the CP genes in safflower are expressed at different stages of petal development. It can be seen from Fig. 4 that most safflower CPs genes are highly expressed during petal decline, except for CtCP3 and CtCP19. It is suggested that the expression of CPs gene in safflower is related to the senescence of plant tissues and organs.

According to previous studies, CP gene may play a role in a variety of adverse environments. We used transcriptome data to analyze the expression changes of these CP genes under different abiotic stresses such as H₂O₂, low temperature, salt and drought. Therefore, we selected a relatively high expressed CtCP1 gene in the fading stage of *C. tinctorius* and verified whether the expression level of this putative gene was fluctuated upon hormonal and abiotic stress treatments at variable time-points by carrying out qRT-PCR assays. As described in Fig. 5, under the influence of external cold treatment, the expression level of CtCP1 gene was slightly increased at 2h time-point, and then significantly up-regulated at 8h time-point and beyond than in control treatment (0h). Following the induction of external PEG, the relative fold change in the expression level of CtCP1 was differentially regulated at different time-points (Fig. 5). Initially, the expression of CtCP1 under PEG induction was down-regulated at 2h after the treatment and then upregulated at 4h time-point and then again decreased at 6h time-point as compare to control treatment (Fig. 5). However, the transcription level of CtCP1 was significantly enhanced and upregulated at 8h and beyond than those treated without PEG induction. Similarly, under continuous H₂O₂ treatment, the expression of CtCP1 gene was slightly up-regulated at 2h after the treatment and then reached to its maximum and peaked at 4h time-point. However, the expression level of CtCP1 was then again down-regulated at 6h and 8h respectively as compare to control treatment (Fig. 5). In contrast, under NaCl treatment, the relative expression level of CtCP1 gene was somewhat increased at 2h time-point but immediately down-regulated at 4h and 6h after the treatment and then interstingly upregulated significantly at 8h time-point than in control treatment (Fig. 5). The general expression trend in the NaCl environment was slightly opposite to that of H₂O₂ condition. Our research findings from these analyses reinforce the evidence that CtCP1 may correspond to the

transcriptional reprogramming molecular mechanism through multiple stages in *C. tinctorius* under various stress conditions.

Cloning of CtCP1 Gene from *Carthamus tinctorius* and obtaining transgenic plants of *Arabidopsis thaliana*

Safflower CP gene plays an important role in various abiotic stresses. We screened a number of differentially expressed candidate genes in four flowering stages and low temperature stress of *Carthamus tinctorius*, among which a *CtCP1* gene encoding papaya cysteine protease was significantly

differentially expressed during petal decline and after low temperature stress. The expression of this sequence increased by 2.5 times after low temperature stress, so we speculate that this gene plays an important role in response to low temperature.

In order to further study the biological function of *CtCP1* in safflower, we constructed an overexpression vector and an antisense vector of *CtCP1*-OE driven by the 35S promoter, and obtained transgenic plants that overexpress and inhibit the expression of *CtCP1*. We selected a *CtCP1* over-expressing line (*CtCP1-e* and *CtCP1-d*) with a good expression level for subsequent experimental research. The expression of the overexpressing *CtCP1*-OE line was significantly higher than that of the WT control (Fig. 6A), and the expression of the suppressed *CtCP1*-Anti line was significantly lower than that of the WT control (Fig. 6B).

Changes of phenotypic and physiological Indexes of CtCP1 Gene in *Carthamus tinctorius* under low temperature stress

From the previous results of qRT-PCR we can find that *CtCP1* is significantly up-regulated under low temperature treatment which indicates that *CtCP1* may play a role in the response to low temperature stress (Fig. 5). We treated wild-type *Arabidopsis thaliana* *CtCP1*-Anti and over-expressed *Arabidopsis thaliana* *CtCP1*-OE in 1/2MS medium at low temperature (-8°C) to observe their phenotypes and determine their physiological indexes. The experimental results (Fig. 6C) show that there is no difference between the first three materials before low temperature treatment. After treatment, the inhibition rate of plant survival by functionally deficient *CtCP1*-Anti was as high as 70%, which was significantly higher than that of wild type WT and overexpressed plant *CtCP1*-OE (30% and 20%, respectively, Fig. 6D). It is suggested that *CtCP1* gene may be a negative regulator of low temperature stress response of *Carthamus tinctorius*.

Cell membrane peroxidation occurs in plants under stress conditions, such as low temperature and drought, and the content of MDA can reflect the peroxidation of plant cell membrane. By measuring MDA, it is found that there is no significant difference in MDA content among the three materials under normal conditions, but the MDA content of plants inhibited by *CtCP1*-Anti after low temperature treatment is significantly lower than that of WT and *CtCP1*-OE (Fig. 7D), indicating that under low temperature treatment, the MDA content of plants inhibited expression is significantly lower than that of WT and *CtCP1*-OE (Fig. 7D). The degree of membrane oxidation of expressed plants inhibited by *CtCP1*-Anti was lower than that of WT and *CtCP1*-OE.

In order to explore the reason for the difference of MDA content in different lines after low temperature treatment, we determined the enzyme activities of SOD, CAT and POD. The results showed that there was no difference between SOD and CAT before low temperature treatment. After low temperature treatment, the enzyme activity of overexpression lines was lower than that of wild type and antisense expression lines (Fig. 7A and C); while POD activity of overexpression lines before and after low temperature treatment was higher than that of wild type and higher than that of antisense expression lines (Fig. 7B). This shows that inhibiting the expression of *CtCP1* gene can maintain the enzyme activity of SOD and CAT and increase the scavenging ability of MDA after low temperature stress in safflower, thus improving the tolerance of plants to low temperature.

When plants are subjected to low temperature stress, the osmotic regulatory substances in the body increase, regulate cell concentration, reduce osmotic potential, maintain normal cell water content, protect cell membrane and protein, and increase cold tolerance. Therefore, we determined the contents of free proline and soluble sugar in plants. The results showed that there was no difference in the contents of free proline and soluble sugar between transgenic lines and wild-type lines before treatment. After low temperature treatment, the contents of proline and soluble sugar in over-expressed lines were significantly lower than those in wild-type lines, while those in antisense expression lines were higher than those in wild-type lines (Fig. 7E, F), which indicated that the expression of *CtCP1* could affect the accumulation of free proline and soluble sugar in plants. To sum up, the results of freeze resistance analysis showed that the freezing resistance of plants inhibited by *CtCP1*-Anti was stronger than that of WT and *CtCP1*-OE overexpressed plants, and *CtCP1* gene played a negative role in the freezing response of plants.

Discussion

Identification and Analysis of CPs Gene Family in *Carthamus tinctorius*

Cysteine protease is a large family of plants, and the whole genome identification of its family members has been reported in *Arabidopsis thaliana* [9], rice [35], citrus [36], castor [37], jatropha [37], papaya [38] and rubber [26]. In this study, we identified 25 safflower CPs, which were divided into 9 subfamilies according to evolutionary relationship, which was consistent with the reports in other species. But the number of CPs members varies greatly from species to species. There are 25 members in safflower, 36 in *Arabidopsis*, 33 in rice, 21 in citrus, 26 in castor, 23 in jatropha, 33 in papaya and 43 in rubber. In addition, the number of members in different subfamilies also varies greatly. Genome-wide replication, tandem replication and large-scale fragment replication may lead to significant changes in the number of genes in each subgroup. Liu et al. [39] revealed that tandem replication plays an important role in papaya specific subfamily amplification. There are 25 CPs members in *Carthamus tinctorius*, which may be due to the fact that *Carthamus tinctorius* is autodiploid [2]. Furthermore, through multiple alignment analysis, 11 pairs of homologous genes showed very consistent evolution in exon-intron structure, conservative motif and gene expression pattern.

In animals, CPs is classified into four categories: cathepsin-L, cathepsin-F, cathepsin-H and cathepsin-B according to the N-terminal motif. "ERFNIN" is the iconic motif of cathepsin-L and cathepsin-H [40, 41], and

"ERFNAQ" is the iconic motif of cathepsin-F [41]. According to the characteristics of N-terminal motifs, RD21, CEP, XCP, XBCP3, THI and SAG12, RD19, ALP, CTB subfamilies in plants correspond to cathepsin-L, cathepsin-F, cathepsin-H and cathepsin-B in animals, respectively. Granulin was first found in animals and reported to have many biological functions [42]. C-terminal extension (Cx5Cx5CCCx7Cx4CCx6CCx5CCx6Cx6C) similar to granulin domain was also found in RD21 and XBCP3 subfamilies. However, not all members of the RD21 subfamily have granule protein domains, which implies the loss of granulin polymorphism in the process of evolution [9].

Expression Analysis of CPs Gene Family in *Carthamus tinctorius* in response to stress

Abiotic stress and biotic stress seriously affect the growth, development and yield of *Carthamus tinctorius*. In recent years, more and more molecular mechanisms of safflower response to different stresses have been found, and there are also reports about CP response to different stress treatments in other species. In wheat, the expression level of TaCP reached the peak [43] at 12 h of PEG treatment, 48 h of salt treatment and 72 h of low temperature treatment. The constitutive expression of sweet potato SPCP2, in *Arabidopsis thaliana* showed more salt tolerance and drought tolerance [44] than the control. The expression of CaCP in pepper was induced by high salt, mannitol and *Phytophthora capsici*. Silencing CaCP, in pepper caused the plant to be more resistant to salt and mannitol stress [34]. However, there is no report on the function of CP in *Carthamus tinctorius*. In this study, we found that the expression of 15 CPs was 2-fold different from that of the control after H₂O₂, low temperature, salt and drought treatments. And the expression of these 9 genes changed significantly after low temperature treatment, indicating that CPs is more sensitive to temperature changes. Two genes were differentially expressed under H₂O₂, salt and drought stress, and seven genes responded to at least two stress treatments. All these indicate that CPs plays an important role in responding to abiotic stress of *Carthamus tinctorius*.

Physiological Mechanism of CtCP1 regulating low temperature Resistance of *Carthamus tinctorius*

After low temperature treatment of transgenic lines and wild-type lines, we found that the survival rate of *CtCP1* overexpression lines was the lowest, followed by wild-type lines, and antisense expression lines had the highest survival rate (Fig. 6). Based on the phenotype, we explored the physiological mechanism of the difference in cold tolerance between transgenic lines and wild-type lines. When plants are subjected to low temperature stress, excessive MDA will be accumulated in the body. However, plants also activate enzymatic and non-enzymatic antioxidant systems to remove excess reactive oxygen species. Tea plants are sensitive to cold stress and show high genetic variability to cold tolerance. Some studies have pointed out that the activation of MDA scavenging genes is the main strategy for tea plants to cope with cold stress in winter [45]. Antioxidant enzymes such as SOD, CAT and POD are the main enzymes for scavenging excess MDA in cells. A variety of genes can improve the stress resistance of plants by regulating the activity of antioxidant enzymes or the expression of coding genes. Overexpression of tomato SINAM1 in tobacco could improve cold resistance [46] by maintaining the activity of antioxidant enzymes and thus reducing the content of MDA, while soybean GmMYB84 could improve the drought resistance of soybean [47] by increasing the activity of antioxidant enzymes and the content of free proline in overexpressed lines. ZmMYB31, a R2R3-

MYB transcription factor, is overexpressed in *Arabidopsis thaliana*. Compared with wild-type lines, overexpressed lines have higher SOD and APX enzyme activities and lower MDA content, thus improving the cold resistance [48] of transgenic *Arabidopsis thaliana*. There are also many genes that regulate MDA content in plants through non-enzymatic antioxidant systems. Wheat TaNAC2 gene was overexpressed in *Arabidopsis thaliana*, and it was found that it could improve *Arabidopsis thaliana* tolerance to low temperature, drought and salt stress [49] by regulating AsA-GSH cycle. After heterologous expression of tomato SINAC2 in *Arabidopsis thaliana*, transgenic *Arabidopsis thaliana* had salt tolerance and drought tolerance, and abundant GSH metabolic genes were found in transcriptome, which led to increased GSH synthesis and reduced oxidative damage [50]. Our results showed that after low temperature treatment, the activities of SOD and CAT, the key enzymes of MDA scavenging in chloroplasts, were the highest in the antisense expression lines and the lowest in the overexpression lines, resulting in the highest MDA content in the *CtCP1* overexpression lines and the lowest in the antisense inhibition lines (Fig. 7), resulting in different cold tolerance of each line.

Osmotic adjustment substances help plants reduce osmotic potential under adversity stress, maintain normal water content, protect proteins and cell membranes, and enable plants to survive. The main osmotic adjustment substances in plants are free proline and soluble sugar. *Elymus europaea* increases the activity of Δ 1-pyrroline-5-carboxylic acid synthase (P5CS) and Δ 1-pyrroline-5-carboxylic acid reductase (P5CR), and reduces proline dehydrogenase (PDH) and Δ 1-pyrroline-5-carboxylic acid dehydrogenase (P5CDH) increases the free proline content in the body under low temperature stress, thereby increasing its own low temperature resistance [51]. We found that after low temperature treatment, although the free proline content in each line increased, the overexpression line increased the least, and the suppressed expression line increased the most (Fig. 7), indicating that *CtCP1* negatively regulates free proline Accumulation of acid.

Conclusion

In this study, safflower was used to analyze the genome-wide evolution information of CPs and the function of *CtCP1* under abiotic stress:

- (1) 25 CPs members were identified in the safflower genome, divided into 9 subfamilies. Gene structure analysis showed that members of different subfamilies have specific structures and conservative motifs, indicating that members of the same family may have evolutionary conservation and functional similarity.
- (2) qRT-PCR at different flowering stages showed that the expression of CPs gene was the highest in the declining stage, and the *CtCP1* gene changed significantly under abiotic stress. We cloned the qRT-PCR of *CtCP1* in different flowering periods and stresses, and the results showed that *CtCP1* expressed the highest in the declining period and low temperature period.
- (3) In order to study the function of *CtCP1* gene, we obtained over-expression *CtCP1* line (OE) and suppressed *CtCP1* expression line (Anti) in *Arabidopsis*. The results of the transgenic lines under low temperature stress showed that suppressing the expression of *CtCP1* enhanced the low temperature resistance of safflower, and overexpression of *CtCP1* weakened the low temperature resistance of safflower.

Abbreviations

CP: Cysteine protease

qRT-PCR: real-time quantitative PCR

Ct: *Carthamus tinctorius* L.

PCD: Programmed cell death

ROS: reactive oxygen species

At: *Arabidopsis thaliana*

CWAF: Cysteine, Tryptophan, Asparagine/ Aspartic acid

SOD: superoxide dismutase

CAT: catalase

POD: peroxidase

MDA: Malondialdehyde

TSS: soluble sugar content

Declarations

Availability of data and materials

Vectors and *Arabidopsis* seeds can be requested from the corresponding authors. The vectors pEASY-T1 and pCAMBIA3301 are also available via Miaoling ([https:// miaolingbio.com](https://miaolingbio.com)).

Authors' contributions

L.X, Y.N and X.L designed the project; H Y. performed the experiments and analyze the data; J.Z, and Y.X carried out formal analysis; N.Q and W.N performed statistical analysis; Y.B conducted software analysis, H.Y wrote the original draft.

Conflicts of Interest:

The authors declare that they have no conflict of interest.

Acknowledgments

We are grateful to Professor Li Haiyan (Hainan University) for providing pCAMBIA3301.

Funding

This work was funded by grants from the National Natural Science Foundation of China (No. 31771868 and 31501366), the Science and Technology Development Project of Jilin Province (No. 20190201172JC and 20190201175JC and 20150623024TC-11), Student innovation and entrepreneurship training program of Jilin Agricultural University.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors agree with submission of this version.

Competing interests

The authors declare that they have no competing interests.

References

1. Zhou X, Tang L, Xu Y, Zhou G, Wang Z. Towards a better understanding of medicinal uses of *Carthamus tinctorius* L. in traditional Chinese medicine: a phytochemical and pharmacological review. *J Ethnopharmacol.* 2014;151:27–43.
2. Yingqi H, Ahmad N, Yuanyuan T, Jianyu L, Liyan W, Gang W, Xiuming L, Yuanyuan D, Fawei W, Weican L, Xiaowei L, Xu Z, Na Y, Haiyan L. Genome-Wide Identification, Expression Analysis, and Subcellular Localization of *Carthamus tinctorius* bHLH Transcription Factors, *Int J Mol Sci*, 20 (2019).
3. Camaş N, Esendal E. Estimates of broad-sense heritability for seed yield and yield components of safflower (*Carthamus tinctorius* L.). *Hereditas.* 2006;143:55–7.
4. Xia J, Zheng M, Li L, Hou X, Zeng W. [Conjugated linoleic acid improves glucose and lipid metabolism in diabetic mice], *Nan fang yi ke da xue xue bao = Journal of Southern Medical University*, 39 (2019) 740–746.
5. Jiang J, He J, Feng Z, Zhang P. Two new quinochalcons from the florets of *Carthamus tinctorius*. *Org Lett.* 2010;12:1196–9.
6. van der Hoorn RA, Kamoun S. From Guard to Decoy: a new model for perception of plant pathogen effectors. *Plant Cell.* 2008;20:2009–17.
7. Grosse-Holz F, Kelly S, Blaskowski S, Kaschani F, Kaiser M, van der Hoorn RAL. The transcriptome, extracellular proteome and active secretome of agroinfiltrated *Nicotiana benthamiana* uncover a large, diverse protease repertoire. *Plant Biotechnol J.* 2018;16:1068–84.
8. Beers EP, Jones AM, Dickerman AW. The S8 serine, C1A cysteine and A1 aspartic protease families in *Arabidopsis*. 65: *Phytochemistry*; 2004. pp. 43–58.
9. Richau KH, Kaschani F, Verdoes M, Pansuriya TC, Niessen S, Stüber K, Colby T, Overkleeft HS, Bogyo M. R.A. Van der Hoorn, Subclassification and biochemical analysis of plant papain-like cysteine proteases displays subfamily-specific characteristics. *Plant Physiol.* 2012;158:1583–99.

10. Martinez M, Cambra I, Carrillo L, Diaz-Mendoza M, Diaz I. Characterization of the entire cystatin gene family in barley and their target cathepsin L-like cysteine-proteases, partners in the hordein mobilization during seed germination. *Plant Physiol.* 2009;151:1531–45.
11. Solomon M, Belenghi B, Delledonne M, Menachem E, Levine A. The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. *Plant Cell.* 1999;11:431–44.
12. Buono RA, Hudecek R, Nowack MK. Plant proteases during developmental programmed cell death. *J Exp Bot.* 2019;70:2097–112.
13. Zhao C, Johnson BJ, Kositsup B, Beers EP. Exploiting secondary growth in Arabidopsis. Construction of xylem and bark cDNA libraries and cloning of three xylem endopeptidases. *Plant Physiol.* 2000;123:1185–96.
14. Zhang D, Liu D, Lv X, Wang Y, Xun Z, Liu Z, Li F, Lu H. The cysteine protease CEP1, a key executor involved in tapetal programmed cell death, regulates pollen development in Arabidopsis. *Plant Cell.* 2014;26:2939–61.
15. Zhao P, Zhou XM, Zhang LY, Wang W, Ma LG, Yang LB, Peng XB, Bozhkov PV, Sun MX. A bipartite molecular module controls cell death activation in the Basal cell lineage of plant embryos. *PLoS Biol.* 2013;11:e1001655.
16. Gilroy EM, Hein I, van der Hoorn R, Boevink PC, Venter E, McLellan H, Kaffarnik F, Hrubikova K, Shaw J, Holeva M, López EC, Borrás-Hidalgo O, Pritchard L, Loake GJ, Lacomme C, Birch PR. Involvement of cathepsin B in the plant disease resistance hypersensitive response. *Plant J.* 2007;52:1–13.
17. McLellan H, Gilroy EM, Yun BW, Birch PR, Loake GJ. Functional redundancy in the Arabidopsis Cathepsin B gene family contributes to basal defence, the hypersensitive response and senescence. *New Phytol.* 2009;183:408–18.
18. Salguero-Linares J, Coll NS. Plant proteases in the control of the hypersensitive response. *J Exp Bot.* 2019;70:2087–95.
19. Ge Y, Cai YM, Bonneau L, Rotari V, Danon A, McKenzie EA, McLellan H, Mach L, Gallois P. Inhibition of cathepsin B by caspase-3 inhibitors blocks programmed cell death in Arabidopsis. *Cell Death Differ.* 2016;23:1493–501.
20. Cai YM, Yu J, Ge Y, Mironov A, Gallois P. Two proteases with caspase-3-like activity, cathepsin B and proteasome, antagonistically control ER-stress-induced programmed cell death in Arabidopsis. *New Phytol.* 2018;218:1143–55.
21. Hensel LL, Grbić V, Baumgarten DA, Bleecker AB. Developmental and age-related processes that influence the longevity and senescence of photosynthetic tissues in Arabidopsis. *Plant Cell.* 1993;5:553–64.
22. James M, Poret M, Masclaux-Daubresse C, Marmagne A, Coquet L, Jouenne T, Chan P, Trouverie J, Etienne P. SAG12, a Major Cysteine Protease Involved in Nitrogen Allocation during Senescence for Seed Production in Arabidopsis thaliana. *Plant Cell Physiol.* 2018;59:2052–63.
23. Poret M, Chandrasekar B, van der Hoorn RAL, Avicé JC. Characterization of senescence-associated protease activities involved in the efficient protein remobilization during leaf senescence of winter oilseed rape. *Plant Sci.* 2016;246:139–53.

24. Poret M, Chandrasekar B, van der Hoorn RAL, Déchaumet S, Bouchereau A, Kim TH, Lee BR, Macquart F, Hara-Nishimura I, Avice JC. A Genotypic Comparison Reveals That the Improvement in Nitrogen Remobilization Efficiency in Oilseed Rape Leaves Is Related to Specific Patterns of Senescence-Associated Protease Activities and Phytohormones. *Front Plant Sci.* 2019;10:46.
25. Zou Z, Liu J, Yang L, Xie G. Survey of the rubber tree genome reveals a high number of cysteine protease-encoding genes homologous to Arabidopsis SAG12. *PLoS One.* 2017;12:e0171725.
26. Zou Z, Xie G, Yang L. Papain-like cysteine protease encoding genes in rubber (*Hevea brasiliensis*): comparative genomics, phylogenetic, and transcriptional profiling analysis. *Planta.* 2017;246:999–1018.
27. Funk V, Kositsup B, Zhao C, Beers EP. The Arabidopsis xylem peptidase XCP1 is a tracheary element vacuolar protein that may be a papain ortholog. *Plant Physiol.* 2002;128:84–94.
28. Felgueiras R, Magalhães R, Silva M, Silva M, Correia M. Transient ischemic attack: Incidence and early risk of stroke in northern Portugal from 1998–2000 to 2009–2011. *International journal of stroke: official journal of the International Stroke Society.* 2020;15:278–88.
29. van Wyk S, Du Plessis M, Cullis C, Kunert K, Vorster B. Cysteine protease and cystatin expression and activity during soybean nodule development and senescence. *BMC plant biology.* 2014;14:294.
30. Xu F, Chye M. Expression of cysteine proteinase during developmental events associated with programmed cell death in brinjal. *The Plant journal: for cell molecular biology.* 1999;17:321–7.
31. Koizumi M, Yamaguchi-Shinozaki K, Tsuji H, Shinozaki K. Structure and expression of two genes that encode distinct drought-inducible cysteine proteinases in Arabidopsis thaliana. *Gene.* 1993;129:175–82.
32. Zang Q, Wang C, Li X, Guo Z, Jing R, Zhao J, Chang X. Isolation and characterization of a gene encoding a polyethylene glycol-induced cysteine protease in common wheat. *Journal of biosciences.* 2010;35:379–88.
33. Chen H, Su C, Lin C, Huang G, Lin Y. Expression of sweet potato cysteine protease SPCP2 altered developmental characteristics and stress responses in transgenic Arabidopsis plants. *Journal of plant physiology.* 2010;167:838–47.
34. Xiao HJ, Yin YX, Chai WG, Gong ZH. Silencing of the CaCP gene delays salt- and osmotic-induced leaf senescence in *Capsicum annuum* L. *Int J Mol Sci.* 2014;15:8316–34.
35. Wang W, Zhou XM, Xiong HX, Mao WY, Zhao P, Sun MX. Papain-like and legumain-like proteases in rice: genome-wide identification, comprehensive gene feature characterization and expression analysis. *BMC Plant Biol.* 2018;18:87.
36. Clark K, Franco JY, Schwizer S, Pang Z, Hawara E, Liebrand TWH, Pagliaccia D, Zeng L, Gurung FB, Wang P, Shi J, Wang Y, Ancona V, van der Hoorn RAL, Wang N, Coaker G. W. Ma, An effector from the Huanglongbing-associated pathogen targets citrus proteases. *Nat Commun.* 2018;9:1718.
37. Zou Z, Huang Q, Xie G, Yang L. Genome-wide comparative analysis of papain-like cysteine protease family genes in castor bean and physic nut. *Sci Rep.* 2018;8:331.
38. Liu J, Sharma A, Niewiara MJ, Singh R, Ming R, Yu Q. Papain-like cysteine proteases in *Carica papaya*: lineage-specific gene duplication and expansion. *BMC Genom.* 2018;19:26.
39. Liu H, Hu M, Wang Q, Cheng L, Zhang Z. Role of Papain-Like Cysteine Proteases in Plant Development. *Front Plant Sci.* 2018;9:1717.

40. Coulombe R, Grochulski P, Sivaraman J, Ménard R, Mort JS, Cygler M. Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *Embo j.* 1996;15:5492–503.
41. Kramer L, Turk D, Turk B. The Future of Cysteine Cathepsins in Disease Management. *Trends Pharmacol Sci.* 2017;38:873–98.
42. Bateman A, Bennett HP. The granulin gene family: from cancer to dementia. *Bioessays.* 2009;31:1245–54.
43. Zang QW, Wang CX, Li XY, Guo ZA, Jing RL, Zhao J, Chang XP. Isolation and characterization of a gene encoding a polyethylene glycol-induced cysteine protease in common wheat. *J Biosci.* 2010;35:379–88.
44. Chen HJ, Su CT, Lin CH, Huang GJ, Lin YH. Expression of sweet potato cysteine protease SPCP2 altered developmental characteristics and stress responses in transgenic *Arabidopsis* plants. *J Plant Physiol.* 2010;167:838–47.
45. Li J, Yang Y, Sun K, Chen Y, Chen X, Li X. Exogenous Melatonin Enhances Cold, Salt and Drought Stress Tolerance by Improving Antioxidant Defense in Tea Plant (*Camellia sinensis* (L.) O. Kuntze), *Molecules*, 24 (2019).
46. Li XD, Zhuang KY, Liu ZM, Yang DY, Ma NN, Meng QW. Overexpression of a novel NAC-type tomato transcription factor, SINAM1, enhances the chilling stress tolerance of transgenic tobacco. *J Plant Physiol.* 2016;204:54–65.
47. Wang N, Zhang W, Qin M, Li S, Qiao M, Liu Z, Xiang F. Drought Tolerance Conferred in Soybean (*Glycine max.* L) by GmMYB84, a Novel R2R3-MYB Transcription Factor. *Plant Cell Physiol.* 2017;58:1764–76.
48. Li H, Ye K, Shi Y, Cheng J, Zhang X, Yang S. BZR1 Positively Regulates Freezing Tolerance via CBF-Dependent and CBF-Independent Pathways in *Arabidopsis*. *Mol Plant.* 2017;10:545–59.
49. Mao C, Ding J, Zhang B, Xi D, Ming F. OsNAC2 positively affects salt-induced cell death and binds to the OsAP37 and OsCOX11 promoters. *Plant J.* 2018;94:454–68.
50. Borgohain P, Saha B, Agrahari R, Chowardhara B, Sahoo S, van der Vyver C, Panda SK. SINAC2 overexpression in *Arabidopsis* results in enhanced abiotic stress tolerance with alteration in glutathione metabolism. *Protoplasma.* 2019;256:1065–77.
51. Fracheboud Y, Jompuk C, Ribaut JM, Stamp P, Leipner J. Genetic analysis of cold-tolerance of photosynthesis in maize. *Plant Mol Biol.* 2004;56:241–53.

Figures

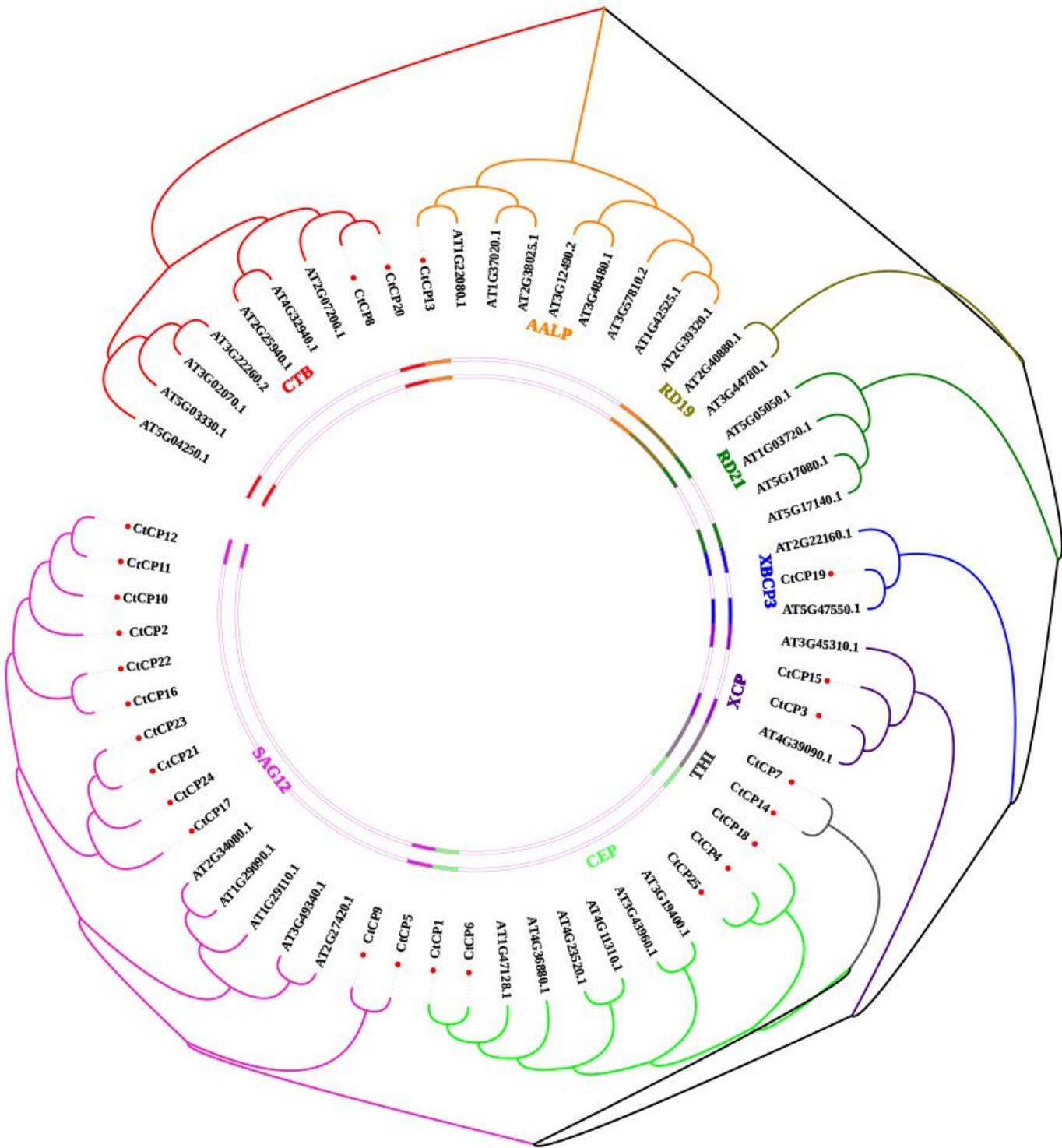


Figure 1

phylogenetic tree analysis of safflower and Arabidopsis CPs. 25 safflower and 36 Arabidopsis CPs, constructed evolutionary trees using MEGA-X. Different colors represent different subfamilies. Safflower sequences are shown with red dots.

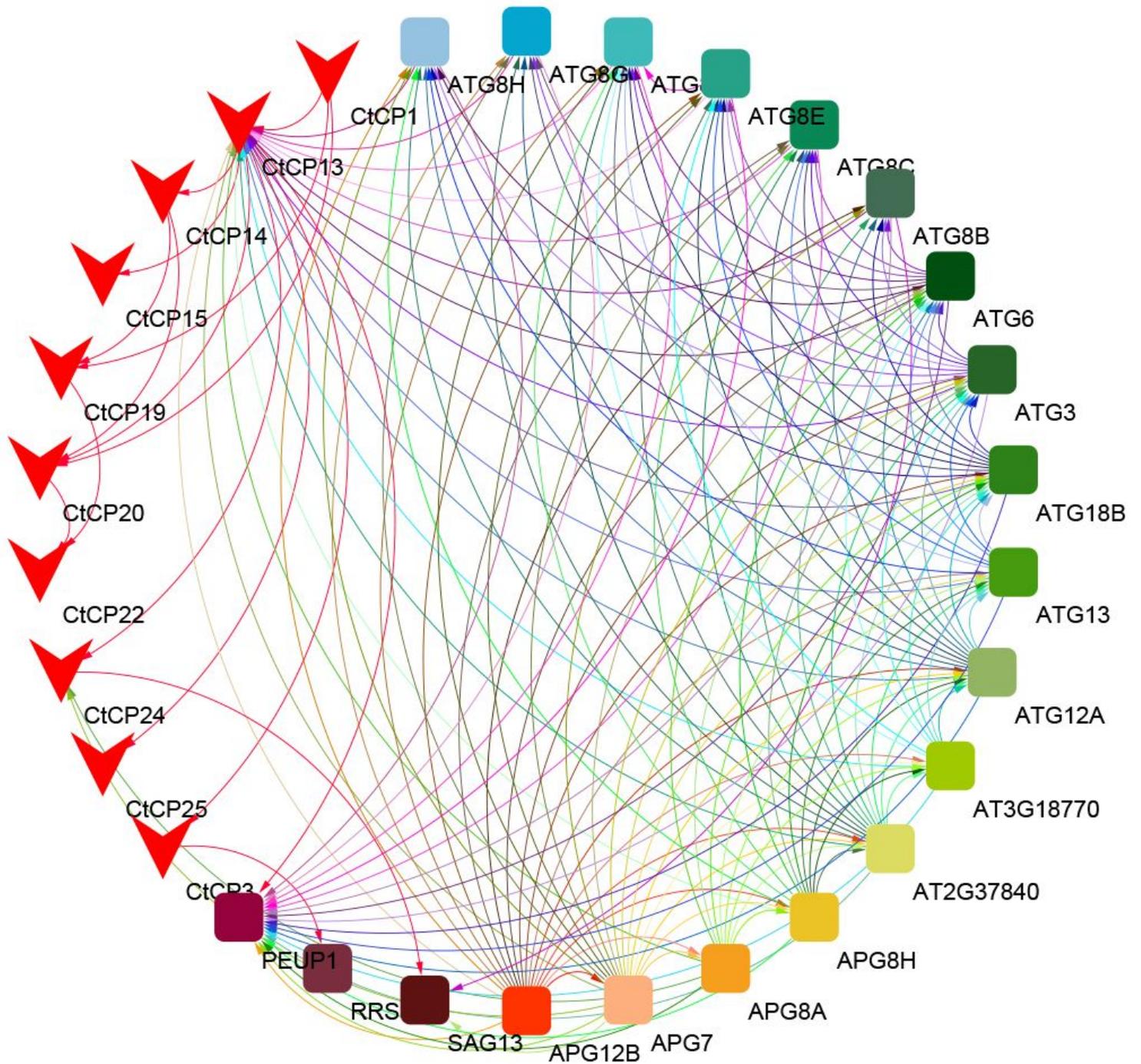


Figure 3

Based on the available information about CtCP homologous genes in *Arabidopsis thaliana*, the protein association network of CtCP. Online tool strings are used to predict the entire network. Lines of different colors represent interactions with different proteins. CtCP proteins are represented by red V, while other proteins are represented by rectangles of different colors.

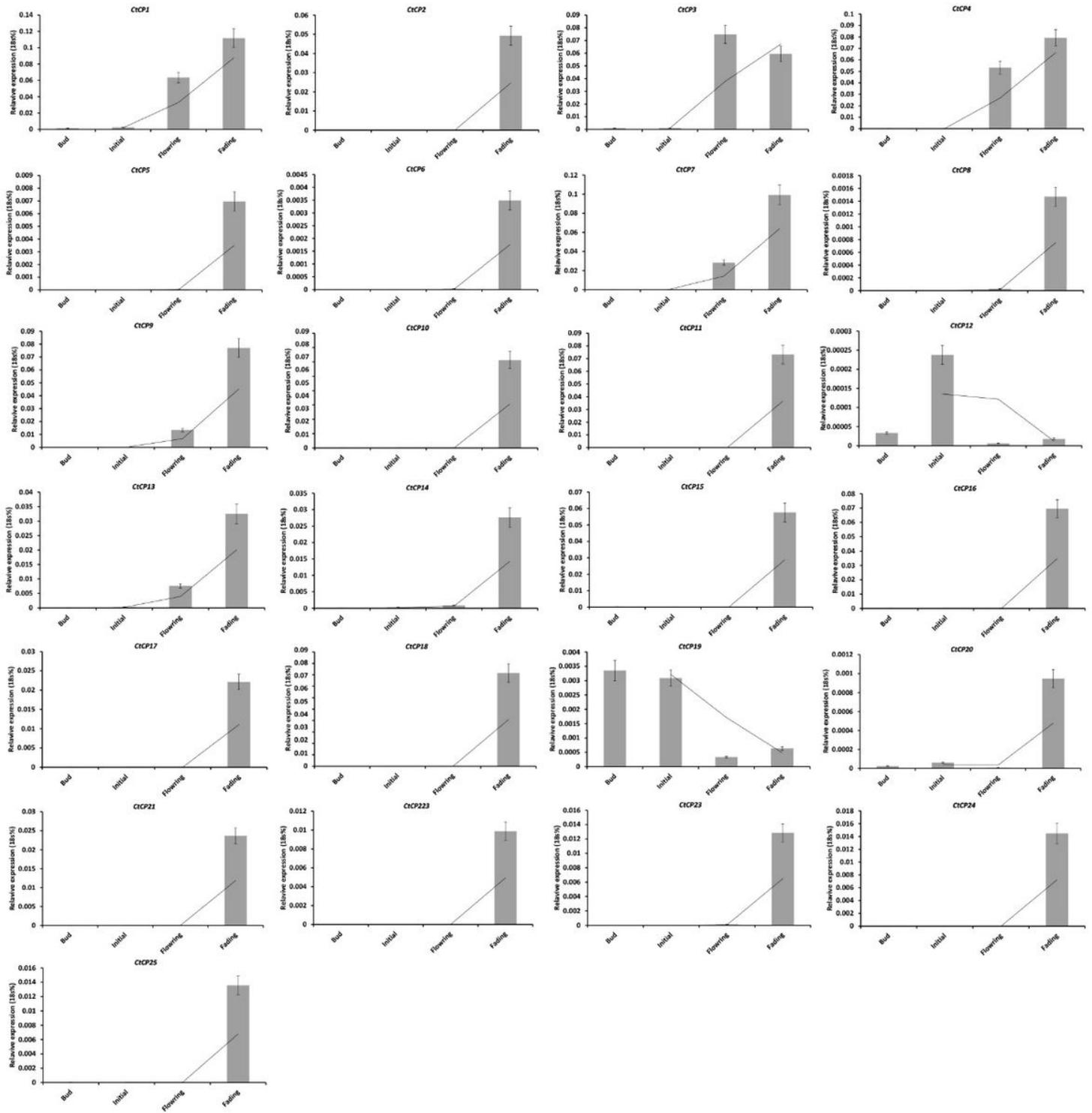
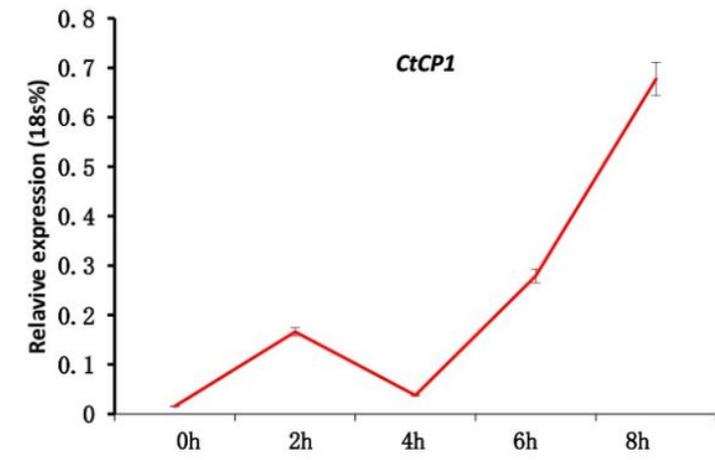
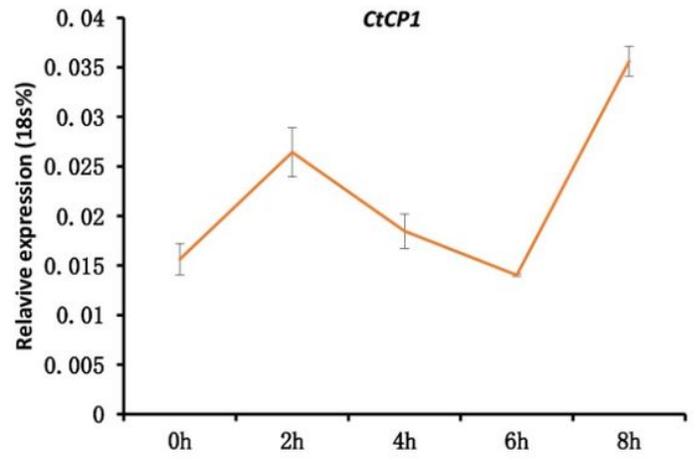


Figure 4

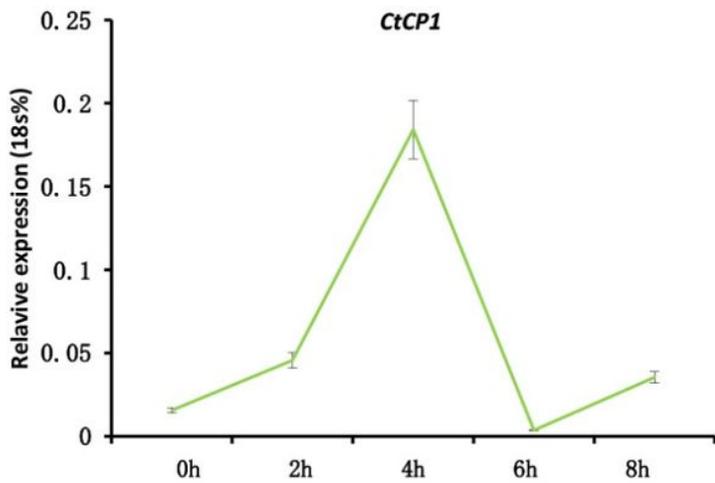
qRT-PCR analysis of CtCP gene in different flowering stages of *Carthamus tinctorius*, the numerical value represents the average \pm standard deviation of three biological repetitive samples.



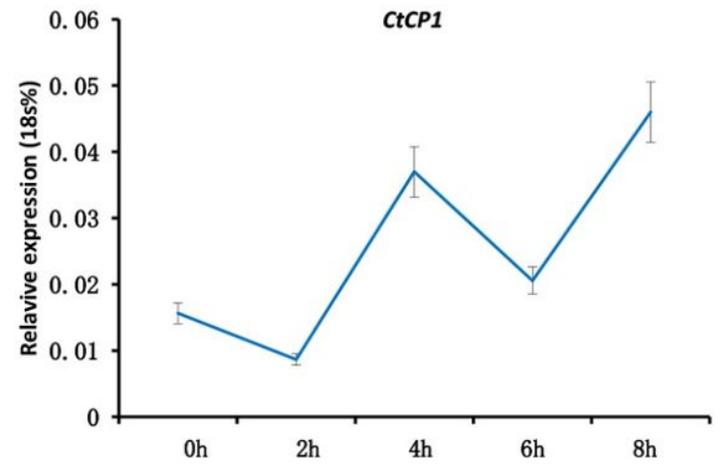
Cold Treatment Time



NaCl Treatment Time



H2O2 Treatment Time



PEG Treatment Time

Figure 5

Expression profile of CtCP genes in response to different abiotic stresses. Color scale denotes FPKM normalized by Genesis software.

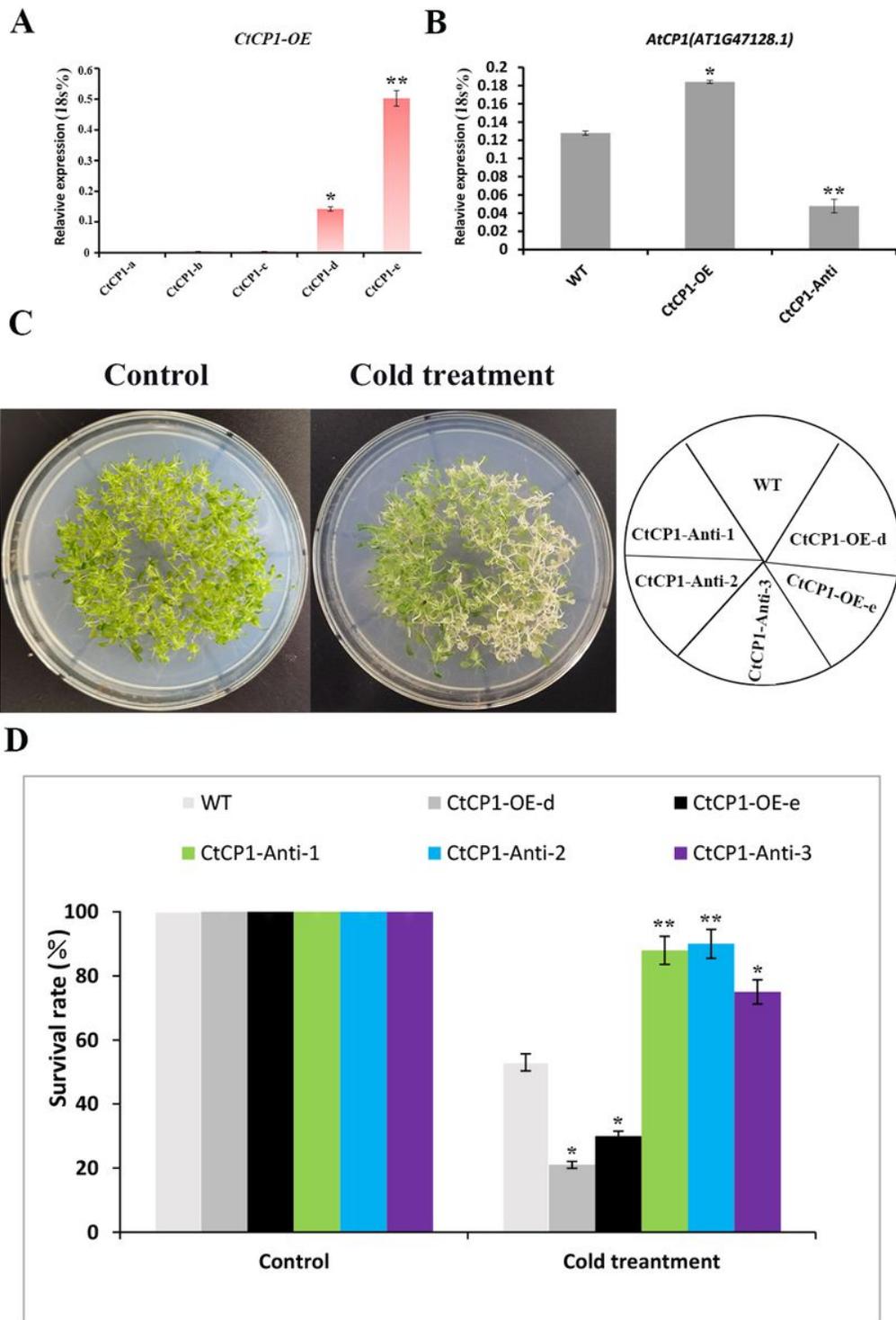


Figure 6

A) Analysis of CtCP1-OE expression in overexpression lines, a, b, c, d and e represent different lines. (B) The expression level of AtCP1 (At1g47128.1) in different lines, CtCP1-OE: CtCP1 overexpression; CtCP1-Anti: CtCP1 antisense suppression; WT: wild type. (C) Control (WT), the appearance of Arabidopsis seedlings with CtCP1-Anti inhibited expression and CtCP1-OE overexpression, with or without (control) cold stress treatment. The plants were grown on 1/2MS plates at 22°C for 12 days, and then cold treated at -8°C for 10 minutes. (D) The survival rate of Arabidopsis seedlings shown in (C). Each experiment is in triplicate. Asterisks indicate

that the value is significantly different from that of the wild-type group under same condition (* $P < 0.05$, ** $P < 0.01$, Student's t-test)

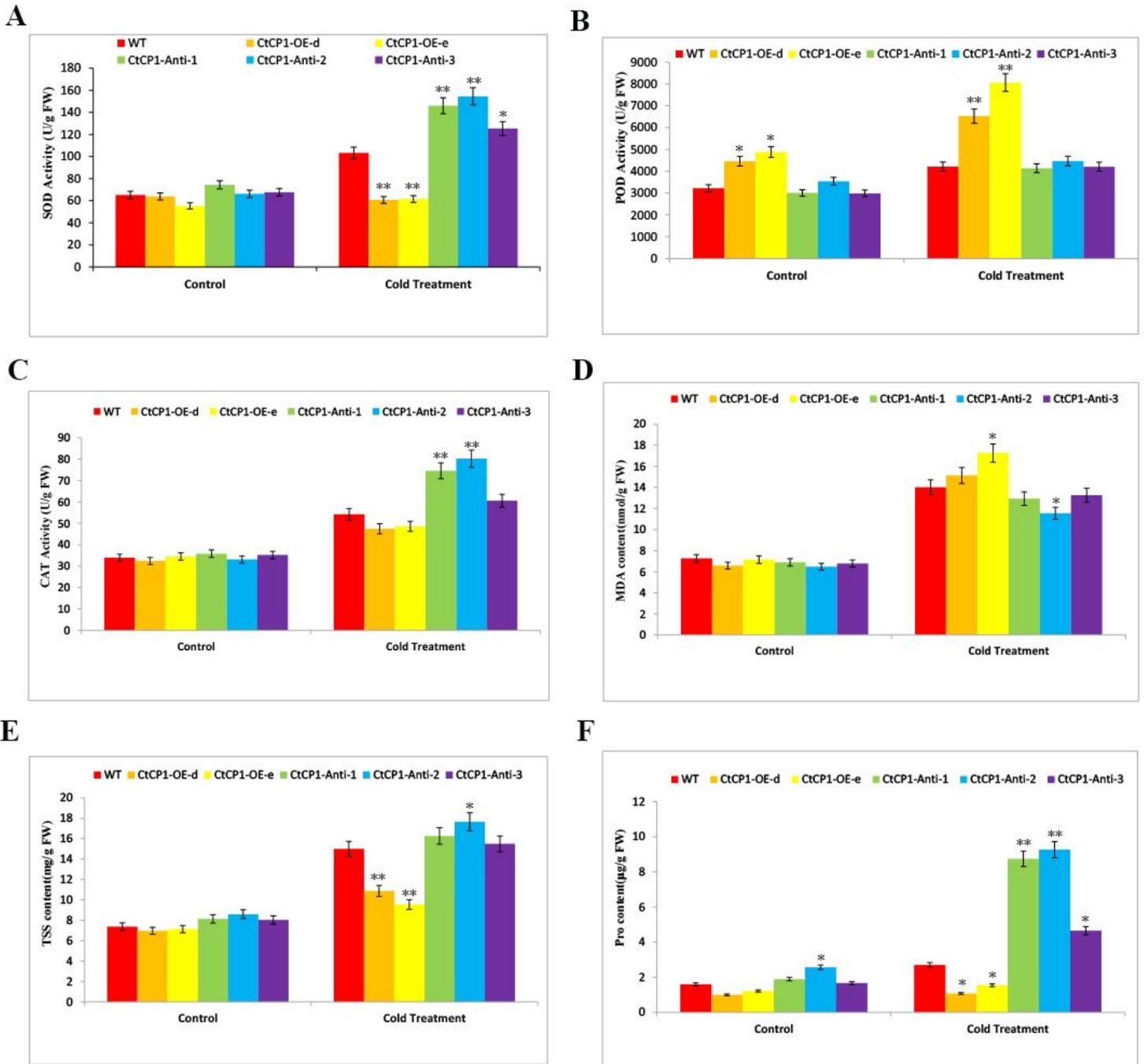


Figure 7

Determination of antioxidant enzyme activity, malondialdehyde, proline and soluble sugar (A) SOD enzyme activity; (B) POD enzyme activity; (C) CAT enzyme activity; (D) MDA content; (E) TSS content; (F) Pro content. Each experiment is in triplicate. Error bars indicate standard deviation. Asterisks indicate that the value is significantly different from that of the wild-type group under same condition (* $P < 0.05$, ** $P < 0.01$, Student's t-test)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.docx](#)
- [TableS2.docx](#)
- [TableS3.docx](#)