Genome-Wide Identification and Analysis of the Cryptochrome/Photolyase Family in brown algae Saccharina japonica

Yukun Wu
Chinese Academy of Fishery Sciences

Pengyan Zhang
Chinese Academy of Fishery Sciences

Zhourui Liang
Chinese Academy of Fishery Sciences

Yanmin Yuan
Chinese Academy of Fishery Sciences

Yi Liu
Ocean University of China

Di Zhang
Chinese Academy of Fishery Sciences

Fuli Liu (liufuli@ouc.edu.cn)
Ocean University of China

Research Article

Keywords: Cryptochrome/Photolyase Family (CPF), blue photoreceptor, genome-wide analysis, Saccharina japonica, gametophyte development

Posted Date: January 4th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2424636/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Cryptochrome/Photolyase Family (CPF) is a widely distributed and highly conserved gene family. The CPF includes photolyases that repair UV-induced DNA damage and blue-light receptors cryptochromes that are known for their photoreceptive functions in terrestrial plants. This study mainly emphasized the CPF members of the economically important brown alga *Saccharina japonica*, and traces the evolutionary relationship and distribution of this family in algae. Eight new CPF members were identified from the genomes of *S. japonica* and *Undaria pinnatifida*, and sixty CPF members from other different species were collected simultaneously for the phylogenetic analysis. Phylogenetic and transcriptional analyses provide evidence for the evolution and function of CPF genes in brown algae. According to the phylogenetic tree, the CPF family was divided into five branches. Photolyase homology domain (PHR) and a non-covalently bound flavin adenine dinucleotide (FAD) molecule are present in the CPF members of two brown algae *S. japonica* and *U. pinnatifida*. But they lack the typical C-terminal extension (CCT) characteristic of plant-like cryptochromes. The brown algae *Ectocarpus siliculosus* and *U. pinnatifida* have both animal-like CPF and plant-like CPF while only plant CPF members were identified in *S. japonica*. The results of RNA-seq and qPCR indicated that *Sj* Cryptochrome-DASH1 (CRY-DASH1) and *Sj* cyclobutane pyrimidine dimer (CPD) genes may not only play a role in light response but also play an important role in the gametophyte development of *S. japonica*. Protein interaction network of *Sj* CPF members showed that there are circadian clock protein Timeless homologous genes in *S. japonica*, which are located in the core position of the protein interaction network. This implied that *Sj* CPF members may transmit information by interacting with downstream Timeless genes. These results provide information for subsequent research on functional and regulatory pathways of CPFs in algae species.

Introduction

Light is one of the important environmental factors affecting physiological state and behavior of organisms. Organisms have evolved a series of light signal pathways to adapt to various subtle light changes in the environment. Various light-sensitive proteins that can sense and transduce light signal, are called photoreceptors (Kianianmomeni, 2014). Cryptochrome/Photolyase Family (CPF), as a class of Blue/ultraviolet-A light photoreceptors, are widespread in all major evolutionary lineages of life from marine algae to terrestrial plants, from primitive multicellular animals to humans. Three major categories of proteins are included in this family: the cyclobutane pyrimidine dimer (CPD) photolyases, the (6–4) pyrimidine-pyrimidone adduct [(6–4) photoproduct] photolyases, and the cryptochromes (CRY) (Xue et al., 2019). CPD photolyases repair cyclobutane pyrimidine (Pyr < > Pyr) dimers (CPDs), while (6 – 4) photolyases repair pyrimidine–pyrimidone (6–4) photoproducts (Pyr[6–4]Pyr) photoproducts (Kavakli et al., 2017). In contrast, cryptochromes, function as photoreceptors and signaling molecules, are among the first photoreceptors evolved in the organism kingdom (Xue et al., 2019). During evolution, some DNA photolyase lost the function of repairing DNA damage, but gradually acquired the function of blue light response and evolved into cryptochrome (Chaves et al., 2011). Currently, the accepted definition of a cryptochrome is a protein with similarity to photolyases that has lost or reduced DNA repair activity and
has gained a novel role in light signaling (Chaves et al., 2011). Cryptochromes normally regulate growth and circadian rhythms in terrestrial plants, while in mammals and insects cryptochromes act as circadian transcriptional repressors, or integral parts of the circadian oscillator (Horst et al., 1999; Anthony et al., 2003; Yuan et al., 2007).

Currently, only a small number of CRY family members in algae have been studied by spectroscopy and photochemistry (König et al., 2017; Coesel et al., 2009; Kiontke et al., 2020; Juhas et al., 2014), and these results indicate that some CPF in algae have both DNA repair activity and signal transduction function activated by blue light. The cryptochrome and DNA photolyases is therefore usually named and studied as a large Cryptochrome/Photolyase Family (CPF) in marine algae research (Fortunato et al., 2015).

CPFs are generally defined by two domains in common: the highly conserved FAD-binding photolyase homology region (PHR) domain and the CRY C-terminal extension (CCT) domain with various lengths. Two chromophore cofactors (Flavin Adenine Dinucleotide, FAD; Pterin or an 8-hydroxy-5-deazaflavin) can bind to the PHR domain (Chaves et al., 2011). The chromophore is the main site of photon absorption and gives CPFs the specific photochemical and photophysical properties (Chen et al., 2013).

Current research on the CPF has focused on terrestrial plants, particularly *Arabidopsis*. Photoactivation of plant CPF begins with photoexcitation or photon absorption and photoreduction of the FAD chromophore (Malhotra et al., 1995), resulting in a conformational change and formation of the cryptochrome homooligomers that interact with other proteins to alter gene expression (Sang et al., 2005; Wang et al., 2016). The biological clock is an endogenous timing mechanism that synchronizes biological processes with the length of the environmental circadian cycle (Harmer et al., 2009). CRY plays an important role in the circadian clock through this mode of protein interaction. It coordinates internal physiological and external environmental signals to provide adaptive advantages by measuring time through a molecular feedback loop with a period of approximately 24 hours (Huang et al., 2016).

Although CPF members are widely present in the biological world, the information on CPF currently available in aquatic organisms is very limited. Unique light properties found in aquatic environments compared to terrestrial environments likely shape the functional diversity of CPF in aquatic organisms (Huang et al., 2016). Marine organisms face more challenges of light than terrestrial species, because of the refraction and absorption of light by micro-particles in seawater (Castro et al., 2004). Marine organisms, especially algae, may experience significant changes in underwater light fields at different depths, in mixed water, in different habitats, or in the presence of other organisms (Marianne et al., 2017). In fact, underwater light is highly absorbed at wavelengths below 250 nm and above 700 nm, resulting in a progressive enrichment of blue-green (400–500 nm) light components with depth. This is possibly one of the main environmental influences that lead to differences in the regulatory mechanisms of marine and terrestrial photoreceptors.

Photosynthetic Stramenopile belongs to an additional SAR (Stramenopiles, Alveolates, Rhizaria) supergroup originating from the secondary endosymbiosis of unicellular red algae in eukaryotic host cells (Keeling et al., 2013). Ochrophyta is a part of Stramenopiles in SAR, including mainly unicellular and
mostly photosynthetic lineages such as diatoms, phaeophyceae, chrysophytes, synurophytes, xanthophytes and many less well-known groups (Charrier et al., 2008). Within this lineage the Phaeophyceae is a special taxon and one of the few eukaryotic lineages to have evolved complex multicellularity (Mark et al., 2014). The current availability of several Phaeophyceae genomes provides a good opportunity for the study of CPF in Ochrophyta (Mark et al., 2010; Ye et al., 2015, Graf et al., 2021). This part has often been lacked in previous studies of marine CPFs, especially for *S. japonica* and *U. pinnatifida* with a heteromorphic life cycle: microscopic gametophyte generation alternating with large sporophyte generation (Bartsch et al., 2008).

In order to better understand the complex evolution and function of CPF in algae, the CPFs in several algae were identified and characterized based on the genome sequence data in this study, focusing on the CPFs form *S. japonica*. Benefiting from updated information on genome data from other species, we extended the study of CPF phylogeny to the representative species of related evolutionary branches, aiming to uncovering the complexity and characterization of CPF evolution and function in different lineages. The obtained information about CPFs in this study will open a new gateway for subsequent CPF studies in future.

**Materials And Methods**

**Data Collection**

A total of 68 CPFs were identified or collected, and then were used in phylogenetic analysis. Sixty of these CPF members are from a family analysis (Paola et al., 2014) on marine CPF lineages. Protein sequences were obtained by deriving gene numbers from NCBI Batch Entrez (https://www.ncbi.nlm.nih.gov/sites/batchentrez) retrieved from this report. The other eight CPFs were identified by this study from two brown algae *S. japonica* (https://bioinformatics.psb.ugent.be/gdb/Saccharina/) (Ye et al., 2015) and *U. pinnatifida* (http://www.magic.re.kr/portal/assembly/MA00358) (Graf et al., 2021). Based on the annotation information of *S. japonica* and *U. pinnatifida* genomes, the local blast in Tbtools (https://github.com/CJ-Chen/TBtools/releases) (Chen et al., 2020) was applied to identify *Sj/Up* CPF with the *Es* CPF members as query sequences.

**Phylogenetic and Structural Analysis of CPF**

Analysis of gene families was mainly carried out with reference to an earlier paper (Wu et al., 2022). All network addresses needed for online analysis are summarized in Supplement Table. 3. Compared with the earlier analysis, we have added an analysis of the protein interaction network here. Prediction of protein interaction networks was accomplished by importing *Sj* CPF protein sequences into STRING (Search Tool for the Retrieval of Interaction Gene/Proteins, https://string-db.org/).

**Materials Culture of *S. japonica*and Expression Analysis of *Sj* CPFs**
The *S. japonica* female gametophyte comes from the gametophyte clones preserved in our laboratory under low temperature and low light conditions (i.e., 4°C, below 5 µmol photons m\(^{-2}\)s\(^{-1}\)). It was pretreated at 10°C for 24 hours in darkness before the start of the experiment. And then, the gametophytes were transferred to be culture under different light quality (white, blue and red) conditions at 10°C, and the transcription levels of genes at 1, 3, 6, and 12 h were detected by RT-PCR method. The female gametophytes were collected at those times and snap-frozen in liquid nitrogen and stored at −80°C.

The differential expression changes of *Sj* CPFs in different developmental stages of *S. japonica* gametophyte were derived from transcriptome data (PRJNA816677) reported by Wu et al (2022). Transcriptome data are divided into four distinct stages according to different time points during gametophyte development: Dormant period (0 d), Vegetative growth period (3 d), Vegetative growth peak (6 d), and Gametogenesis period (9 d).

**RNA extraction and RT-PCR**

Specific qRT-PCR primers (Supplement table 1) for the *Sj* CPF gene were designed by NCBI Primer-BLAST (https://www.ncbi.nlm.nih.gov/Tools/Primer-BLAST/, accessed on June 12, 2022). β-actin (GenBank accession number: FJ375360) was selected as a housekeeping gene because previous studies showed that it is stably expressed under light conditions (Deng et al., 2012). RNA was extracted using Polysaccharide Polyphenol/Complex Plant RNA Rapid Extraction Kit (SparkJade, Qingdao, China) and SP ARKscript II RT kit with gDNA Eraser (SparkJade, Qingdao, China) and cDNA was prepared using the SP ARKeasy. RT-PCR was performed using CHAMQ SYBR COLOR QPCR Master Mix (Vazyme, Nanjing, China) and Applied Biosystems StepOnePlus™ Real-Time PCR System (Applied Biosystems, Waltham, America). Expression levels were determined by averaging the values of three replicates for each experimental condition.

**Results**

**Identification and Characterization of CPF in algae Species**

Eight CPF members were identified from the brown algal genome by homologous alignment: four in *S. japonica* and four in *U. pinnatifida*. Additionally, four and five CPFs were collected from the brown algae *E. siliculosus* (Mark, 2010) and diatom *Phaeodactylum tricornutum* respectively (Martino et al., 2007). The gene numbers, genome locations, open reading frame lengths and exon numbers of CPF members of the four algae species was listed in Table 1. The ORF length of CPF genes in the four algae ranged from 492 bp (*U. pinnatifida* CPD2, Gene ID: ML_UPxx_s01_p14962) to 3528 bp (*S. japonica* CRY-DASH2, Gene ID: SJ12406) with an average of 1778 bp. CPF genes of marine algae are scattered on different chromosome or scaffolds without aggregation.

The information on protein levels of CPFs in four typical stramenopiles species were provided in Table 2. These seventeen CPF members encode proteins ranging from 164 amino acids (*U. pinnatifida* CPD2,
Gene ID: MI_UPxx_s01_p14962) to 1176 amino acids (S. japonica CRY-DASH2, Gene ID: SJ12406). The protein molecular weights ranged from 18.138 kDa to 122.044 kDa, with an average of 64.6 kDa, and most of the CPFs were between 50-70kDa.

None of the identified CPF members have a transmembrane structure as predicted by DeepTMHMM. The predicted results of subcellular localization indicate that Pt CPF and Es CPF are mainly located in chloroplast, while Up CPF is mainly distributed in nucleolus and Sj CPF1 may mainly be distributed in cytoplasm. The two main protein folding modes, α-helix and random coil, accounting for 33.77% - 46.38% and 36.95% - 47.99%, respectively. They form the main folding modes in the CPF protein with a proportion of 78.89% - 87.69%. The other two protein folding modes, β-turn and extended strand, accounted for 4.11%-8.77% and 6.03%-13.54% respectively. The number of phosphorylation sites in CPFs from the four species ranged from 9 to 128. Serine is the most numerous of the three phosphorylation sites (Serine, Threonine and Tyrosine).

**Characterization of CPF Gene and Protein Structure**

The exon number of CPF genes varied markedly ranging from 1 to 13 in the four algae (Fig. 1, Table 1). The exon numbers of CPF genes in the three brown algae were higher than that in the diatom P. tricornutum. The highest number (13 exons) of exon presented in S. japonica CRY-DASH1, while the lowest number of exon existed in U. pinnatifida CPD2 (5 exons). Compared to the three brown algae, the diatom P. tricornutum possesses less exons in CPF genes. Except that Pt CPF1 and Pt CPF2 have five and four exons, respectively, the other three CPF genes have only one exon (Table 1, Fig. 1). In order to further understand the evolution-function-structure relationship of CPF, the conserved domain information (from Protein Families database, Pfam, Fig. 2) and conserved motif sites (from Multiple Expectation Maximization for Motif Elicitation, MEME, Fig. 3) of CPF genes of several algae species were analyzed. The CPF members of different species collected and identified mainly contain two conserved domains: DNA photolyase domain and FAD binding domain. Pt CPD4 and Up CRY-DASH lack the critical structural domain DNA photolyase. Typical characteristic structures of C-terminal abhydrolase and cryptochrome C-terminal characteristic domain (This domain is normally missing in stramenopiles) were identified in Sj CRY-DASH and At CRY1, respectively. Most of the DNA photolyase domains in these identified CPFs are located closer to the N-terminus while the FAD binding domain is usually closer to the C-terminus.

CPFs have intrinsically disordered regions (IDRs) with varying lengths and high sequence diversity. The major IDR of CPFs appear in the middle part of the two conserved domains. Ten most conserved motifs were predicted from the protein sequences of sixteen typical algal CPFs (Fig. 3), among which motif 1 and motif 5 are widely distributed in the CPFs.

**Phylogenetic Analysis**

In order to understand the evolution and possible function of CPFs in stramenopiles species, CPFs from 14 typical organisms were collected and subject to phylogenetic analysis (Fig. 4). The 14 organisms were selected from diverse lineages, including 6 model organisms and eight species of aquatic algae from...
opisthokont, archaeplastida, stramenopiles and cryptophyta (Fig. 4). According to the phylogenetic tree and previous study (Paola, 2014) on the classification of CPF family, the collected 68 CPFs were divided into seven groups: Algae CRY-DASH-like, Plants-like CPF/CRY and CRY-DASH, Plants CRY, CPD Photolyases, ((6-4) Photolyases, Animals-like CPF/CRY, Animalia CRY. E. siliculosus, S. japonica and U. pinnatifida are three closely related species that belong to Ochrophyta, but the CPF members of the three species are not all clustered in the same evolutionary clade. Sj CRY-DASH1, Sj CRY-DASH2 and Up CRY-DASH were clustered in the algal CRY-DASH group, and no Es CPF members were found in the algae CRY-DASH group. Sj CPF, Es CRY2 and Es CRY3 were clustered into the algal CPF group, and no Up CPF members were found in this group. Most members of this group are derived from algae, but it also contains two CRY-DASH from higher plants. CPD is the only evolutionary group found in all three species of brown algae. Es CRY1 and Up CRY were clustered into (6-4) photolyase group, which is a special type of CPD, and S. japonica lacks genes in this group.

The (6-4) photolyase group of algae and fish species in a subset of aquatic organisms is very closely related to CRY of animal species. They are assigned to the same large clade. Two genes Cr CPH1 and Vc CPH1 in two chlorophyta (green algae) species Chlamydomonas reinhardtii and volvox carterii were assigned to the lineage of plant CRY in the phylogenetic tree. Notably, CRY-DASH in Arabidopsis thaliana and Oryza sativa are the only two non-algal CPF members clustered into algae CPF.

**Upstream Cis-Acting Elements Analysis and Protein Interaction Network of CPF**

Cis-acting elements were predicted by intercepting 3200 nucleotide sequences upstream of CPF genes (Fig. 5). The representative 6 types of cis-acting elements are low-temperature, abscisic acid, circadian control, light, MeJA responsiveness and MYBhv1 binding site. These elements are regularly arranged in the promoter region. The exclusive binding sites for MYB were found in the upstream promoters of Es CRY1, Es CRY2, Sj CRY-DASH1, Sj CRY-DASH2. There are two clock-related sites embedded in Es CPD and Sj CPD upstream, both of which belong to the cyclobutene pyrimidine dimer (CPD).

The protein interaction network of CPF members were analyzed by importing the identified Es-Sj CPF protein sequence into STRING (Search Tool for the Retrieval of Interaction Gene/Proteins) (Fig. 6). Since no protein interaction data related to S. japonica in the STRING database, the protein interaction database of E. siliculosus, a relative species of S. japonica, was used to predict the possible interaction network of CPF members in S. japonica. The prediction results indicated that Es CRY2 and Es CRY3 may directly interact with four (ACV, Casein kinase, D8LBE5, Timeless, D7G629, EF-hand domain pair, D8LPS0, Hypothetical leucine rich repeat calmodulin binding protein) and five (ACV, Casein kinase, D8LBE5, Timeless, D7G629, EF-hand domain pair, CK, Casein kinase, D8LPS0, Hypothetical leucine rich repeat calmodulin binding protein) proteins (Fig. 6, Supplement table 2), respectively. There are 22 proteins that form direct and secondary interactions with Es CPF members, and 15 of them form direct interactions with Es Timeless.

**Expression Analysis of SjCPFs at Different Developmental Stages of S. japonica Gametophyte**
To explore the possible function of \( Sj \) CPF genes in \( S. japonica \), expression variation of \( Sj \) CPF genes was analyzed based on transcriptome data from different development stage of \( S. japonica \) gametophyte (Fig. 7). The expression levels of the three genes \( Sj \) CPF, \( Sj \) CRY-DASH1 and \( Sj \) CPD significantly changed in different developmental stages of the gametophyte. The expression of \( Sj \) CRY-DASH2 was stably and maintained at a very low level. Compared with the initial state, the \( Sj \) CPF gene showed a significant decrease (fold change > 2 and p-value < 0.05) when gametophyte entered the latter three stages (3 Day, 6 Day, 9 Day). \( Sj \) CRY-DASH1 and \( Sj \) CPD genes significantly increased in the stage from day 0 to day 3. Then the CPD gene gradually rose to the peak on the sixth day, but returned to the expression level of the dormant stage (0 d) during the gametophyte development stage.

**Light-Dependent Expression of CPFs in \( S. japonica \)**

In order to explore the expression of CPF genes in response to light period and quality in gametophyte of \( S. japonica \), the changes of transcription levels of \( Sj \)CPF genes at five different time points from dark to different light quality were detected (Fig. 8). The qPCR results of the two genes \( Sj \) CPF1 and \( Sj \) CRY-DASH2 showed that when the gametophyte moved from darkness to any light quality, the expression of them dropped sharply, and the expression level remained at a very low level for the next 12 h. Different from the expression patterns of the previous two genes, \( Sj \) CPD and \( Sj \) CRY-DASH1 showed obvious differences in expression patterns under different light qualities. The expression levels of \( Sj \) CRY-DASH1 and \( Sj \) CPD genes did not fluctuate significantly under continuous red light for 12 h. However, their expression patterns changed significantly under continuous exposure to white light. The expression of \( Sj \) CRY-DASH1 and \( Sj \) CPD gradually increased from 0 to 3 h and then decreased significantly from 3 to 6 h. Although blue light was included in white light, the expression patterns of \( Sj \) CRY-DASH and \( Sj \) CPD genes under white light appeared different from those under pure blue light. In continuous 12 h of blue light exposure, the \( Sj \) CRY-DASH1 gene first increased significantly (fold change > 3 and p-value < 0.05), and then at the sixth hour it decreased to a level similar to that at 0 h, and finally increased significantly at the 12 h (fold change > 2 and p-value < 0.05). However, the \( Sj \) CPD gene, which has a similar expression pattern to \( Sj \) CRY-DASH1 under white light, maintained a low expression level under pure blue light until the 12 h, and then increased significantly (compared to the initial stage fold change > 2 and p-value < 0.05).

**Discussion**

CPF is a strongly conserved and widespread gene family in the biological world. Its main member CRY is one of the earliest evolved photoreceptors (Han et al., 2019). In the studies on photosynthetic plants, cryptochromes have been proved to play a crucial role in the growth and development of plants mainly through photo-responsive protein–protein interaction as the signal transduction mode (Wang et al., 2017). At present, the understanding of cryptochromes in brown macroalgae is still vague. Although the genomes of several brown algae have been available (Ye et al., 2015, Graf et al., 2021), there is still no follow-up research related to CPF. In order to understand the evolution and function of CPFs in brown algae, detailed gene family analysis in representative brown alga were carried out. Moreover, expression
of CPF genes in response to different light quality and during different development stages were determined in *S. japonica* gametophyte, aiming to explore the possible functions of CPFs in brown algae.

### Distinct CPF member Composition and Conserved Gene Domains in Brown Algae

Three major categories of proteins are included in CPF family: the cyclobutene pyrimidine dimer (CPD) photolyases, the (6–4) pyrimidine-pyrimidone adduct [(6–4) photoproduct] photolyases, and the crytochromes (CRY) (Paola et al., 2014). The number of CPF genes varies little among different lineages, generally not more than ten, i. g. 5 in *A. thaliana*, 10 in *Danio rerio*, 6 in *P. tricornutum* (Baranasic et al., 2022, The Arabidopsis Genome Initiative, 2000, Martino et al., 2007). The number of CPF in three brown algae in this study is 4, which may indicate that the number of CPF in brown algae is conservative.

The domain characteristics of these CPF members are very conserved, often containing a photolyase homology region (PHR) that can bind to a chromophore (Essen et al., 2006). Algal CPF members are often missing a functional C-terminal domain that is critical in terrestrial plant CRY (Paola et al., 2014). *Es* CRY1 and *Up* CRY are phylogenetically closer to animal crytochrome and they are clustered into clades of class (6-4) photelyase-CRY. This result is similar to earlier studies in diatoms (Coesel et al., 2009). But other CPF members *Sj* CPF/CRY-DASH1 and *Sj* CRY-DASH2 are closer to the CRY genes in higher plants. These indicated that there is an algal-specific differentiation of CPF members. In previous study, it was found that the number of CPF homologous to animals and plants was almost equal in *T. pseudonana* (Coesel et al., 2009). The most straightforward interpretation is that these “animal-like” genes in brown algae were derived from the heterotrophic secondary host, although scenarios involving gene loss in the plant/red algal lineage cannot be ruled out (Armbrust, 2004).

The partial sequence of SJ04101 gene which was named as *Sj* CPF in this study was cloned and named as *Sj* CRY-DASH through homologous sequence alignment in previous study (Li et al., 2011). Although, it has been verified that SJ04101 has a very close relationship with the plant CRY-DASH gene, subsequent studies showed that the CPF members of this clade in marine algae act as a novel ultraviolet/blue light sensor in marine environments (Coesel et al., 2009). Therefore, the SJ40101 gene was renamed as *Sj* CPF in our study.

### Protein Interaction Networks of *Sj* CPFs

In the STRING database, there is information on protein interaction of the model organism of brown algae *E. siliculosus*. Additional annotation of interacting proteins of the *Sj* CPF gene based on annotation information from the genomes of *S. japonica* and *E. siliculosus* (Mark, 2010, Ye, 2015). Interestingly, not all CPF members have a huge interaction network, which may be due to the special evolutionary status of *S. japonica* and *E. siliculosus* and the weakness of molecular research on these two algae. A particularly interesting result emerged in the *Sj* CPF protein interaction network: the D8LBE5 protein was located at the core of the interaction network. D8LBE5 is one of the first molecular cogs identified in the *Drosophila* biological clock (Myers et al., 1995, Koike et al., 1998, Sangoram et al., 1998, Takumi et al., 1999). It was
named *Timeless* for its function in mediating light entrainment and temperature compensation of the molecular clock and for its key role in the circadian rhythm of *Drosophila* (Cai et al., 2021).

It is worth noting that the *Timeless* gene is conserved in terrestrial animals including vertebrates, and plays a very important role in the biological clock system. It is very interesting that a gene that plays an important role in the circadian clock system of higher animals appears in the photoautotrophic kelp. The CPF gene is one of the extremely conserved gene families in both animal and plant. Based on the protein interaction network of CPF, it can be speculated that *Sj* CPF may have an interaction with *Timeless* (TIM) proteins similar to that in animals. TIM function is extensively regulated by posttranslational modifications (PTMs) in *Drosophila* (Takumi, 1999). Casein kinase 2 (CK2) promotes the entry of TIM into the nucleus by phosphorylation (Cai et al., 2020). *CRY* controls TIM proteasomal degradation in the cytoplasm primarily during the daytime (Lee et al., 1996). However, although the *Sj* Timeless gene is located in the core position of the predicted *Sj* CPF protein interaction network, there is no direct evidence that Timeless genes play a role in the biological clock in Stramenopiles species. This requires further evidence.

**Expression of *Sj* CPF in response to the light and during gametophyte development**

RT-qPCR results demonstrated that the light-dependent transcriptional response of different members of the *Sj* CPF family varied remarkably. *Sj* CPD and *Sj* CRY-DASH1 responded to light changes from dark to blue and white light. The expression trend of these two proteins was almost unaffected by red light, but showed significant fluctuation under blue or white light. Research on the developmental physiology of *S. japonica* gametophytes showed that blue light can effectively promote the sexual development and reproduction of gametophytes, while red light almost completely inhibits the sexual development and reproduction but can keep the vegetative growth of gametophytes (Lüning et al., 1981). Therefore, it is speculated that the two genes (*Sj* CPD and *Sj* CRY-DASH1) that can respond to blue light and white light may be related to the development of *S. japonica* gametophytes. However, this need more evidence, and this thus will be a future research direction for CPFs in *S. japonica*.

Another notable point is that two cryptochromes (CRY1 and CRY2) are encoded in *Arabidopsis*, and the light responses they mediated are mostly functionally redundant (Wang, 2020). The main difference between the two is that CRY2 degrades rapidly under blue light while CRY1 is more stable (Malathy et al., 2021). Our results indicated that the gene expression of *Sj* CPF1 and *Sj* CRY-DASH2 significantly decreased from dark to blue light. Moreover, when entering the red light environment, they also decreased to a certain extent, but the decrease is smaller than that in the blue light environment. Thus, it is speculated that CPF members in *S. japonica* probably function with a mechanism similar to that in higher plants.

Transcriptome data indicated that *Sj* CRY-DASH1 shares a similar expression pattern with *Sj* CPD, which can be clearly distinguished from the expression patterns of the other two CPF members in *S. japonica*. The transcription levels of CRY-DASH1 and CPD genes were at high levels from the third day until the sixth and ninth days and then decreased during the gametophyte development. The two time points of
the sixth day and the ninth day just correspond to the process from the end of the vegetative growth period into the gametogenesis. This suggests that CRY-DASH1 and CPD may play roles during the transition from vegetative growth into gametogenesis of *S. japonica* gametophyte.

**Conclusion**

CPF is an ancient DNA repair and light-responsive protein family with so far mostly unclear functions and regulatory pathways. Here we identified CPF members in *S. japonica* and *U. pinnatifida* by gene family analysis and determined the gene expression of *Sj* CPF members at the transcriptional level. The gene expression of *Sj* CPF members can respond to blue light and differential expressed during different developmental stages of *S. japonica* gametophyte, indicating *Sj* CPF members probably play roles in *S. japonica* gametophyte. Moreover, it was found that *Sj* CPF has a interaction with Timeless, implying CPFs probably function in the circadian regulation in brown algae.. However, the biophysical and functional properties of many CPF members in algae is very limited. This requires further studies on the function and mechanism of CPFs in brown algae, which can help us further understand the light-responsive regulatory mechanism of CPFs.

**Declarations**

**Supplementary Materials:** The following supporting information can be downloaded at: www

**Author Contributions:** F.-L.L., Z.-P.Y and W.-Y.K. conceived and designed the research. Z.-P.Y and W.-Y.K. performed the transcriptome and gene expression analysis and wrote the original draft. Z.-R.L., Y.L., P.-Y.Z. Y.-M.Y. and D.Z. helped with the sampling and field work. P.-Y.Z. helped to perform the experiment. F.-L.L. revised the manuscript. All authors contributed to revising the manuscript. F.-L.L. and W.-Y.K. contributed equally. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the agricultural variety program of Shandong Province 2021LZGC004, Teachers Scientific Research Start-up Funding Project of Ocean University of China (862201013155), National Key R&D Program of China (2018YFD0900305) and China Agriculture Research System (CARS-50).

**Institutional Review Board Statement:** Not applicable

**Informed Consent Statement:** Not applicable

**Data Availability Statement:** RNA-Seq data can be found with accession number PRJNA816677. The RNA-Seq data is publicly available on National Center for Biotechnology Information. The other data presented in this study are available in Supplementary Materials.

**Acknowledgments:** This work was funded by the agricultural variety program of Shandong Province 2021LZGC004, National Key RandD Program of China (2018YFD0900305) and China Agriculture
Research System (CARS-50). The authors are grateful to the anonymous reviewers for their constructive comments on this paper and thanks to the marine enterprises in Penglai, China for providing us with high-quality kelp gametophyte germplasm resources.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


http://doi.org/10.1111/j.1529-8817.2007.00384.x
https://doi.org/10.1101/2022.04.08.487609
https://www.nature.com/articles/s41588-022-01089-w
https://doi.org/10.1016/j.sbi.2006.01.004
https://www.science.org/doi/10.1126/science.270.5237.805
https://doi.org/10.1111/febs.16253


https://www.science.org/doi/10.1126/science.271.5256.1740


**Tables**

Tables 1 and 2 are available in the Supplementary Files section.

**Figures**
Figure 1

Exon-intron structure diagram of three brown algae. The thick black lines represent exons, and the thin lines represent introns. In order to better observe the exon structure of genes among each other, the distance of each intron is adjusted to be the same in the figure.
Schematic diagram of the conservative domains of representative algal CPF members. Visualization of four conserved domains screened from conserved domain data exported from the Batch CD-search tool.

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
<th>Motif Locations</th>
</tr>
</thead>
<tbody>
<tr>
<td>AaCRY1</td>
<td>5.37e-110</td>
<td></td>
</tr>
<tr>
<td>AaCRY2</td>
<td>1.60e-86</td>
<td></td>
</tr>
<tr>
<td>CrCPH1</td>
<td>6.33e-68</td>
<td></td>
</tr>
<tr>
<td>CrCPDtypeII</td>
<td>9.37e-157</td>
<td></td>
</tr>
<tr>
<td>CrCRY-DASH2</td>
<td>3.55e-78</td>
<td></td>
</tr>
<tr>
<td>CrCPF1</td>
<td>1.30e-92</td>
<td></td>
</tr>
<tr>
<td>CrCRY-DASH1</td>
<td>1.27e-129</td>
<td></td>
</tr>
<tr>
<td>OiCPF2</td>
<td>2.94e-129</td>
<td></td>
</tr>
<tr>
<td>OiCPF5</td>
<td>1.62e-146</td>
<td></td>
</tr>
<tr>
<td>OiCPF1</td>
<td>3.31e-93</td>
<td></td>
</tr>
<tr>
<td>OiCPF3</td>
<td>1.91e-60</td>
<td></td>
</tr>
<tr>
<td>OiCPF4</td>
<td>2.84e-154</td>
<td></td>
</tr>
<tr>
<td>PtCPF2</td>
<td>1.31e-125</td>
<td></td>
</tr>
<tr>
<td>PtCPD3</td>
<td>2.36e-159</td>
<td></td>
</tr>
<tr>
<td>PtCPF1</td>
<td>2.40e-102</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3**

Motif composition structure of CPF members in representative organisms. Visualized motifs on gene sequences with different colors and lengths.

**Figure 4**

Phylogenetic relationship of CPF proteins from several typical organisms. The phylogenetic tree was constructed based on the proteins sequence alignment of the CFP proteins. The 68 CPF proteins are divided into seven groups (Algae CRY-DASH-like, Plants-like CPF/CRY and CRY-DASH, Plants CRY, CPD
Photolyases, ((6-4) Photolyases, Animals-like CPF/CRY, Animalia CRY) based on the phylogenetic tree. Each group is marked with a different colour. CPF members in *S. japonica* and *U. pinnatifida* are in bold. The detailed information of proteins used in this phylogenetic analysis were provided in Supplementary Table 2.

**Figure 5**

Six typical cis-acting elements in the upstream promoter regions of two brown algae *E. siliculosus* and *S. japonica* CPF genes.
Figure 6

CPF member protein interaction network. The protein interaction network obtained by importing $Sj$ CPF members into the STRING database (Search Tool for the Retrieval of Interaction Gene/Proteins). An empty "glass sphere" means that there is no 3D structure of the protein in the database. The straight line (edge) between the circle nodes represents the interaction between the two proteins connected by the straight line.
Figure 7

Differential expression of four *SjCPF* genes in different developmental stages of *S. japonica* gametophyte.
Figure 8

Expression variation of *Sj* CPF genes under different light quality irradiation. The abscissa represents different light quality, and the ordinate represents the relative expression of genes. The expression levels of the four genes in the zeroth hour under different light qualities were normalized to 1.

Supplementary Files

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

- TableandSupplementTable.xlsx