



# Systems scale characterization of circadian rhythm pathway in *Camellia sinensis*



Gagandeep Singh<sup>1</sup>, Vikram Singh<sup>1</sup>, Vikram Singh\*

Centre for Computational Biology and Bioinformatics, School of Life Sciences, Central University of Himachal Pradesh, Dharamshala 176206, India

## ARTICLE INFO

### Article history:

Received 13 January 2021  
Received in revised form 14 December 2021  
Accepted 18 December 2021  
Available online 05 January 2022

### Keywords:

*Camellia sinensis*  
Circadian rhythm  
PPI network  
Hidden Markov model  
Expression analysis

## ABSTRACT

Tea (*Camellia sinensis*) is among the most valuable commercial crops being a non-alcoholic beverage having antioxidant properties. Like in other plants, circadian oscillator in tea modulates several biological processes according to earth's revolution dependent variations in environmental cues like light and temperature. In the present study, we report genome wide identification and characterization of circadian oscillator (CO) proteins in tea. We first mined the genes (24, in total) involved in circadian rhythm pathway in the 56 plant species having available genomic information and then built their hidden Markov models (HMMs). Using these HMMs, 24 proteins were identified in tea and were further assessed for their functional annotation. Expression analysis of all these 24 CO proteins was then performed in 3 abiotic (A) and 3 biotic conditions (B) stress conditions and co-expressed as well as differentially expressed genes in the selected 6 stress conditions were elaborated. A methodology to identify the differentially expressed genes in specific types of stresses (A or B) is proposed and novel markers among CO proteins are presented. By mapping the identified CO proteins against the recently reported genome wide interologous protein–protein interaction network of tea (TeaGPIN), an interaction sub-network of tea CO proteins (TeaCO-PIN) is developed and analysed. Out of 24 CO proteins, structures of 4 proteins could be successfully predicted and validated using consensus of three structure prediction algorithms and their stability was further assessed using molecular dynamic simulations at 100 ns. Phylogenetic analysis of these proteins is performed to examine their molecular evolution.

© 2021 Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Circadian oscillators (CO) play important role in the internal time estimation in an organism by synchronizing the biological events in co-ordination with day and night cycles [1]. Circadian oscillators are crucial for several processes that include growth, development and other physiological events under various environmental conditions [2]. Circadian rhythm is an internal time-keeping mechanism by which all organisms including plants anticipate the environmental changes like daily light–dark cycles, temperature fluctuations to synchronise internal biological events with these external changes. These rhythms are ubiquitous and are helpful in expressing several temporally separated biological traits through the cycles spanning 24 h' time period [3]. Eukaryotic circadian oscillators rely on transcriptional and translational feedback loops [4,5] regulated by more than 20 transcription factors that

constitute core plant clock network [6]. Early in the morning before sunrise, CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) and LONG ELONGATED HYPOCOTYL (LHY) genes downregulate the expression of PSEUDO-RESPONSE REGULATOR 5 (PRR5), PRR7, PRR9, TIMING OF CAB EXPRESSION 1 (TOC1/PRR1), GIGANTEA (GI) and genes belonging to evening complex (EC), namely, LUX ARRHYTHMO (LUX), EARLY FLOWERING 3 (ELF3), ELF4 [6]. PRR (morning and midday phase), CCA1 HIKING EXPEDITION (CHE) genes subsequently express and restrict the CCA1 and LHY expression by repressing them during day time. At afternoon, REVEILLE 8 (RVE8), RVE4, NIGHT LIGHT-INDUCIBLE AND CLOCK-REGULATED GENE 1 (LNK1) and LNK2 factors induce expression of PRR5, TOC1 and ELF4 genes [6]. Moreover, at evening TOC1 negatively regulate GI, LUX, ELF4 and PRR5 [2]. EC maintains negative regulation of GI, PRR7 and PRR9's expression that, in turn, alleviate CCA1 and LHY repression during the night.

The growth and development of a plant is directly affected by these rhythms that are helpful in modulating various behavioural activities [7]. There are several studies in which plant COs are reported to regulate important physiological processes like

\* Corresponding author.

E-mail address: [vikramsingh@cuhimachal.ac.in](mailto:vikramsingh@cuhimachal.ac.in) (V. Singh).

<sup>1</sup> Equal contribution.

photosynthesis, flowering time, root growth, sugar metabolism, hormonal signalling, nutrient biogenesis and plant immunity [8]. Under free-running (lacking any external cues) conditions, behavioural activities of plants are well-studied and it is being understood that multiple clocks in different tissues and even within the cells operate in unison, thereby increasing the circadian system complexity [9]. To understand the inner working of oscillations in the model plant *Arabidopsis thaliana*, several approaches including mathematical modelling and phenotypic expression were employed in last decade.

Tea is among the most popular non-alcoholic beverages that is obtained from newborn leaves of the plant *Camellia sinensis*. Due to its flavour and health promoting functions, tea is the most consumed drink after water across the world [10]. To meet the consumer demands, studies based on the combination of various genetic and environmental factors are being performed to increase the tea production. Using different tea processing methods, a variety of tea products are also being developed having several palatabilities. Various transcriptomic studies have been conducted to examine the stress specific expressions, and recently a high quality draft genome of tea has also been sequenced and annotated for understanding its specific traits [11].

In the present study, we have made *in-silico* efforts for the systems-level mining and characterization of genes inducing circadian rhythms in the tea plant. To explore and examine the systems level architecture of protein–protein subnetwork of circadian rhythm clock associated genes (TeaCR-PIN), genome-wide interologous interactome map (TeaGPIN) of tea was considered [12]. Further, all the interacting proteins of developed subnetwork (nodes and edges) were selected for co-expression and differential-expression analysis by examining their expression in 3 abiotic and 3 biotic stress conditions. Furthermore, pathway enrichment analysis was also performed for the co-expressed modules of TeaCR-PIN [13]. Moreover, structural prediction and analysis of these proteins was performed and stability of proteins was accessed by molecular dynamics and simulations. Additionally, phylogenetic analysis was also performed in order to find the most accurate evolutionary estimates of these proteins in closely associated plant species.

## 2. Materials and methods

### 2.1. 2.1. Data extraction

Proteome information of 56 plants with available genomes was downloaded from UniProt database (<https://www.uniprot.org/>). Based on the available information in literature about proteins related to circadian rhythms in two widely studied plant species *Arabidopsis thaliana* and *Oryza sativa*, 24 proteins were enlisted [2,6]. Sequences of all the selected proteins were retrieved and subjected to Blast-P with all the downloaded plants having genome information. For each protein, top hits having at least 40% sequence identity, 50% query coverage and E – value  $\leq 10^{-5}$  were selected and aligned using Clustal Omega [14]. Hidden Markov models for each selected protein were built using HMMER (<http://hmm-janelia.org/>). Furthermore, tea proteins were downloaded from TPIA database [15] and all the protein sequences were scanned at domain level with in-house made HMM profiles to identify the core circadian oscillator proteins in *Camellia sinensis*. Each CO protein was then queried to InerProScan v5.48–83 (<https://www.ebi.ac.uk/interpro/>) to identify their respective domains.

### 2.2. Expression analysis of core clock genes

The expression analysis of identified CO proteins was carried out by leveraging the RNA Seq data of *C. sinensis* transcriptomes

(Supplementary Table S1) obtained from 3 abiotic (SRP047312 [16]; SRP055910 [17]; ERP012919 [18]) and 3 biotic stress conditions (SRP060335 [19]; SRP063593 [20]; SRP067826 [21]). All pair-end filtered reads of 44 samples were mapped separately to CO proteins using BOWTIE2 tool [22]. Further, for calculating the relative expression of mapped CO genes, normalized transcripts per kilobase million (TPM) values per sample were obtained for each selected protein [23]. Furthermore, expression analysis of obtained CO genes was conducted by calculating Pearson's correlation coefficient (PCC) for the abiotic and biotic conditions, separately. Pairwise co-expression for all the 24 genes was computed using Pearson correlation coefficient (PCC)

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

where n is the number of samples (21 for abiotic and 23 for biotic conditions), x and y represent vectors of RPKM values corresponding to genes x and y.

In order to check if there exists some characteristic expression profile specific to abiotic (A) and biotic (B) stress types, we computed Pearson's correlation coefficient (PCC) for each gene-pair to establish a stress type specific expression relationship, where all the stress conditions belonging abiotic stress are categorized as stress type A while all the conditions belonging to biotic stress are categorized as stress type B. PCC was computed for every gene pair in stress types A and B, separately. A pair of genes found to share similar expression patterns in both stress types (*i.e.* same pair found to have  $r \geq 0.5$  in both stress types A and B or  $r \leq -0.5$  in both stress types A and B representing that both genes are either simultaneously upregulated or downregulated) was considered to be stress type specific co-expressed (STSCE) in the expression matrix. A gene pair was considered to be stress type specific differential expressed (STSDE), if

- (i) the  $r$  value for that pair is found to be high positive ( $r \geq 0.5$ ) in one type of stress (either A or B) and high negative ( $r \leq -0.5$ ) in the other stress type.
- (ii) the  $r$  value for that pair is found to be high ( $|r| \geq 0.5$ ) in one type of stress (either A or B) and low ( $|r| \leq 0.2$ ) in the other stress type.

All the gene pairs were termed as neutrally expressed if they have weak correlation ( $|r| \leq 0.2$  in both the types A and B) and considered to have stress type independent expression (STIE).

### 2.3. Construction and analysis of PPI sub-network of CO proteins (TeaCO-PIN)

Additionally, to find the interacting partners of CO proteins, genome wide interologous protein–protein interaction network (TeaGPIN) was selected. By mapping all the CO proteins to TeaGPIN [12], a protein–protein interaction subnetwork of these proteins (TeaCO-PIN) was developed that was visualized and analyzed using cytoscape [24]. Moreover, TeaCO-PIN was also subjected to weighted gene co-expression network analysis (WGCNA) tool [25] for identifying the co-expressing modules. DAVID Bioinformatics Resources 6.8 [26] was used to perform pathway enrichment analysis of these modules.

### 2.4. Functional annotation of clock associated proteins

For functional categorization of the identified proteins associated with core clock in *Camellia sinensis*, each of protein was subjected to Blast-P with non-redundant (NR) database at NCBI [27], UniProt database [28] and TAIR database [29] with E-value of  $10^{-5}$ . Further, gene ontology (GO) terms for CO proteins were

predicted using AgriGo database [30] and WEGO tool [31]. For gaining the information of particular pathways in which these proteins are associated with, all the obtained sequences were mapped to KEGG database [32].

### 2.5. In-silico structural modeling of CO proteins

To find the three-dimensional structural patterns of CO proteins, homology modeling method was considered. All the proteins were subjected to Blast-P with protein data bank (PDB) to identify the templates for structure identification. Three structure prediction softwares namely, modeller 9.21 [33], CPH models [34] and Phyre2 tools [35] were used to build structures of all these proteins. All the modeled proteins were subjected to validation by structure analysis and verification server (SAVES) to find the structural errors like overall stereochemical quality, macromolecular volume, parameters of residues, non-bonded interactions and compatibility of models [36–38]. To find the residues in most favored regions, PROCHECK analysis was performed to analyze the Ramachandran plot [39].

### 2.6. Molecular dynamics simulation

Proteins with validated structures after homology modeling were selected for molecular dynamics and simulation by GROMACS 5.0 (Groningen Machine for Chemical Simulation) package using the GROMOS96 53a6 force field [40]. Command `pdb2gmx` was used in order to generate protein topology files. Solvation of proteins was performed and solvated proteins were fixed in cubic box with a distance of 1.0 nm between the protein surface and edges of the box. The particle-mesh ewald (PME) electrostatic and periodic boundary conditions were then applied in all the directions [40]. As per the requirements in the protein under consideration, the entire system was neutralized by adding  $\text{Na}^+$  counter ions. To avoid the steric clashes and high energy interactions, energy minimization was performed for the whole system at 50,000 steps of steepest descent. During MD simulations, production and equilibration phases were configured. To equilibrate the system, it was administered to the simulations at constant number, volume and temperature [NVT] and constant number, pressure and temperature [NPT] at 300 K for 100 ps. Finally, every system was subjected to MD production run at 1 bar pressure and 300 K temperature for 100 ps. The atom coordinates were recorded at every 100 ns throughout the MD simulation.

### 2.7. Phylogenetic analysis of CO proteins

Top hits for all the four structurally validated CO proteins after Blast-P results with downloaded plant genomes were considered in order to extract their sequences. We obtained a multiple sequence alignment (MSA) for each protein by aligning its homologs in the template organisms using ClustalW software with default parameters [41]. Phylogenetic trees from each alignment were then constructed, using the neighbor-joining (NJ) method that uses the Jones-Thornton-Taylor (JTT) model, 1,000 bootstraps and Poisson model with uniform rates for substitution as well as pairwise deletion.

## 3. Results and discussions

While the demand of *Camellia sinensis* (tea) is continuously increasing, this crop is also under continuous threat due to various abiotic and biotic stresses [12]. The substantial role of circadian rhythms in determining the results of host-pathogen interactions in plants, and hence contributing towards increase in the fitness

of plants is being widely studied [42]. Due to irregular environmental conditions, plants are facing continuous stress to survive. In response to particular stress (abiotic or biotic), circadian rhythm induces the expression in particular genes of interest as demonstrated in *Arabidopsis thaliana*, where genes are expressed rhythmically [43–45]. Thus, we attempt to identify the proteins related to circadian rhythm and their expression through an exhaustive computational framework. The overall workflow chart is given in Fig. 1.

### 3.1. Data mining and analysis

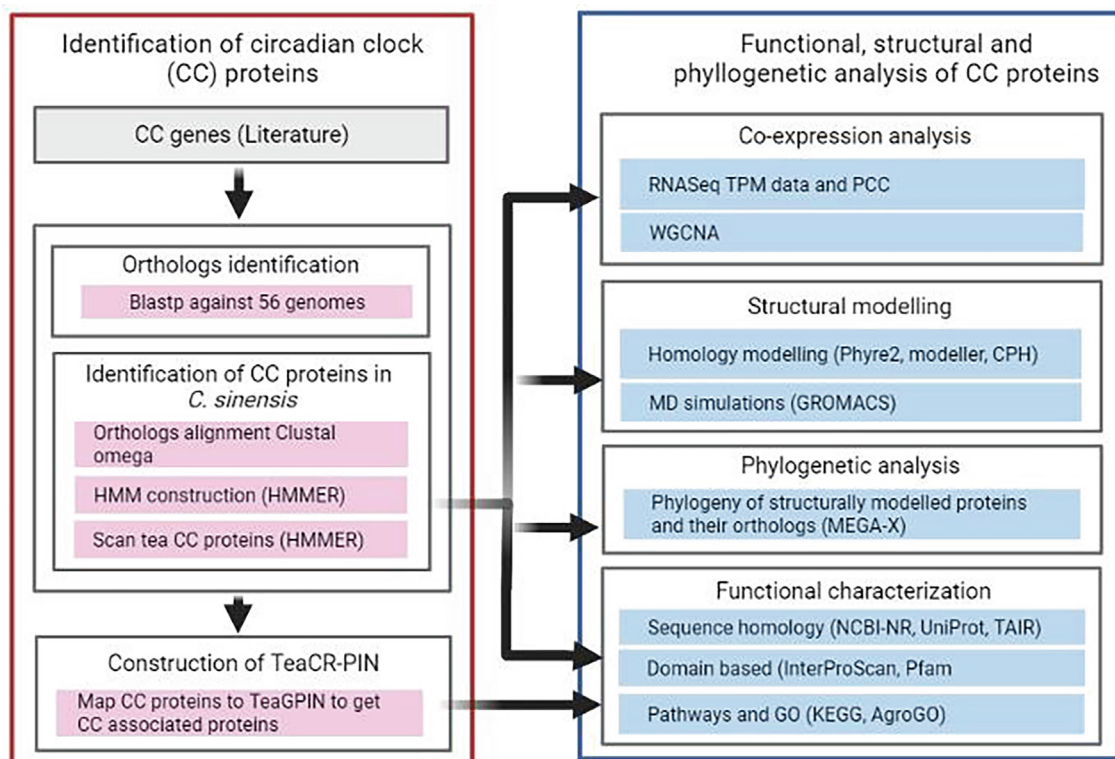
By the extensive search of literature, 24 proteins related to core CO were selected and an attempt has been made to identify these proteins in *Camellia sinensis*. We have selected 56 plants with available genome information and searched all the identified genes in the proteomes of these plants through Blast-P. For the selected 24 proteins, a total of 2,543 hits were found in the 56 plants under study at an  $E$ -value  $\leq 10^{-5}$  (see Supplementary Table S2). By imposing a criterion of at least 40% sequence identity and 50% query coverage we selected 963 top-hits corresponding to all the CO proteins, separately. HMM profiles for all the 24 CO proteins were built by extracting their sequences from all the plant species under consideration in which hits were found to meet the above criteria. The *Camellia sinensis* proteome was scanned against these HMMs, and on the basis of returned top hits, 24 proteins related to circadian rhythm were identified in tea for the further analysis (details are given in Supplementary Table S3). Since, gene LWD1 is partially redundant with LWD2, CCA1 with LHY, RVE8 with RVE4 and RVE6, different members of PRR family and LUX with NOX and they express during different times of the day or form complexes [6] so we selected the second best hit for LWD2, REV4, PRR5, REV6, NOX, PRR3 and LHY proteins. These results are further confirmed by NCBI-Nr, GO (Supplementary Fig. S1) and KEGG annotations (Supplementary Table S4). Pfam, InterProScan databases also confirmed the presence of respective domains in the predicted CO proteins (Supplementary Table S4).

### 3.2. Expression analysis of core clock genes

Expression patterns emerging from expression profiles of all the CO genes were analysed separately (3 abiotic and 3 biotic stress conditions). After mapping of pair end reads, we obtained TPM values for every gene in order to normalize the total number of mapped reads and reads mapped to a gene [23]. The expression profile of each CO gene in the form of heat map in all the 44 samples is given in Fig. 2A. As can be seen from Fig. 2A, expression patterns of considered proteins are not consistent throughout the various conditions selected in this work as the genes are having either high-expression or low-expression with respect to time at different conditions of abiotic and biotic stresses. So, we were interested to figure out if there are some gene pairs which are having distinctive expression patterns in abiotic v/s biotic conditions. For that, expression profiles of CO genes were analysed by computing pairwise Pearson's correlation coefficient for abiotic and biotic conditions separately. Entire data is given in the Supplementary Table S5. It was interesting to observe that there exist clearly distinctive expression patterns among various gene-pairs in abiotic v/s biotic stress conditions. In Fig. 2B, we give representative pairs of genes belonging to various classes. In the following, we discuss distinctive expression profiles of some of those gene pairs.

Among 276 protein pairs 40 pairs are STSCE, 43 pairs are STSDE, 21 pairs are STSIE and 172 pairs are found to have moderate correlation. Protein pair TEA012471.1 (LNK2) and TEA010432.1 (LNK1) was found to have high positive PCC in both abiotic and biotic conditions respectively and hence belong to stress type specific co-





**Fig. 1.** Overall methodology for genome wide identification, characterization (functional and network level), co-expression analysis and structural modelling of circadian rhythm proteins in *Camellia sinensis*.

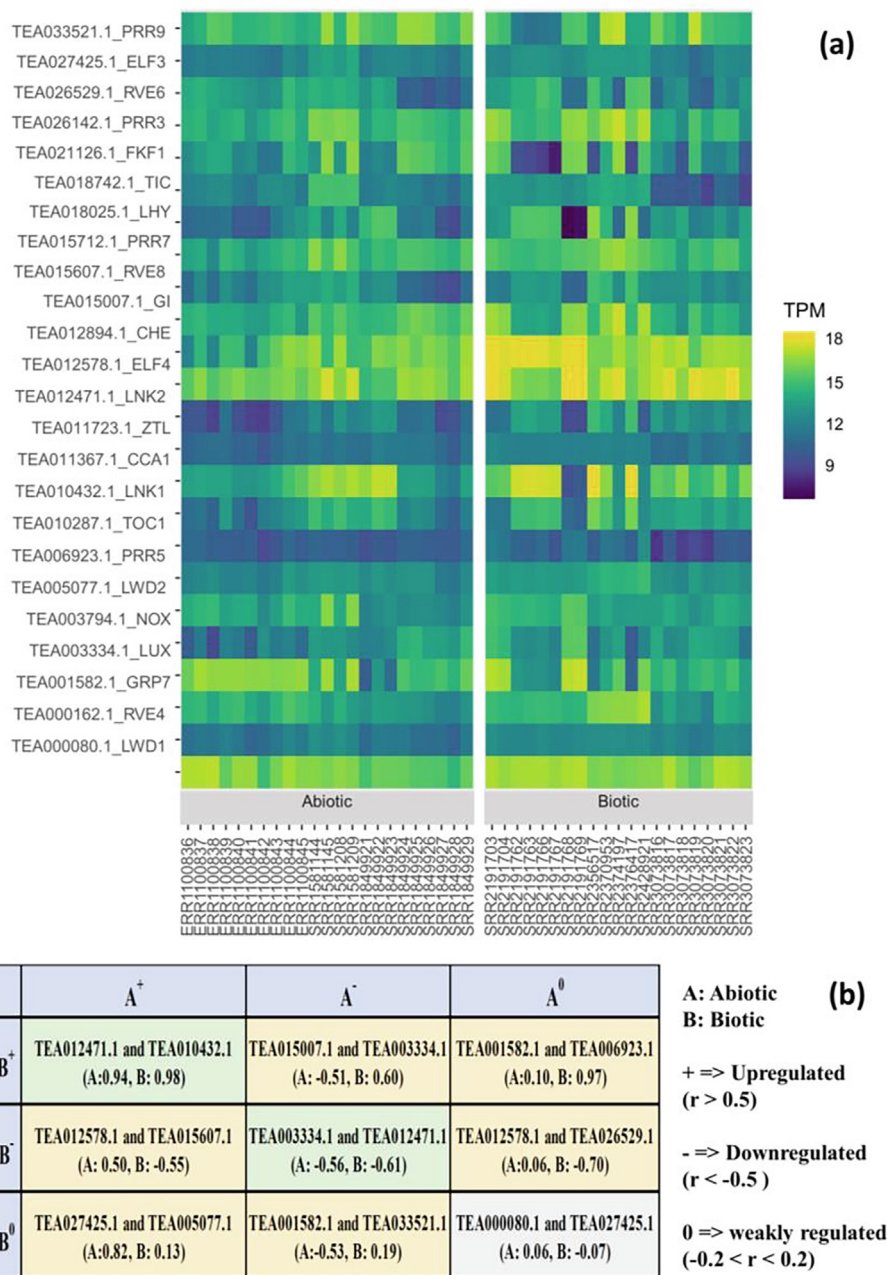
expression (STSCE) category. Night light-inducible and clock-regulated 1 (LNK1) and LNK2 genes express during day time and function to integrate the light signals with the CO [6]. LNKs act as transcriptional coactivators of afternoon (PRR5) and evening phased genes including TOC1 and ELF4 [46]. Both these LNKs are partially redundant which regulate important physiological events including internal time keeping by linking circadian and light (especially phytochrome) signaling and increase plant fitness in both abiotic and biotic pathologies [47]. Similarly, TEA018742.1 (TIC) and TEA027425.1 (ELF3) was also found to belong to STSCE category, as both of these proteins were having PCC values greater than + 0.5. Reactive oxygen species (ROS) production has been reported previously in response to both abiotic and biotic stresses. They act as signaling molecules to strengthen plant defence against the induced stress. ROS gene expression is gated by core clock genes [48]. Up-regulations of TIC and ELF3 have also been previously reported during the stress conditions [49] and mutant plants for the respective genes have shown to regulate ROS stress responses [50]. This clearly suggests that both ELF3 and TIC are stress type specific co-expressing genes, their relationship is well captured by our method.

Furthermore, protein pair TEA012578.1 (ELF4) and TEA011367.1 (CCA1) was found to be stress type specific differential-expression (STSDE) as it weakly correlated (0.177) in abiotic conditions while strongly but negatively correlated (-0.74) in biotic stress conditions. It has been established that CCA1 and LHY complex represses the transcription of evening complex (EC) including LUX, NOX, ELF3, ELF4 by binding to evening elements in their promoter regions. EC provides a positive feed-back to CCA1 by repressing the expression of PRR9 and PRR7 at late night and thereby indirectly promoting CCA1 expression [6]. Since most of the biotic stresses occur under natural conditions, ELF4 and CCA1 expressions are negatively regulated. However, rhythmic expression of CBF1, 2, 3 (low temperature responsive) genes has

been shown to be regulated by CCA1 [51] only. Because expression of stress responsive genes is gated by environmental signals, under abiotic stresses a weak correlation between ELF4 and CCA1 has been observed. Protein pair TEA015712.1 (PRR7) and TEA033521.1 (PRR9) also belongs to STSDE () category. In the current study, we have considered abiotic stresses, like, temperature, draught, winter dormancy. It has been shown that abscisic acid (ABA) accumulates under these types of stresses which is gated by circadian oscillator genes [52]. PRR family genes including PRR5, PRR7 and PRR9 have been proposed to control this rhythmic production of ABA [53]. Moreover, PRR9 is morning phased gene while PRR7 is a midday phased gene which expresses sequentially during the day and suppress each other's expression. At morning PRR9 expresses and with progression of day its level starts to diminish with elevated level of PRR7 expression. As biotic stresses mostly occur under free-running-conditions so PRR genes follow their natural temporal pattern of expression [6]. Thus, these pairs of proteins may be considered as the key biomarkers with respect to both types of stresses (abiotic as well as biotic).

### 3.3. Identification and characterization of proteins interacting with tea circadian rhythm clock proteins (TeaCO-PIN)

In order to find the interacting partners of CO proteins, all the identified proteins were mapped to genome wide interologous network (TeaGPIN) [12]. A total of 12 proteins were successfully mapped that were found to interact with 166 other proteins (Supplementary Table S5). Incorporating the first-degree interactions of all these nodes, an interaction subnetwork having 260 edges was developed and is termed as TeaCO-PIN. Among the 12 mapped core clock proteins, 11 were found to directly interact with each other through 15 interactions. Protein pair TEA00080.1, TEA005077.1 was found to be LWD1 and LWD2 respectively, both of which are associated with transmission of light signals to CO and transcrip-



**Fig. 2.** (a) Expression analysis (TPM values) of 24 te CO genes in RNA-Seq datasets of three abiotic (23 samples) and three biotic (21 samples) stress conditions. (b) Representative gene pairs belonging to various stress type specific expressions. STSCE pairs are in green colored cells, STSDE pairs are in yellow colored cells and STIE pair is in grey colored cell. In each cell, PCC values of the proteins in A as well as B types of stresses are mentioned in parenthesis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

tional activation of CCA1, PRR9, PRR5 and PRR1 at morning [54]. We also obtained protein GI (TEA015007.1) paired with LUX (TEA003334.1), TOC1 (TEA010287.1) and FKF1 (TEA021126.1) proteins. GI is an evening acting gene that is repressed in morning by CCA1 and LHY complex while in the evening by LUX and TOC1. GI acts as an integral part of core oscillator as it is involved in processes like providing diurnal and temperature cues to the core clock, photoperiodic flowering, abiotic stresses etc [55]. Core clock protein CCA1 (TEA011367.1) has been found to be paired with TOC1 (TEA010287.1), PRR7 (TEA015712.1) and FKF1 (TEA021126.1). The first transcription translation feedback loop is comprised of three proteins CCA1, LHY and TOC1. CCA1 and LHY complex repress the expression of TOC1 while the latter has been shown

to repress CCA1 and LHY expression [56]. Other members of PRR family have also been repressed by CCA1 and LHY which subsequently express during the day and in-turn repress the expression CCA1 and LHY complex [56]. Furthermore, we obtained REV8 (TEA015607.1) TOC1 (TEA010287.1) pair and RVE8 has been found to be a direct activator of TOC1 while TOC1 repress the expression of RVE8 [57]. LUX (TEA003334.1) protein has been found to interact with TOC1 (TEA010287.1) FKF1 (TEA021126.1) while it has been shown that TOC1 directly repress transcription of LUX [6,58]. Clock protein TOC1 (TEA010287.1) has been found to be interacting with NOX (TEA003794.1), FKF1 (TEA021126.1) and CHE (TEA012894.1). NOX binds with LUX in evening complex and is also repressed by TOC1 [6,59]. As can be observed from these

results most of the core clock has been reconstructed from the protein–protein interaction network showing accuracy of the method.

Furthermore, 166 direct interactors of core clock proteins are subjected to functional annotations through Blast-P with NCBI non-redundant database and TAIR database (Supplementary Table S6). Then domain-based characterization was carried out using InterProScan and Pfam databases followed by KEGG and Gene Ontology (GO) annotations (Supplementary Table S6). We found that most of the proteins are associated with the input and downstream pathways (output pathways) to core circadian oscillator. The oscillator is entrained mainly by environmental cues like light and temperature. Light signals are transduced to the CO in plants are primarily by phytochromes, cryptochromes and ZEITLUPE (ZTL)/F-BOX1 (FKF1)/LOV KELCH PROTEIN2 (LKP2) family proteins [60]. Phytochromes perceive red light while cryptochromes and proteins of ZTL family are both blue light photoreceptors [61]. There exist five variants of phytochromes (phyA-E) in *Arabidopsis* among which mainly phyA and phyB entrain core clock under red light [62]. We obtained proteins TEA031363.1 and TEA019169.1 which are predicted to be phyB. They interact with TEA015007.1 (GI), TEA010287.1 (TOC1), TEA021126.1 (FKF1) and TEA011367.1 (CCA1). PhyB has been shown earlier that it interacts CCA1, TOC1, GI, LUX and these interactions are under the regulation of red light [63]. We have also predicted other phytochrome proteins including two phyA TEA005460.1, TEA002223.1 proteins and one TEA001678.1 phyE protein. Proteins TEA032791.1, TEA023230.1, TEA012616.1, TEA012558.1 are predicted to be cry1 and TEA009050.1 as cry2 which perceive blue light and transduce that signal to core oscillator [61]. TEA010741.1, TEA033721.1 and TEA033133.1 are predicted as proteins containing basic helix-loop-helix (bHLH) domain and are found to interact with proteins TEA000080.1 (LWD1), TEA005077.1 (GI), TEA011367.1 (CCA1). It has been reported that proteasomal degradation of GI and ELF3 is performed by CONSTITUTIVE PHOTOMORPHOGENIC1 (COP1), E3 ubiquitin ligase activity of that is repressed by both phytochromes and cryptochromes in a light dependent manner [64–66]. PhyB repress expression of COP1 and PHYTOCHROME INTERACTING FACTORS (PIFs), transcription factors belonging to the basic helix-loop-helix (bHLH) family [67]. Proteins TEA022021.1 (cullin-associated and neddylation dissociated), TEA018433.1 (defective in cullin neddylation protein *i.e.* DUF298), TEA031936.1 (cullin1), TEA020303.1 (cullin4), TEA013377.1 (flavin-binding 2C kelch repeat) and both TEA017928.1, TEA000229.1 (kelch repeat) act as E3 ubiquitin ligases in ZTL mediated light signal transduction to core clock [68,69].

The signal transduced by these input pathways is conveyed to central oscillator that controls the rhythms of various output pathways. We obtained members of CONSTAS family proteins namely CONSTANS-like 2 (TEA016175.1 and TEA027544.1), CONSTANS-like 9 (TEA018780.1 and TEA016616.1) and CONSTANS-LIKE 4-like (TEA014116.1). All of these proteins are observed to interact with either or both of the TEA015007.1 (GI) TEA021126.1 (FKF1) proteins in TeaCRPIN. The CYCLING DOF FACTOR (CDF) family proteins have been reported to suppress transcription of CO which are under the control of blue light photoreceptor FKF1 [70]. FKF1 perceives blue light and forms an active complex with GI, FKF1-GI, that recognises CDFs family proteins as substrates for proteasomal degradation [71,72]. Once activated CO proteins induces the expression of FLOWERING LOCUS T (FT), predicted to be TEA015714.1, transcripts which act as florigen and control the expression of floral identity genes like APETALA1 And LEAFY [73]. We were able to map most of enzymes involved in anthocyanin (flavonoids) biosynthesis pathway, that has shown to be under the control of core circadian oscillator through CHALCONE SYNTHASE (TEA034019.1 and TEA023333.1) [74]. Other pathway proteins include chalcone isomerase (TEA033023.1), flavanone 3-

hydroxylase (TEA023790.1), dihydroflavonol 4-reductase (TEA032730.1 and TEA024758.1), Glycosyltransferase (TEA009682.1) anthocyanidin reductase (TEA022960.1 and TEA009266.1) [75]. Anthocyanins play a wide spectrum of roles plants like; imparting colours to various parts of plants including fruit, flowers, leaves, root etc.; interactions between pollen-pistil, plant-insect; stress tolerance including both abiotic (UV protection) and biotic (insect repellent) [76].

TeaCR-PIN having CO proteins and their first interactors obtained from genome wide tea interologous network (TeaGPIN) is shown in Fig. 3 and the complete information about all of its nodes and edges is given in Supplementary Table S5. Furthermore, by means of WGCNA, all the 166 proteins of TeaCR-PIN were analysed that resulted in 6 co-expressed modules. Proteins TEA021126.1 (FKF1) and TEA012894.1 (CHE) are found to be belonging to module 1; proteins TEA015712.1 (PRR7), TEA010287.1 (TOC1) and TEA001582.1 (GRP7) belong to module 2; while proteins TEA018025.1 (LHY) and TEA003334.1 (LUX) are found to be present in module 3. Pathway enrichment analysis performed by DAVID Bioinformatics Resource revealed that “circadian rhythm in plants” is highly enriched in all the 3 modules further affirming the findings of our study.

### 3.4. In-silico modeling of proteins and MD simulation

Phyre2, modeller9.21, and CPH model methods which are based on homology modeling were employed for structural modeling. Out of 24 CO proteins, structures of 4 proteins were successfully predicted after selecting suitable templates (TEA001582.1:4F02.D, TEA011723.1:5GQT.A TEA021126.1: 5GQT.A and TEA033521.1:1M5T.A). Furthermore, all these 4 models were validated by checking their stereochemical qualities using PROCHECK through Ramachandran plot analysis that revealed 100%, 99.2%, 98.9% and 98.5% residues of proteins TEA001582.1, TEA011723.1 TEA021126.1 and TEA033521.1, respectively, are in the most favored and allowed regions. All the Ramachandran plot results are shown in Supplementary Fig. S2. In order to check the compactness, stability and structural behavior of modeled proteins, molecular dynamics (MD) simulation and energy minimization was performed. Root mean square Deviation (RMSD) plots clearly revealed that structures of TEA021126.1 and TEA011723.1 are highly stable with little or no deviation. However, structures of TEA001582.1 and TEA033521.1 are found to have slight RMSD fluctuations. Initially RMSD value for TEA001582.1 fluctuates till approximately 60 ns after that no variation in RMSD is observed and the structure becomes stable. On the other hand, RMSD values for TEA033521.1 remains stable till approximately 75 ns, however, a sharp increase is observed after that suggesting unstability of the structure. Plots of RMSD values obtained for all the structures are shown in Fig. 4.

Protein TEA001582.1 encodes for glycine rich RNA-binding protein 7 (GRP7). GRP7 has important role in adaptation of cold stress by harboring RNA chaperone activity and also has crucial role in regulation of circadian rhythm by oscillating with evening peaks as reported in *Arabidopsis thaliana* [77,78]. Protein TEA011723.1 and protein TEA021126.1 encoding for flavin binding, kelch repeat, F-box1 (FKF1) regulate the degradation of other CO proteins like TOC1 and PRR5 which is crucial to determine circadian rhythm oscillation period [69]. Protein TEA033521.1 encodes for pseudo-response regulator 9 (PRR9) which is a core circadian component having important role in leaf senescence through positive regulation of ORE1 through feed-forward pathway by means of direct transcriptional and post-transcriptional regulation as reported in *Arabidopsis thaliana* [79].





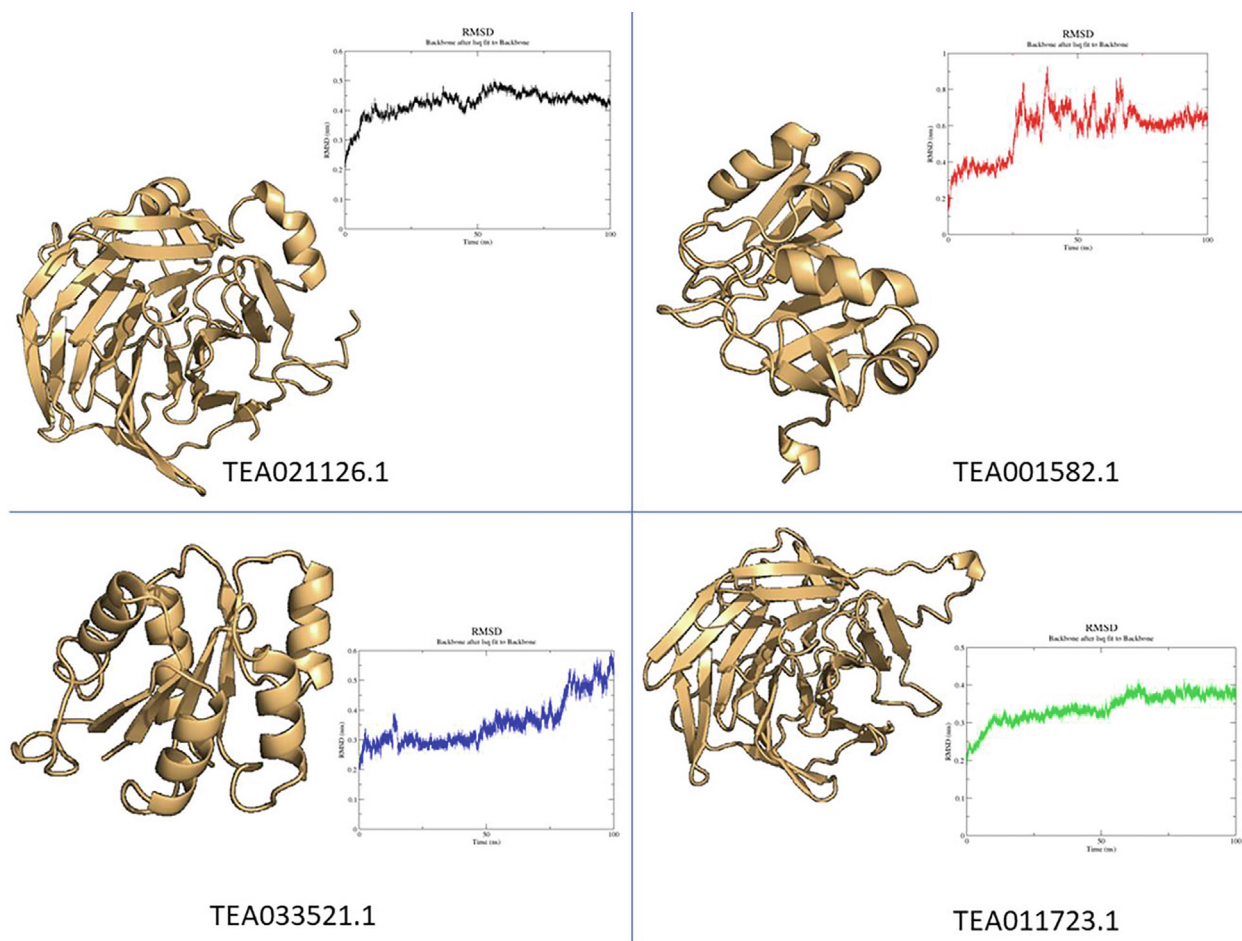


Fig. 4. Modelled structures of 4 tea CO proteins along with RMSD plots obtained from their molecular dynamics simulations at 100 ns.

the plant is subjected to stresses belonging to either A or B types, so that these patterns may be proposed as the A or B stress type markers. For that, we proposed a novel methodology, and we were able to report some pairs of CO genes which are found to have stress type specific co-expression or differential expression which may be focused in the future studies to develop transgenic varieties to encounter the A or B stress types. Additionally, during *in-silico* structural modeling, structures of 4 proteins were successfully predicted and analyzed *via* molecular dynamics (MD) simulation. Identification of core circadian rhythm proteins in tea might be helpful in balancing abiotic and biotic stress to regulate immune responses by over-expression or down-expression of corresponding genes. To the best of our knowledge, this is the first report in tea in which genome-wide circadian rhythm related proteins are identified and extensively analyzed. All the HMMs are being made available (Supplementary File S1) for the identification and characterization of proteins associated with CO pathway in newly sequenced plants, by adopting the proposed methodology.

#### Author contribution

VS\* conceptualized and designed the research framework as well as supervised the entire study. GS and VS<sup>†</sup> performed the computational experiments. GS, VS<sup>†</sup> and VS\* analysed the data, interpreted results, wrote and finalized the manuscript.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgement

We thank Central University of Himachal Pradesh for providing the required infrastructure and computational facilities. VS<sup>†</sup> thanks Council of Scientific and Industrial Research (CSIR), India for providing Senior Research Fellowship (SRF). GS thanks Mr. Ashish Panghalia for useful discussions.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2021.12.026>.

#### References

- [1] Dodd AN, Salathia N, Hall A, Kévei E, Tóth Réka, Nagy F, et al. Plant Circadian Clocks Increase Photosynthesis, Growth, Survival, and Competitive Advantage. *Science* 2005;309(5734):630–3.
- [2] Lu H, McClung CR, Zhang C. Tick tock: circadian regulation of plant innate immunity. *Annu Rev Phytopathol* 2017;55(1):287–311.



- [3] Brunner M, Káldi K. Interlocked feedback loops of the circadian clock of *Neurospora crassa*. *Mol Microbiol* 2008;68(2):255–62.
- [4] Gonze D, Goldbeter A. Circadian rhythms and molecular noise. *Chaos* 2006;16(2):026110. <https://doi.org/10.1063/1.2211767>.
- [5] Singh RB, Singh V, Ramaswamy R. Stochastic synchronization of circadian rhythms. *J Syst Sci Complex* 2010;23(5):978–88. <https://doi.org/10.1007/s11424-010-0208-x>.
- [6] Nohales MA, Kay SA. Molecular mechanisms at the core of the plant circadian oscillator. *Nat Struct Mol Biol* 2016;23(12):1061–9.
- [7] Greenham K, McClung CR. Integrating circadian dynamics with physiological processes in plants. *Nat Rev Genet* 2015;16(10):598–610.
- [8] De Caluwé J, Xiao Q, Hermans C, Verbruggen N, Leloup J-C, Gonze D. A compact model for the complex plant circadian clock. *Front Plant Sci* 2016;7:74.
- [9] Endo M, Shimizu H, Nohales MA, Araki T, Kay SA. Tissue-specific clocks in *Arabidopsis* show asymmetric coupling. *Nature* 2014;515(7527):419–22.
- [10] Danesi F, Saha S, Kroon PA, Glibetić M, Konić-Ristić A, D'Antuono LF, et al. Bioactive-rich *Sideritis scardica* tea (mountain tea) is as potent as *Camellia sinensis* tea at inducing cellular antioxidant defences and preventing oxidative stress. *J Sci Food Agric* 2013;93(14):3558–64.
- [11] Wei C, Yang H, Wang S, Zhao J, Liu C, Gao L, et al. Draft genome sequence of *Camellia sinensis* var. *sinensis* provides insights into the evolution of the tea genome and tea quality. *Proc Natl Acad Sci* 2018;115(18):E4151–8.
- [12] Singh G, Singh V, Singh V. Genome-wide interologous interactome map (TeaGPIIN) of *Camellia sinensis*. *BioRxiv* 2019:696062. <https://doi.org/10.1101/696062>.
- [13] Singh V, Singh G, Singh V. TulsipIN: an interologous protein interactome of *Camellia sinensis*. *BioRxiv* 2019:680025. <https://doi.org/10.1101/680025>.
- [14] Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011;7(1):539. <https://doi.org/10.1038/msb.2011.75>.
- [15] Xia E, Li F, Tong W, Li P, Wu Q, Zhao H, et al. Tea Plant Information Archive (TPIA): A comprehensive genomics and bioinformatics platform for tea plant. *Plant Biotechnol J* 2019.
- [16] Paul A, Jha A, Bhardwaj S, Singh S, Shankar R, Kumar S. RNA-seq-mediated transcriptome analysis of actively growing and winter dormant shoots identifies non-deciduous habit of evergreen tree tea during winters. *Sci Rep* 2014;4:5932.
- [17] Li C-F, Xu Y-X, Ma J-Q, Jin J-Q, Huang D-J, Yao M-Z, et al. Biochemical and transcriptomic analyses reveal different metabolite biosynthesis profiles among three color and developmental stages in 'Anji Baicha' (*Camellia sinensis*). *BMC Plant Biol* 2016;16(1). <https://doi.org/10.1186/s12870-016-0885-2>.
- [18] Shi J, Ma ChengYing, Qi DanDan, Lv HaiPeng, Yang T, Peng QunHua, et al. Transcriptional responses and flavor volatiles biosynthesis in methyl jasmonate-treated tea leaves. *BMC Plant Biol* 2015;15(1). <https://doi.org/10.1186/s12870-015-0609-z>.
- [19] Zhang Q, Cai M, Yu X, Wang L, Guo C, Ming R, et al. Transcriptome dynamics of *Camellia sinensis* in response to continuous salinity and drought stress. *Tree Genet Genomes* 2017;13(4). <https://doi.org/10.1007/s11295-017-1161-9>.
- [20] Li X, Lin Y, Zhao S, Zhao X, Geng Z, Yuan Z, et al. Transcriptome changes and its effect on physiological and metabolic processes in tea plant during mechanical damage. *For Pathol* 2018;48(4):e12432. <https://doi.org/10.1111/efp.2018.48.issue-410.1111/efp.12432>.
- [21] Jayaswami K, Mahajan P, Singh G, Parmar R, Seth R, Raina A, et al. Transcriptome analysis reveals candidate genes involved in blister blight defense in tea (*Camellia sinensis* (L) Kuntze). *Sci Rep* 2016;6(1). <https://doi.org/10.1038/srep30412>.
- [22] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9(4):357–9.
- [23] Wagner GP, Kin K, Lynch VJ. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci* 2012;131(4):281–5.
- [24] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13(11):2498–504. <https://doi.org/10.1101/gr.1239303>.
- [25] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf* 2008;9:559.
- [26] Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2008;37:1–13.
- [27] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990;215(3):403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- [28] Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, et al. UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 2004;32:D115–9.
- [29] Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, et al. The *Arabidopsis* information resource: making and mining the “gold standard” annotated reference plant genome. *Genesis* 2015;53:474–85.
- [30] Du Z, Zhou X, Ling Y, Zhang Z, Su Z. agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res* 2010;38:W64–70.
- [31] Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, et al. WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res* 2006;34(Web Server):W293–7.
- [32] Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 2017;45(D1):D353–61.
- [33] Webb B, Sali A. Protein structure modeling with MODELLER. *Funct. Genomics. Springer*; 2017. p. 39–54.
- [34] Nielsen M, Lundegaard C, Lund O, Petersen TN. CPHmodels-3.0—remote homology modeling using structure-guided sequence profiles. *Nucleic Acids Res* 2010;38:W576–81.
- [35] Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 2015;10(6):845–58. <https://doi.org/10.1038/nprot.2015.053>.
- [36] Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* 1993;2(9):1511–9.
- [37] Hooft RWW, Vriend G, Sander C, Abola EE. Errors in protein structures. *Nature* 1996;381:272.
- [38] Pontius J, Richelle J, Wodak SJ. Deviations from standard atomic volumes as a quality measure for protein crystal structures. *J Mol Biol* 1996;264(1):121–36.
- [39] Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 1993;26(2):283–91.
- [40] Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: a message-passing parallel molecular dynamics implementation. *Comput Phys Commun* 1995;91(1–3):43–56.
- [41] Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 2018;35:1547–9.
- [42] Roden LC, Ingle RA. Lights, rhythms, infection: the role of light and the circadian clock in determining the outcome of plant–pathogen interactions. *Plant Cell* 2009;21(9):2546–52.
- [43] Sauerbrunn N, Schlaich NL. PCC1: a merging point for pathogen defence and circadian signalling in *Arabidopsis*. *Planta* 2004;218(4):552–61.
- [44] Weyman PD, Pan Z, Feng Q, Gilchrist DG, Bostock RM. DEA1, a circadian- and cold-regulated tomato gene, protects yeast cells from freezing death. *Plant Mol Biol* 2006;62(4–5):547–59.
- [45] Michael TP, Breton G, Hazen SP, Priest H, Mockler TC, Kay SA, et al. A morning-specific phytohormone gene expression program underlying rhythmic plant growth. *PLoS Biol* 2008;6(9):e225.
- [46] Xie Q, Wang P, Liu X, Yuan Li, Wang L, Zhang C, et al. LNK1 and LNK2 are transcriptional coactivators in the *Arabidopsis* circadian oscillator. *Plant Cell* 2014;26(7):2843–57. <https://doi.org/10.1105/tpc.114.126573>.
- [47] Rugnone ML, Faigon Soverna A, Sanchez SE, Schlaen RG, Hernando CE, Seymour DK, et al. LNK genes integrate light and clock signaling networks at the core of the *Arabidopsis* oscillator. *Proc Natl Acad Sci U S A* 2013;110(29):12120–5. <https://doi.org/10.1073/pnas.1302170110>.
- [48] Spoel SH, van Ooijen G. Circadian redox signaling in plant immunity and abiotic stress. *Antioxidants Redox. Signal* 2014;20(18):3024–39. <https://doi.org/10.1089/ars.2013.5530>.
- [49] Grundy J, Stoker C, Carré IA. Circadian regulation of abiotic stress tolerance in plants. *Front Plant Sci* 2015. <https://doi.org/10.3389/fpls.2015.00648>.
- [50] Lai AG, Doherty CJ, Mueller-Roeber B, Kay SA, Schippers JHM, Dijkwel PP. Circadian Clock-Associated 1 regulates ROS homeostasis and oxidative stress responses. *Proc Natl Acad Sci U S A* 2012;109(42):17129–34. <https://doi.org/10.1073/pnas.1209148109>.
- [51] Harmer SL, Hogenesch JB, Straume M, Chang HS, Han B, Zhu T, et al. Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. *Science* (80–) 2000;290. Doi: 10.1126/science.290.5499.2110.
- [52] Lee KH, Piao HL, Kim H-Y, Choi SM, Jiang F, Hartung W, et al. Activation of Glucosidase via Stress-Induced Polymerization Rapidly Increases Active Pools of Abscisic Acid. *Cell* 2006;126(6):1109–20. <https://doi.org/10.1016/j.cell.2006.07.034>.
- [53] Fukushima A, Kusano M, Nakamichi N, Kobayashi M, Hayashi N, Sakakibara H, et al. Impact of clock-associated *Arabidopsis* pseudoresponse regulators in metabolic coordination. *Proc Natl Acad Sci U S A* 2009;106(17):7251–6. <https://doi.org/10.1073/pnas.0900952106>.
- [54] Wu JF, Wang Y, Wu SH. Two new clock proteins, LWD1 and LWD2, regulate *Arabidopsis* photoperiodic flowering. *Plant Physiol* 2008;148. Doi: 10.1104/pp.108.124917.
- [55] Park DH, Somers DE, Kim YS, Choy YH, Lim HK, Soh MS, et al. Control of circadian rhythms and photoperiodic flowering by the *Arabidopsis* GIGANTEA gene. *Science* (80–) 1999;285. Doi: 10.1126/science.285.5433.1579.
- [56] Huang W, Pérez-García P, Pokhilko A, Millar AJ, Antoshechkin I, Riechmann JL, et al. Mapping the core of the *Arabidopsis* circadian clock defines the network structure of the oscillator. *Science* (80–) 2012;335. <https://doi.org/10.1126/science.1219075>.
- [57] Rawat R, Takahashi N, Hsu PY, Jones MA, Schwartz J, Salemi MR, et al. REVEILLE8 and PSEUDO-RESPONSE REGULATOR5 form a negative feedback loop within the *Arabidopsis* circadian clock. *PLoS Genet* 2011;7(3):e1001350. <https://doi.org/10.1371/journal.pgen.1001350>.
- [58] Helfer A, Nusinow DA, Chow BY, Gehrke AR, Bulyk ML, Kay SA. LUX ARRHYTHMO encodes a nighttime repressor of circadian gene expression in the *Arabidopsis* core clock. *Curr Biol* 2011;21(2):126–33. <https://doi.org/10.1016/j.cub.2010.12.021>.
- [59] Dai S, Wei X, Pei L, Thompson RL, Liu Y, Heard JE, et al. BROTHER OF LUX ARRHYTHMO is a component of the *Arabidopsis* circadian clock. *Plant Cell* 2011;23. <https://doi.org/10.1105/tpc.111.084293>.

- [60] Inoue K, Araki T, Endo M. Integration of input signals into the gene network in the plant circadian clock. *Plant Cell Physiol* 2017;58. <https://doi.org/10.1093/pcp/pcx066>.
- [61] Somers DE, Devlin PF, Kay SA. Phytochromes and Cryptochromes in the Entrainment of the Arabidopsis Circadian Clock. *Science* 1998;282(5393):1488–90.
- [62] Devlin PF, Kay SA. Cryptochromes are required for phytochrome signaling to the circadian clock but not for rhythmicity. *Plant Cell* 2000;12(12):2499–509. <https://doi.org/10.1105/tpc.12.12.2499>.
- [63] Yeom M, Kim H, Lim J, Shin A-Y, Hong S, Kim J-I, et al. How Do Phytochromes Transmit the Light Quality Information to the Circadian Clock in Arabidopsis? *Molecular Plant* 2014;7(11):1701–4.
- [64] Yu J-W, Rubio V, Lee N-Y, Bai S, Lee S-Y, Kim S-S, et al. COP1 and ELF3 Control Circadian Function and Photoperiodic Flowering by Regulating GI Stability. *Mol Cell* 2008;32(5):617–30. <https://doi.org/10.1016/j.molcel.2008.09.026>.
- [65] Liu B, Zuo Z, Liu H, Liu X, Lin C. Arabidopsis cryptochrome 1 interacts with SPA1 to suppress COP1 activity in response to blue light. *Genes Dev* 2011;25(10):1029–34. <https://doi.org/10.1101/gad.2025011>.
- [66] Sheerin DJ, Menon C, zur Oven-Krockhaus S, Enderle B, Zhu L, Johnen P, et al. Light-activated phytochrome A and B interact with members of the SPA family to promote photomorphogenesis in arabidopsis by reorganizing the COP1/SPA complex. *Plant Cell* 2015;27(1):189–201. <https://doi.org/10.1105/tpc.114.134775>.
- [67] Leivar P, Monte E. PIFs: Systems Integrators in Plant Development. *Plant Cell* 2014;26(1):56–78.
- [68] Choi CM, Gray WM, Mooney S, Hellmann H. Composition, Roles, and Regulation of Cullin-Based Ubiquitin E3 Ligases. *Arab B* 2014;12:e0175. <https://doi.org/10.1199/tab.0175>.
- [69] Baudry A, Ito S, Song YH, Strait AA, Kiba T, Lu S, et al. F-box proteins FKF1 and LKP2 act in concert with ZEITLUPE to control Arabidopsis clock progression. *Plant Cell* 2010;22:606–22.
- [70] Fornara F, Panigrahi KCS, Gissot L, Sauerbrunn N, Rühl M, Jarillo JA, et al. Arabidopsis DOF Transcription Factors Act Redundantly to Reduce CONSTANS Expression and Are Essential for a Photoperiodic Flowering Response. *Dev Cell* 2009;17(1):75–86. <https://doi.org/10.1016/j.devcel.2009.06.015>.
- [71] Imaizumi T, Schultz TF, Harmon FG, Ho LA, Kay SA. FKF1 F-Box Protein Mediates Cyclic Degradation of a Repressor of CONSTANS in Arabidopsis. *Science* 2005;309(5732):293–7.
- [72] Sawa M, Nusinow DA, Kay SA, Imaizumi T. FKF1 and GIGANTEA Complex Formation Is Required for Day-Length Measurement in Arabidopsis. *Science* 2007;318(5848):261–5.
- [73] Shim JS, Kubota A, Imaizumi T. Circadian clock and photoperiodic flowering in arabidopsis: CONSTANS is a Hub for Signal integration. *Plant Physiol* 2017;173(1):5–15. <https://doi.org/10.1104/pp.16.01327>.
- [74] Thain SC, Murtas G, Lynn JR, McGrath RB, Millar AJ. The circadian clock that controls gene expression in Arabidopsis is tissue specific. *Plant Physiol* 2002;130. <https://doi.org/10.1104/pp.005405>.
- [75] Dao TTH, Linthorst HJM, Verpoorte R. Chalcone synthase and its functions in plant resistance. *Phytochem Rev* 2011;10(3):397–412.
- [76] Pervaiz T, Songtao J, Faghihi F, Haider MS, Fang J. Naturally Occurring Anthocyanin, Structure, Functions and Biosynthetic Pathway in Fruit Plants. *J Plant Biochem Physiol* 2017;05(02). <https://doi.org/10.4172/2329-9029.1000187>.
- [77] Kim JY, Kim WY, Kwak KJ, Oh SH, Han YS, Kang H. Glycine-rich RNA-binding proteins are functionally conserved in arabidopsis thaliana and oryza sativa during cold adaptation process. *J Exp Bot* 2010. <https://doi.org/10.1093/jxb/erq058>.
- [78] Schmal C, Reimann P, Staiger D, Rao CV. A circadian clock-regulated toggle switch explains AtGRP7 and AtGRP8 oscillations in Arabidopsis thaliana. *PLoS Comput Biol* 2013;9(3):e1002986.
- [79] Kim H, Kim HJ, Vu QT, Jung S, McClung CR, Hong S, et al. Circadian control of ORE1 by PRR9 positively regulates leaf senescence in Arabidopsis. *Proc Natl Acad Sci* 2018;115(33):8448–53.
- [80] Sachetto-Martins G, Franco LO, de Oliveira DE. Plant glycine-rich proteins: a family or just proteins with a common motif? *Biochim et Biophys Acta (BBA) - Gene Struct Express* 2000;1492(1):1–14.