Inference of Gene Regulatory Network Uncovers the Linkage Between Circadian Clock and Crassulacean Acid Metabolism in *Kalanchoë fedtschenkoi*

Robert C. Moseley[1], Francis Motta[2], Gerald A. Tuskan[3,4], Steve Haase[1], Xiaohan Yang[3,4,5*]

[1]Department of Biology, Duke University, Durham, North Carolina, USA
[2]Department of Mathematical Sciences, Florida Atlantic University, Boca Raton, Florida, 33431 USA
[3]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
[4]DOE Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
[5]The Bredesen Center for Interdisciplinary Research and Graduate Education, University of Tennessee, Knoxville, TN, USA

*** Correspondence:**
Dr. Xiaohan Yang
yangx@ornl.gov

Abstract

The circadian clock drives time-specific gene expression, allowing for associated biological processes to be active during certain times of the 24 h day. Crassulacean acid metabolism (CAM) photosynthetic plants represent an interesting case of circadian regulation of gene expression as CO_2 fixation and stomatal movement in CAM plants display strong circadian dynamics. The molecular mechanisms behind how the circadian clock enabled these physiological differences is not well understood. Therefore, we set out to investigate whether core circadian elements in CAM plants were re-phased during evolution, or whether networks of phase-specific genes were simply connected to different core elements. We utilized a new metric for identifying candidate core genes of a periodic gene network and then applied the Local Edge Machine (LEM) algorithm to infer regulatory relationships between the candidate core clock genes and orthologs of known core clock genes in *K. fedtschenkoi*. We also used LEM to identify stomata-related gene targets for *K. fedtschenkoi* core clock genes and constructed a subsequent gene regulatory network. Our results provide new insights into the mechanism of circadian control of CAM-related genes in *K. fedtschenkoi*, facilitating the engineering of CAM machinery into non-CAM plants for sustainable crop production in water-limited environments.

1 Introduction

The circadian clock is a vital time-keeping mechanism that synchronizes periodic environmental signals to an organism's physiology, allowing for biological processes to function in a timely manner. This mechanism is very important in plants due to their sessile nature.

Numerous environmental signals and stressors to plants are cyclic, such as light availability, temperature, and predation. The circadian clock thus enables plants to activate the appropriate processes in response to these repeating variables.

Plants that photosynthesize through crassulacean acid metabolism (CAM) are great examples of how plants synchronize biological processes to their environment. CAM plants exhibit improved photosynthetic efficiency due to a temporal separation of CO_2 fixation and improved water-use efficiency due to inverted stomatal conductance, relative to C_3 plants [1]. These traits coupled with the global issue of increased frequency and intensity of drought [2,3] have generated an increase in CAM research with the goal of engineering these traits into C_3 plants, to enable better drought responses and/or improved drought tolerance [4,5]. Currently, it is theorized that the temporal separation of CO_2 fixation is under control of the circadian clock [6,7] and that the inverted stomatal conductance could be a result of a change in clock regulation [8]. However, the events that lead to these drastic physiological differences seen in CAM plants via the circadian clock is not well understand.

To explain potential ways the circadian clock was involved in the evolution of CAM, the first step is to understand components that make up the circadian clock network. The circadian clock is a complex network that has at its center a small regulatory network of core clock genes, generally referred to as the core clock [9]. Core clock genes are defined as highly connected transcription factors (TFs) which subsequently create positive and negative feedback loops. This network of interlocking feedback loops causes the core clock genes to be rhythmically expressed. Connected to the core clock are additional genes, usually TFs, as the core clock TFs regulate not only themselves but also genes outside the core clock. This transmits the rhythmicity of gene expression generated by the core clock to additional networks, resulting in these specific networks to have rhythmically expressed genes. These genes and the networks they are in are generally referred to as clock-regulated (Fig. 1A). Eventually, the phenotypes connected to these networks display rhythmicity as well. Using these definitions one can being to generate testable hypotheses on how the circadian clock could have been involved in the evolution of CAM.

For example, stomatal movement has been shown to be under the control of the circadian clock [10], therefore, the inversion of stomatal movement seen in CAM plants, relative to C_3, could have occurred from rewiring between the core clock and the gene regulatory network (GRN) controlling stomatal movement. Specifically, the stomatal movement GRN could be under the control of another core clock gene in the core clock (Fig. 1B) or the original core clock gene was rewired within the core clock network (Fig. 1C), altering timing in the stomatal movement GRN.

A more intriguing hypothesis is that CAM plants use different genes than C_3 plants in the core clock network. This isn't to say that the CAM core clock network is constructed differently or consist of functionally different genes, but rather has conserved network topology and functionally similar, non-orthologous genes. This hypothesis is based on the idea that network topology is as equally, if not more important, in GRNs as the network components themselves [11,12]. The conservation of topology and sequence divergence in components in circadian clock network across species has been well documented [9,13,14]. This presents the idea that the stomatal movement GRN in CAM plants could be regulated by unknown core clock genes.

To test these hypotheses, construction of gene network models that incorporate the underlying temporal dynamics is needed. Traditional methods to build models, such as ChiP-chip, ChiP-seq, and mutant expression profiling, can be laborious and can miss the dynamics of the network. Fortunately, high-throughput technologies have allowed for tractable methods of measuring transcription levels in time-course experiments [15-21]. These data exhibit the underlying

temporal dynamics of gene expression and new computational tools have taken advantage of this property to help infer and build gene network models [22,23].

Therefore, we utilized time-course transcriptome data from the CAM plant *Kalanchoë fedtschenkoi* [21] to infer the regulatory relationships between the core clock network and stomatal-related genes. Through network inference, several genes were identified as potentially new core clock genes in *K. fedtschenkoi*. Additionally, stomata-related genes, including genes with rescheduled gene expression, were predicted to be regulated by core clock genes in *K. fedtschenkoi*. The circadian clock plays a crucial role in the physiological response to various environmental stresses in plants, such as drought [24] and our results provide a circadian clock network model to experimentally test various hypotheses on circadian control of stomatal movement in CAM. Fully elucidating the links between the circadian clock and CAM will be key for successful engineering of CAM into C_3 plants for improved drought response and tolerance.

2 Results

2.1 Candidate core clock transcription factors in *Kalanchoë fedtschenkoi*

DLxJTK was used to rank the rhythmic *K. fedtschenkoi* gene list from Moseley, et al. [25] to pull out potential core clock genes. The top 60 TFs ranked by DLxJTK were selected from the full list of DLxJTK ranked genes (Supplementary Table S1) and were used as candidate core clock TFs. The candidate core clock TFs covered a majority of the phases of the day and displayed a bimodal distribution with peaks occurring before subjective night and before subjective morning (Supplementary Fig S1). These results are consistent with phase call distributions of circadian genes seen in other plant species, as well as non-plant species [16,19,26]. To determine if any of the *K. fedtschenkoi* TFs were orthologous to *A. thaliana* TFs that have been annotated as circadian-related, ortholog groups (OGs) constructed in Yang, et al. [21] were investigated. Only 36 of the 60 *K. fedtschenkoi* TFs were placed in OGs with 75 *A. thaliana* genes. Of the 75 *A. thaliana* genes, 13 were found to be associated with circadian rhythm (Supplementary Table S2). A majority of the TFs belonged to the C2H2, MYB-HB, and C2C2-CO families containing 15, 9, and 9 genes, respectively (Supplementary Fig. S1).

After applying a cutoff of 0.7 to remove low probability regulatory relationships, all known core clock TFs were predicted to regulate at least one candidate core clock TF (Fig. 2). One ortholog of LNK2 (Kaladp0099s0129) was predicted to regulate 20 candidate clock-regulated TFs while the orthologs of LUX (Kaladp0033s0047) and LNK1 (Kaladp0607s0046) were predicted to regulate eight different candidate core clock TFs each. All but 6 candidate core clock TFs were found to be activated or repressed by known core clock TFs, while 13 candidate core clock TFs were found to activate or repress known core clock TFs. Only one candidate core clock TF (Kaladp0748s0043) was predicted to regulate more than one known core clock TF (Fig. 2). Using a likelihood ranking cutoff of 0.7, eight candidate core clock TFs were identified as high-confidence core clock TFs in *K. fedtschenkoi* (Fig. 3A and Supplementary Table S3). The eight high-confidence candidate core clock TFs were phased to three separate phases of the day (i.e., morning, midday, evening). To annotate each of the eight high-confidence candidate core clock TFs, *A. thaliana* orthologs were identified by placement in OGs. All *A. thaliana* orthologs identified were rhythmic [25]. Descriptions of the genes are below and separated into three categories corresponding to phase of the day (i.e., morning, midday, evening) of max gene expression as follows.

2.2 Morning phased candidate core clock transcription factors

Three of the high-confidence candidate core clock TFs (Kaladp0011s1342, Kaladp0009s0042, and Kaladp1154s0002) in *K. fedtschenkoi* were phased to the morning (Table 1). Among these three, Kaladp0011s1342 and Kaladp0009s0042 were not placed in an OG with an *A. thaliana* gene. Therefore, their respective protein sequence was used to search the NCBI non-redundant protein BLAST database using an E-value cutoff of 1e-5. Kaladp0011s1342 was phased to 2 h before the beginning of the light period and found to have a similar protein sequence with two *A. thaliana* proteins, AT3G58120 (BZIP61) and AT2G42380 (BZIP34) (Table 1 and Supplementary Fig. S2). BZIP TFs are known to regulate pathogen defense, light and stress signaling, seed maturation and flower development [27]. BZIP34 has been predicted to be involved in the regulation of lipid metabolism and/or cellular transport [28]. BZIP34 and BZIP61 were both rhythmic and were phased to four and eight h after the beginning of the light period, respectively (Table 1 and Supplementary Fig. S2). Kaladp0009s0042's protein sequence lacked homology with any protein sequences in *A. thaliana*. The protein sequence of Kaladp0009s0042 was found to contain a Dof (DNA-binding with one finger) domain. Additionally, the remaining *K. fedtschenkoi* gene in this group, Kaladp1154s0002, was found in an OG containing three *A. thaliana* genes encoding for the Dof domain-containing proteins, including cycling DOF factor 1 (AT5G62430; CDF1), 2 (AT5G39660; CDF2), and 3 (AT3G47500; CDF3). CDF1, CDF2, and CDF3 are involved in various signaling pathways, including photoperiodic and light signaling, stress responses and circadian clock regulation [29]. CDF1 transcription has been reported to be repressed by the circadian clock pseudo-response regulator protein family [30-34] and activated by the circadian clock genes CCA1 and LHY [35], resulting in CDF1 gene expression at dawn. All three *A. thaliana* orthologs were rhythmic and phased to dawn (Table 1 and Supplementary Fig. S2). CDF1 protein accumulation is also regulated by the circadian clock through protein stability via complex formation with gigantean (GI) or flavin-binding, Kelch repeat, F-box 1 (FKF1) [29,36,37]. However, feedback into the clock has not been reported for the CDFs. Both Kaladp0009s0042 and Kaladp1154s0002 gene expression peaked at dawn (Table 1 and Supplementary Fig. S2).

2.3 Midday phased candidate core clock transcription factors

Four of the high-confidence candidate core clock TFs (Kaladp0878s0025, Kaladp0674s0030, Kaladp0748s0043, and Kaladp0674s0182) in *K. fedtschenkoi* were phased to midday (Table 1 and Supplementary Fig. S2). Kaladp0878s0025 had one *A. thaliana* ortholog (AT1G07050), which encodes for a Constans-like protein encoding gene and is a predicted target of the clock regulator GI [38]. The transcript expression of Kaladp0878s0025 was phased to 8 h after the beginning of the light period, whereas its *A. thaliana* ortholog had gene expression phased to 12 h after light (Table 1 and Supplementary Fig. S2).

Kaladp0674s0030 had two *A. thaliana* orthologs, AT5G63160 and AT3G48360, with both encoding for members of the Bric-a-Brac/Tramtrack/Broad (BTB) gene family, specifically BT1 and BT2, respectively. Only BT2 had gene expression data in the Mockler, et al. [39] dataset and was found to be rhythmic with gene expression phased to 20 h after light. Kaladp0674s0030 was phased to six h after light, displaying a shift in expression between the two species (Table 1 and Supplementary Fig. S2). BT2 is known to activate telomerase expression in mature *A. thaliana* leaves, play a critical role in nitrogen-use efficiency in *A. thaliana* and *Oryza sativa*,

suppress sugar and ABA responses, and positively regulate certain auxin responses in plants [40,41]. Additionally, BT2 is regulated diurnally and controlled by the circadian clock, with maximum expression in the dark [41]. It has been suggested that the pattern of gene expression for BT2 mRNA could be linked to the availability of photosynthate, which is a product of photosynthesis [40].

Kaladp0748s0043 was found to be orthologous to the *A. thaliana* plant homeobox family protein BELL1 (BEL1), which is a key regulator of ovule development and needed for auxin and cytokinin signaling pathways for correct patterning of the ovule [42]. Kaladp0748s0043's gene expression was phased to 6 h after light, whereas BEL1 in *A. thaliana* was phased to 12 h after light (Table 1 and Supplementary Fig. S2).

Kaladp0674s0182 was orthologous to RING/U-box superfamily proteins in *A. thaliana*, which are typically E3 ubiquitin ligases. Due to this ambiguity in function, Kaladp0674s0182 will not be used in further analysis.

2.4 Evening phased candidate core clock transcription factors

Only one high-confidence candidate core clock TF (Kaladp0007s0017) in *K. fedtschenkoi* had gene expression phased to the evening (Table 1 and Supplementary Fig. S2). Kaladp0007s0017 had two *A. thaliana* orthologs, which encode for jasmonate (JA)-associated MYC2-like proteins 1 (AT2G46510; JAM1) and 2 (AT1G01260; JAM2). JAM1 has been reported as the balancing component opposite of the MYC2 TF in the JA signaling pathway [43]. Specifically, JAM1 and MYC2 are induced by JA and share many of the same target genes. Where MYC2 activates transcription of multiple genes, including JAM1, JAM1 negatively influences gene expression by physically inferring with MYC2 binding to promoter regions of target genes [43]. The target genes for both TFs are considered "early-responsive JA genes" as changes in gene expression of target genes occur within 1 h of JA detection [44]. JA signaling is linked to activation of defense pathways and has been reported to be under the control of the circadian clock through regulation of MYC2 via repression of transcription [45]. Additionally, JAM1 has been reported to participate in ABA signaling as a positive regulator as overexpression of the gene in *A. thaliana* increased drought tolerance [45]. JAM1 and JAM2 in *A. thaliana* had gene expression phased to 12 and 9 h after light, respectively, whereas Kaladp0007s0017 had gene expression phased to 14 h after light (Table 1 and Supplementary Fig. S2).

2.5 Core clock regulation of stomata-related genes in *K. fedtschenkoi*

The fact that stomatal movement has been inverted in CAM plants raises the question whether the circadian clock, specifically core clock TFs, played a role in this inversion through rescheduling of gene expression. To investigate this question, regulatory relationships were inferred between the 7 high-confidence candidates, plus the known core clock TFs in *K. fedtschenkoi* (Supplementary Table S4), and the 1,605 stomata-related genes in *K. fedtschenkoi*, which were identified as rhythmic in a separate study [25]. Four high-confidence candidate core clock TFs in *K. fedtschenkoi*, Kaladp0007s0017, Kaladp1154s0002, Kaladp0674s0030, and Kaladp0878s0025, were found in the target list (Supplementary Table S1 of Moseley, et al. [8]) of rhythmic stomata-related genes and were subsequently removed as targets.

LEM was used to infer regulatory relationships between core clock TFs and rhythmic stomata-related genes. Using a cutoff of 0.7 on the LEM output related to the probability of a TF regulating a gene, 582 of the 1,605 stomata-related genes were inferred to be regulated by core clock TFs (Supplementary Table S4 and S5). A visualization of the overall network can be seen in Figure 4. Core clock genes are known to activate or repress genes, depending on their mode of regulation, during specific phases of the day. For instance, CCA1 and LHY repress genes that are expressed during the evening [46]. To determine if LEM predicts the appropriate phase of regulation for core clock genes, the phase calls of target genes for core clock TFs with ≥10 target genes were examined in diel plots (Fig. 4 and 5a). LNK1 and LNK2 are thought to activate gene expression of targets during the afternoon and evening [46], and in line with this, a majority of the predicted stomata-related gene targets for both LNK genes in *K. fedtschenkoi* were phased to the afternoon and evening (Fig. 5a). PRR7 is known to repress genes during dawn and in the morning [46]. Indeed, predicted target genes of PRR7 in *K. fedtschenkoi* were phased to dawn and the morning (Fig. 5a). Lastly, ELF4 is known to repress genes in the morning and evening by forming a complex with ELF3 and LUX [47]. In *K. fedtschenkoi* LEM only predicted morning-phased gene targets for ELF4 in *K. fedtschenkoi* (Fig. 5a). The remaining components of the evening complex were examined as well to see if any of their targets were phased to the evening. None of the remaining evening complex components, including the two other ELF4 genes, had target genes phased to the evening (Supplementary Table S6). Additionally, ELF4, ELF3, and LUX did not share similar targets.

Candidate core clock TFs were allowed in the LEM model to be activators or repressors. Most TFs were inferred to be both activators or repressors, so it is unclear whether they are acting as one or the other (Supplementary Table S5). However, high-confidence candidate clock TFs Kaladp0007s0017 and Kaladp9878s0025 were primarily predicted as activators of gene expression and candidate core clock TF Kaladp1154s0002 was predicted primarily as a repressor of gene expression (Supplementary Table S5). A majority of the targets for all high-confidence candidate core clock TFs were phased to either dusk or dawn (Fig. 5b).

To determine what biological functions in stomata-related processes are under the control of the circadian clock, enrichment of associated gene ontology terms was performed. A majority of the biological functions enriched in the 582 rhythmic stomata-related genes were associated to protein phosphorylation (Fig. 6). To determine if any *K. fedtschenkoi* genes are related to *A. thaliana* genes annotated or known as stomata-related, OGs were examined again. Within the 582 *K. fedtschenkoi* genes, 49 were placed in OGs that contained *A. thaliana* genes that were either annotated or known as stomata-related genes (Supplementary Table S5). All the candidate core clock TFs and four known core clock TFs were predicted to regulate at least one of the 49 *K. fedtschenkoi* genes. The remaining stomata-related genes were identified in a separate study [8] as new stomata-related genes and all known and candidate core clock TFs were predicted to regulate at least one new stomata-related gene (Supplementary Table S5).

2.6 Regulation of rescheduled stomata-related genes

Five stomata-related *K. fedtschenkoi* genes identified in Yang, et al. [21] as having undergone rescheduling of gene expression relative to their C_3 orthologs were inferred to be regulated by core clock TFs, with only one of them being orthologous to a known stomata-related gene in *A. thaliana* (Table 2). Twelve stomata-related *K. fedtschenkoi* genes identified in Moseley, et al. [8] as having undergone rescheduling of gene expression relative to their C_3

orthologs were inferred to be clock-regulated, with only three being orthologous to annotated or known stomata-related genes in *A. thaliana* (Table 2). Several of these genes were predicted to encode for protein kinases and transporters. High-confidence candidate core clock TF Kaladp0011s1342 was predicted to regulate the most stomata-related genes that displayed re-scheduling (Table 2). Interestingly, Kaladp0011s1342 also displayed rescheduling of gene expression relative to its two orthologs in *A. thaliana* (Supplementary Fig. S2) and one of the targets of Kaladp0011s1342 was the rescheduled catalase 2 gene identified in Moseley, et al. [8].

3 Discussion

Through a combination of a new metric to identify potential core clock genes and the gene regulation inference algorithm, LEM, this study predicted several novel candidate core clock TFs in *K. fedtschenkoi*. These seven core clock gene candidates predicted in *K. fedtschenkoi* are located in the transcriptional feedback loops (Fig. 3B), which are consistent with the feedback-loop architecture of the core clock [46]. *A. thaliana* orthologs of several *K. fedtschenkoi* candidate core clock TFs were found to have been reported to be connected to the circadian clock and are involved in various signaling pathways. Additionally, most of the candidate core clock TFs were either phased to the morning or evening (Table 1), consistent with other reported circadian genes [16,19]. The discovery of these new potential core clock candidates supports our previous hypothesis that there could be unknown circadian genes in *K. fedtschenkoi* [21]. Recently, LEM was used to identify new core circadian clock genes in mouse [22], in which four out of the top ten genes were validated as clock-regulated genes via RNAi knockdown approach. The success of this application of LEM in mouse makes the candidate core clock TFs identified here in *K. fedtschenkoi* high-confidence candidates for future experimental work.

Figure 1 illustrates two models explaining how the circadian clock could alter gene expression and therefore physiology of an organism. The first model (Fig. 1B) illustrates that a different core clock gene could have taken over the regulation of a physiological process, thus changing when the process occurs. The second model (Fig. 1C) illustrates that a change in the timing of a physiological process could be a result of the core clock gene, that regulates the process, being rewired in the core clock network. Evidence for both models explaining how stomatal movement was inverted in CAM plants by the circadian clock was found, suggesting that both mechanisms could have aided in the evolution of CAM. For instance, ELF4 was found to have rescheduled expression relative to its *A. thaliana* ortholog [25] and in the current study, ELF4 was predicted to regulate a rescheduled stomata-related gene (Table 2), in line with the second model (Fig. 1C).

It has been suggested that there could be new core clock genes in *K. fedtschenkoi* [21] and in this study, several TFs were predicted with high-confidence to be new core clock genes in *K. fedtschenkoi* (Table 1 and Fig. 4). This presented a new model by which the circadian clock could have altered stomatal movement in CAM plants via the new core clock genes regulating stomatal-related genes. Several rescheduled stomatal-related genes were inferred to be regulated by the predicted core clock genes (Table 2), supporting the model of new core clock genes regulating stomatal movement. For example, a duplicated CAT2 gene in *K. fedtschenkoi* was found to have rescheduled gene expression and proposed to be involved in the inversion of stomatal movement in *K. fedtschenkoi* [8]. The rescheduled CAT2 gene was inferred to be regulated by the predicted core clock gene Kaladp0011s1342. Evidence supporting this model and the two models illustrated in Figure 1 suggest that the core clock played a role in reversing

the day/night stomatal movement pattern in CAM photosynthesis species in comparison with C_3 photosynthesis species through a combination of the three mechanisms.

The candidate core clock TFs identified in this study may be used by the core clock network to integrate rhythmicity into the various signaling pathways they control or, they may be diurnally regulated genes which functionally affect the core clock. Given the experimental conditions, we cannot say these are not diurnally regulated genes, and thus they may represent clock regulatory elements that alter the transcript dynamics of known core clock genes by integrating external light stimuli. Experimental work is needed to determine their essentiality with the core clock. For instance, protoplast transient reporter gene expression assay [48] can be used to validate the role of these TFs in the regulation of circadian rhythm. Specially, promoter fusion constructs can be used that contain the promoter of the candidate core clock TF driving the transcription of a fluorescent protein. Transfection of protoplasts and subsequent recording of fluorescence over 48 hours would enable a quick means to determine involvement with the clock. A *K. fedtschenkoi* protoplast protocol has not been published, though an *A. thaliana* protoplast assay may work based on the concept of core clock genes being highly connected. An additional study could examine the impact of the candidate core clock TF on the core clock network. Here, a two-promoter construct would be made with one promoter being a constitutive promoter driving the transcription of the respective candidate TF and the second promoter being the promoter of a known core clock gene that the candidate TF is predicted to target. The known core clock promoter would then drive the transcription of a fluorescent protein. If the candidate core clock TF is integrated into the core clock network, expression of the fluorescent protein will be altered relative to the expression of the core clock gene associated with the core clock promoter driving the fluorescent protein.

Within the clock-regulated stomata-related genes, several GO terms associated with phosphorylation were significantly enriched (Fig. 6). Additionally, several rescheduled stomata-related genes were identified as protein kinases (Table 2). Phosphorylation allows for rapid regulation of protein function and is known to play a significant role in stomatal movement. Furthermore, an extensive array of phosphorylation and dephosphorylation events occur in guard cells [49]. Signaling pathways for stomatal closure, e.g., ABA, and for stomatal opening, e.g., blue light, both involve protein kinases phosphorylating anion channels (ABA signaling: [50-53]) and H^+-ATPases (blue light signaling: [54-56]). Evidence of gene expression rewiring of protein kinases involved in stomatal movement has also been identified in the CAM plant *Agave americana* [15]. Taken together, these results suggest that the clock played a role in the inversion of stomatal movement, potentially by rescheduling phosphorylation events of stomata-related genes. The direct substrates of these kinases and how they affect stomatal movement remains unknown. However, the channels and ATPase identified in Table 2 serve as good candidates to test as substrates for the protein kinases identified here.

Through gene regulatory network analysis, this study predicted a set of novel TFs that could be important components of either the core clock or the networks attached to the core clock in the CAM species *K. fedtschenkoi*. These candidate core clock TFs, if validated by experiments in the future, would significantly advance our understanding of the regulatory mechanism in CAM systems. Also, our analysis of the regulatory relationship between the novel TFs and stomata-related genes revealed that clock-facilitated rescheduling of protein kinases involved in stomatal movement aided the inverted stomatal movement seen in CAM plants, via connecting with different core clock TFs. These results provide new knowledge to inform

genetic improvement of drought resistance in C_3 photosynthesis plants for sustainable food and bioenergy production on dry and marginal lands.

4 Materials and Methods

4.1 Time-course gene expression data

The diel expression data for *K. fedtschenkoi* and *Arabidopsis thaliana* were obtained from Yang, et al. [21] and Mockler, et al. [39], respectively. The *A. thaliana* dataset obtained from Mockler, et al. [39] was the photo/thermocycle dataset. The *K. fedtschenkoi* expression data were collected at 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, and 24 h whereas the *A. thaliana* data were collected at 4, 8, 12, 16, 20, and 24 h after the starting of the light period. Since the *A. thaliana* gene expression data was measured at 4-h intervals and the *K. fedtschenkoi* data was measured at 2-h intervals, the *A. thaliana* data was adjusted to arrive at expression profiles for all *A. thaliana* and *K. fedtschenkoi* genes on the same time scale. Here, the piecewise cubic Hermite interpolating polynomial (pchip) interpolation function in the pandas Python library was used to sample the *A. thaliana* data to simulate gene expression levels at additional time points so that both time-course data sets consisted of the same time intervals: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, and 24 h after the starting of the light period. Pchip was preferred over the more common method of cubic spline interpolation due to cubic spline's tendency to overshoot which introduces oscillations. Additionally, pchip maintains the shape of the data and has been used on microarray time course data sets [57,58]. Additionally, *K. fedtschenkoi* genes with a max FPKM<1 were considered noise and removed. The rhythmic *K. fedtschenkoi* and *A. thaliana* gene sets were taken from [25].

4.2 Identifying candidate core clock genes

Identifying core clock genes using time-course data can be difficult due to similarities in their gene expression profiles with the gene expression profiles of clock-regulated genes. However, core clock genes in various species have been found to have the highest amplitudes and the most statistically significant rhythms [22,59,60]. The periodicity detection methods of de Lichtenberg (DL) [61] and JTK-CYCLE (JTK) [62] are assays that take into account the amplitude of time-course gene expression and if the period of expression matches to the period length in question, respectively. Therefore, to better identify core clock genes, a new metric was established, termed DLxJTK, that combines these two features of DL and JTK (S. Haase and F. Motta, personal communication). DLxJTK uses the p-values for amplitude from DL and for periodicity from JTK for each gene and has been used in mammalian and fungal systems with success (S. Haase and F. Motta, personal communication). The DLxJTK formula is:

$$DLxJTK = P_{per}\,P_{amp}\left(\left(1+\frac{P_{per}}{0.001}\right)^2\right)\left(\left(1+\frac{P_{amp}}{0.001}\right)^2\right)$$

where P_{per} is the JTK p-value for periodicity and P_{amp} is the DL p-value for amplitude. The output of DLxJTK is a ranked ordered list of genes, with core clock genes being near the top of the list. DLxJTK was applied to the *K. fedtschenkoi* rhythmic gene list from Moseley, et al. [25] and the top 60 TFs were used for further analysis.

Previously, a method, Local Edge Machine (LEM), was described that enabled the discovery of new components of the mouse circadian clock network [22]. To identify high-confidence core clock genes in *K. fedtschenkoi*, LEM was used to infer regulatory relationships between the top 60 candidate core clock TFs and TFs orthologous to known core clock genes in *A. thaliana* (from here on referred to as "known core clock TFs") in two steps. Firstly, LEM was used to identify if any of the 60 candidate core clock TFs were regulated by the known core clock TFs. The known core clock transcriptional activators used were reveille 8 (RVE8), RVE6, light-regulated WD 1 (LWD1), LWD2, night light-inducible and clock-regulated 1 (LNK1), and LNK2 [46]. The known core clock transcriptional repressors used were circadian clock associated 1 (CCA1), late elongated hypocotyl (LHY), timing of cab expression 1 (TOC1), CCA1 hiking expedition (CHE), LUX, NOX, pseudo-response regulator 9 (PRR9), PRR7, PRR5, ELF3, and ELF4 [46]. LUX is only active after forming the evening complex with early flowering 3 (ELF3) and 4 (ELF4) [47,63], which do not bind to DNA [64]. ELF3 and ELF4 were included for this reason. LEM was set to only use the respective mode of gene regulation for each known core clock TF used. All candidate core clock TFs were set as targets for the known core clock TFs.

Secondly, since core clock TFs are known to regulate other known core clock genes, LEM was run again but with the top 60 candidate core clock TFs as potential regulators of the known core clock TFs. To identify high-confidence candidate core clock TFs, a measure of likelihood, described in McGoff, et al. [22], was used for each candidate core clock TF. This measure is calculated by taking the maximum LEM probability that the candidate core clock TF was a regulator of any known core clock TF and multiplying it by the maximum LEM probability that the candidate core clock TF was regulated by any known core clock TF.

4.3 Identify core clock-regulators of stomata-related genes

LEM was applied to identify potential regulatory relationships between core clock TFs and stomata-related genes. Known core clock TFs plus the candidate core clock TFs were used as potential regulators of stomata-related genes. Stomata-related genes in *K. fedtschenkoi* were identified as genes that are orthologous to an *A. thaliana* gene that is either annotated as stomata-related or is known as stomata-related. Orthology between species was based on placement within the same ortholog group. Additionally, new *K. fedtschenkoi* stomata-related genes predicted in Moseley, et al. [8] were included as well.

4.4 Gene Ontology analysis

Gene Ontology (GO) terms for the *K. fedtschenkoi* and *A. thaliana* were obtained from Phytozome v12.1 [65]. *K. fedtschenkoi* genes encoding putative transcription factors were retrieved from [21]. Using ClueGO [66], observed GO biological process were subjected to the right-sided hypergeometric enrichment test at medium network specificity selection and p-value correction was performed using the Holm-Bonferroni step-down method [67]. There was a minimum of 3 and a maximum of 8 selected GO tree levels, while each cluster was set to include a minimum of between 3% and 4% of genes associated with each term. GO term fusion and grouping settings were selected to minimize GO term redundancy and the term enriched at the highest level of significance was used as the representative term for each functional cluster. The GO terms with p-values less than or equal to 0.05 were considered significantly enriched.

4.5 Comparative analysis of gene expression

To calculate time-delay between time-course gene expression profiles, the diel expression data were normalized by Z-score transformation. Pair-wise circular cross correlation was calculated for the orthologous gene pairs of interest for all possible time delays using the SciPy library (https://www.scipy.org/) in Python. Circular cross correlation produces a correlation coefficient between two genes (e.g., gene 1 and gene 2) as a function of the lag. With each correlation coefficient, a lag value was given. The lag values were then converted into hours, giving an estimate on time delay. The time delay at which the correlation was maximum was selected as the estimated delay between the two genes. Spearman's rank correlation coefficient was then calculated between gene 1's expression data and the shifted expression data of gene 2 by its estimated time delay.

4.6 Data Availability

All datasets generated for this study are included in the manuscript and the supplementary files.

5 Author Contributions

RCM and XY conceived the research. RCM performed all data analyses and wrote the manuscript. FM, SH, GAT and XY provide input during the study and edited the manuscript.

6 Competing Interest

The author(s) declare no competing interests.

7 Funding

This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. This research was supported by the U.S. Department of Energy, Office of Science, Genomic Science Program under Award Number DE-SC0008834 and the DOE Center for Bioenergy Innovation. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract Number DE-AC05-00OR22725.

8 Acknowledgements

The authors wish to thank David Weston, Margaret E. Staton, and Jay Chen for critical comments on the manuscript.

9 References

1 Ehleringer, J. R. & Monson, R. K. Evolutionary and Ecological Aspects of Photosynthetic Pathway Variation. *Annu. Rev. Ecol. Syst.* **24** (1993).

2 Cook, B., Smerdon, J. E., Seager, R. & Coats, S. Global warming and 21st century drying. *Climate Dynamics* **43**, 2607-2627 (2014).

3 Dai, A. Increasing drought under global warming in observations and models. *Nature Climate Change* **3**, 52-58, doi:10.1038/nclimate1633 (2012).

4 Borland, A. M. *et al.* Engineering crassulacean acid metabolism to improve water-use efficiency. *Trends Plant Sci* **19**, 327-338, doi:10.1016/j.tplants.2014.01.006 (2014).

5 Yang, X. *et al.* A roadmap for research on crassulacean acid metabolism (CAM) to enhance sustainable food and bioenergy production in a hotter, drier world. *New Phytologist* **207**, 491-504 (2015).

6 Hartwell, J. The co-ordination of central plant metabolism by the circadian clock. *Mechanistic and Functional Studies of Proteins* **33**, 945-948 (2005).

7 Hartwell, J. The circadian clock in CAM plants. *Endogenous Plant Rhythms*, 211-236 (2006).

8 Moseley, R. C., Tuskan, G. A. & Yang, X. Comparative Genomics Analysis Provides New Insight Into Molecular Basis of Stomatal Movement in Kalanchoë fedtschenkoi. *Frontiers in Plant Science* **10**, doi:10.3389/fpls.2019.00292 (2019).

9 Harmer, S. L. The circadian system in higher plants. *Annu Rev Plant Biol* **60**, 357-377, doi:10.1146/annurev.arplant.043008.092054 (2009).

10 Hubbard, K. E. & Webb, A. A. in *Rhythms in Plants: Dynamics Responses in a Dynamic Environment* (eds S. Mancuso & S. Shabala) 231-255 (Springer, 2016).

11 Kholodenko, B. Cell Signalling Dynamics in Time and Space. *Nat Rev Mol Cell Biol* **7**, 165-176 (2006).

12 Tiana, G., Krishna, S., Pigolotti, S., Jensen, M. H. & Sneppen, K. Oscillations and temporal signalling in cells. *Phys Biol* **4**, R1-17, doi:10.1088/1478-3975/4/2/R01 (2007).

13 McClung, C. R. Plant Circadian Rhythms. *The Plant Cell* **18**, 792-803 (2006).

14 Young, M. & Kay, S. A. Time Zones: A Comparative Genetics of Circadian Clocks. *Nature Reviews: Genetics* **2**, 702-715 (2001).

15 Abraham, P. E. *et al.* Transcript, protein and metabolite temporal dynamics in the CAM plant Agave. *Nat Plants* **2**, 16178, doi:10.1038/nplants.2016.178 (2016).

16 Filichkin, S. A. *et al.* Global profiling of rice and poplar transcriptomes highlights key conserved circadian-controlled pathways and cis-regulatory modules. *PLoS One* **6**, e16907, doi:10.1371/journal.pone.0016907 (2011).

17 Higashi, T. *et al.* Detection of Diurnal Variation of Tomato Transcriptome through the Molecular Timetable Method in a Sunlight-Type Plant Factory. *Front Plant Sci* **7**, 87, doi:10.3389/fpls.2016.00087 (2016).

18 Koda, S. *et al.* Diurnal Transcriptome and Gene Network Represented through Sparse Modeling in Brachypodium distachyon. *Front Plant Sci* **8**, 2055, doi:10.3389/fpls.2017.02055 (2017).

19 Michael, T. P. *et al.* Network discovery pipeline elucidates conserved time-of-day-specific cis-regulatory modules. *PLoS Genet* **4**, e14, doi:10.1371/journal.pgen.0040014 (2008).

20 Ming, R. *et al.* The pineapple genome and the evolution of CAM photosynthesis. *Nat Genet* **47**, 1435-1442, doi:10.1038/ng.3435 (2015).

21 Yang, X. *et al.* The Kalanchoë genome provides insights into convergent evolution and building blocks of crassulacean acid metabolism. *Nature Communications* **8**, 1-15, doi:10.1038/s41467-017-01491-7 (2017).

22 McGoff, K. A. *et al.* The Local Edge Machine: inference of dynamic models of gene regulation. *Genome Biol* **17**, 214, doi:10.1186/s13059-016-1076-z (2016).

23 Cummins, B., Gedeon, T., Harker, S. & Mischaikow, K. in *Computational Methods in Systems Biology*. (eds J. Feret & H. Koepple) 300-308 (Springer).

24 Grundy, J., Stoker, C. & Carre, I. A. Circadian regulation of abiotic stress tolerance in plants. *Front Plant Sci* **6**, 648, doi:10.3389/fpls.2015.00648 (2015).

25 Moseley, R. C. *et al.* Conservation and Diversification of Circadian Rhythmicity Between a Model Crassulacean Acid Metabolism Plant Kalanchoe fedtschenkoi and a Model C3 Photosynthesis Plant Arabidopsis thaliana. *Front Plant Sci* **9**, 1757, doi:10.3389/fpls.2018.01757 (2018).

26 Zhang, R., Lahens, N., Ballance, H., Hughes, M. E. & Hogenesch, J. B. A circadian gene expression atlas in mammals: Implications for biology and medicine. *PNAS* **111**, 16219-16224 (2014).

27 Jakoby, M. *et al.* bZIP transcription factors in Arabidopsis. *Trends in Plant Science* **7**, 106-111 (2001).

28 Gibalova, A. *et al.* AtbZIP34 is required for Arabidopsis pollen wall patterning and the control of several metabolic pathways in developing pollen. *Plant Molecular Biology* **70**, 581-601 (2009).

29 Fornara, F. *et al.* Arabidopsis DOF transcription factors act redundantly to reduce CONSTANS expression and are essential for a photoperiodic flowering response. *Developmental Cell* **17**, 75-86 (2009).

30 Huang, W. *et al.* Mapping the core of the Arabidopsis circadian clock defines the network structure of the oscillator. *Science* **336**, 75-79 (2012).

31 Ito, S. *et al.* Insight into missing genetic links between two evening-expressed pseudo-response regulator genes TOC1 and PRR5 in the circadian clock-controlled circuitry in Arabidopsis thaliana. *Plant Cell Physiol* **49**, 201-213 (2008).

32 Nakamichi, N. *et al.* Transcriptional repressor PRR5 directly regulates clock-output pathways. *PNAS* **109**, 17123-17128 (2012).

33 Nakamichi, N. *et al.* Arabidopsis clock-associated pseudo-response regulators PRR9, PRR7 and PRR5 coordinately and positively regulate flowering time through the canonical CONSTANS-dependent photoperiodic pathway. *Plant Cell Physiology* **48**, 822-832 (2007).

34 Niwa, T. *et al.* Genetic linkages of the circadian-clock associated genes, TOC1, CCA1 and LHY, in the photoperiodic control of flowering time in Arabidopsis thaliana. *Plant Cell Physiol* **48**, 925-937 (2007).

35 Seaton, D. D. *et al.* Linked circadian outputs control elongation growth and flowering in response to photoperiod and temperature. *Mol Syst Biol* **11**, 776, doi:10.15252/msb.20145766 (2015).

36 Imaizumi, T., Schultz, T., Harmon, F., Ho, L. & Kay, S. FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis. *Science* **309**, 293-297 (2005).

37 Song, Y. H., Smith, R. W., To, B. J., Millar, A. J. & Imaizumi, T. FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. *Science* **336**, 1045-1049, doi:10.1126/science.1219644 (2012).

38 Keurentjes, J. *et al.* Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. *PNAS* **104**, 1708-1713 (2007).

39 Mockler, T. C. *et al.* The Diurnal Project: Diurnal and Circadian Expression Profiling, Model-based Pattern Matching, and Promoter Analysis. *Cold Spring Harbor Symposia on Quantitative Biology* **72**, 353-363 (2007).

40 Araus, V. *et al.* Members of BTB Gene Family of Scaffold Proteins Suppress Nitrate Uptake and Nitrogen Use Efficiency. *Plant Physiol* **171**, 1523-1532, doi:10.1104/pp.15.01731 (2016).

41 Mandadi, K. K., Misra, A., Ren, S. & McKnight, T. D. BT2, a BTB protein, mediates multiple responses to nutrients, stresses, and hormones in Arabidopsis. *Plant Physiol* **150**, 1930-1939, doi:10.1104/pp.109.139220 (2009).

42 Bencivenga, S., Simonini, S., Benkova, E. & Colombo, L. The transcription factors BEL1 and SPL are required for cytokinin and auxin signaling during ovule development in Arabidopsis. *Plant Cell* **24**, 2886-2897, doi:10.1105/tpc.112.100164 (2012).

43 Fonseca, S. *et al.* bHLH003, bHLH013 and bHLH017 are new targets of JAZ repressors negatively regulating JA responses. *PLoS One* **9**, e86182, doi:10.1371/journal.pone.0086182 (2014).

44 Nakata, M. *et al.* A bHLH-type transcription factor, ABA-INDUCIBLE BHLH-TYPE TRANSCRIPTION FACTOR/JA-ASSOCIATED MYC2-LIKE1, acts as a repressor to negatively regulate jasmonate signaling in arabidopsis. *Plant Cell* **25**, 1641-1656, doi:10.1105/tpc.113.111112 (2013).

45 Li, H., Sun, J., Xu, Y., Jiang, H. & Li, C. The bHLH-type transcription factor AtAIB positively regulates ABA response in Arabidopsis. *Plant Molecular Biology* **65**, 655-665 (2007).

46 Nohales, M. A. & Kay, S. A. Molecular mechanisms at the core of the plant circadian oscillator. *Nat Struct Mol Biol* **23**, 1061-1069, doi:10.1038/nsmb.3327 (2016).

47 Nusinow, D. A. *et al.* The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth. *Nature* **475**, 398-402, doi:10.1038/nature10182 (2011).

48 Hansen, L. L. & van Ooijen, G. Rapid Analysis of Circadian Phenotypes in Arabidopsis Protoplasts Transfected with a Luminescent Clock Reporter. *J Vis Exp*, doi:10.3791/54586 (2016).

49 Wang, R. *et al.* Common and unique elements of the ABA-regulated transcriptome of Arabidopsis guard cells. *BMC Genomics* **12** (2011).

50 Grondin, A. *et al.* Aquaporins Contribute to ABA-Triggered Stomatal Closure through OST1-Mediated Phosphorylation. *Plant Cell* **27**, 1945-1954, doi:10.1105/tpc.15.00421 (2015).

51 Mustilli, A. C. Arabidopsis OST1 Protein Kinase Mediates the Regulation of Stomatal Aperture by Abscisic Acid and Acts Upstream of Reactive Oxygen Species Production. *The Plant Cell Online* **14**, 3089-3099, doi:10.1105/tpc.007906 (2002).

52 Sirichandra, C. *et al.* Phosphorylation of the Arabidopsis AtrbohF NADPH oxidase by OST1 protein kinase. *FEBS Lett* **583**, 2982-2986, doi:10.1016/j.febslet.2009.08.033 (2009).

53 Vahisalu, T. *et al.* SLAC1 is required for plant guard cell S-type anion channel function in stomatal signalling. *Nature* **452**, 487-491, doi:10.1038/nature06608 (2008).

54 Kinoshita, T. & Shimazaki, K. Blue light activates the plasma membrane H+-ATPase by phosphorylation of the C-terminus in stomatal guard cells. *The EMBO Journal* **18**, 5548-5558 (1999).

55 Inoue, S. *et al.* Blue light-induced autophosphorylation of phototropin is a primary step for signaling. *PNAS* **105**, 5626-5631 (2008).

56 Wang, Y. *et al.* Overexpression of plasma membrane H+-ATPase in guard cells promotes light-induced stomatal opening and enhances plant growth. *Proc Natl Acad Sci U S A* **111**, 533-538, doi:10.1073/pnas.1305438111 (2014).

57 Dong, B. *et al.* Predicting Housekeeping Genes Based on Fourier Analysis. *PloS ONE* **6**, 1-11, doi:10.1371/

10.1371/journal.pone.0021012.g001 (2011).

58 Luo, Z., Azencott, R. & Zhao, Y. Modeling miRNA-mRNA interactions: fitting chemical kinetics equations to microarray data. *BMC Systems Biology* **9**, 1-15 (2014).

59 Anafi, R. C. *et al.* Machine learning helps identify CHRONO as a circadian clock component. *PLoS Biol* **12**, e1001840, doi:10.1371/journal.pbio.1001840 (2014).

60 Sharma, A., Wai, C. M., Ming, R. & Yu, Q. Diurnal Cycling Transcription Factors of Pineapple Revealed by Genome-Wide Annotation and Global Transcriptomic Analysis. *Genome Biol Evol* **9**, 2170-2190, doi:10.1093/gbe/evx161 (2017).

61 de Lichtenberg, U. *et al.* Comparison of computational methods for the identification of cell cycle-regulated genes. *Bioinformatics* **21**, 1164-1171, doi:10.1093/bioinformatics/bti093 (2005).

62 Hughes, M. E., Hogenesch, J. B. & Kornacker, K. JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J Biol Rhythms* **25**, 372-380, doi:10.1177/0748730410379711 (2010).

63 Mizuno, T. *et al.* Ambient temperature signal feeds into the circadian clock transcriptional circuitry through the EC night-time repressor in Arabidopsis thaliana. *Plant Cell Physiol* **55**, 958-976, doi:10.1093/pcp/pcu030 (2014).

64 Helfer, A. *et al.* LUX ARRHYTHMO encodes a nighttime repressor of circadian gene expression in the Arabidopsis core clock. *Curr Biol* **21**, 126-133, doi:10.1016/j.cub.2010.12.021 (2011).

65 Goodstein, D. M. *et al.* Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res* **40**, D1178-1186, doi:10.1093/nar/gkr944 (2012).

66 Bindea, G. *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091-1093, doi:10.1093/bioinformatics/btp101 (2009).

67 Holm, S. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics* **6**, 65-70 (1979).

Table 1. Candidate core clock transcription factors in *Kalanchoë fedtschenkoi*.

"Kafe shift" is the number of h the *K. fedtschenkoi* gene expression profile shifted from its *A. thaliana* ortholog's gene expression profile. The shift was calculated by subtracting the phases calls of each ortholog. "Spear shift" is the Spearman rank correlation coefficient between orthologs after shifting the *K. fedtschenkoi* gene expression profile by the "Kafe shift".

Kafe Gene ID	Kafe Phase Call	Arth Ortholog ID	Arth Gene Symbol	Arth Gene Desc.	Arth Phase Call	Kafe Shift (hrs)	Spear Shift
Kaladp0748s0043	6	AT5G41410	BEL1	POX (plant homeobox) family protein	12	-6	0.97
Kaladp0007s0017	14	AT2G46510	JAM1	ABA-inducible BHLH-type transcription factor	12	+2	0.91
Kaladp0007s0017	14	AT1G01260	JAM2	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	8	+6	0.83
Kaladp0011s1342	22	AT2G42380	BZIP34	Basic-leucine zipper (bZIP) transcription factor family protein	4	-8	0.96
Kaladp0011s1342	22	AT3G58120	BZIP61	Basic-leucine zipper (bZIP) transcription factor family protein	8	-8	0.97
Kaladp1154s0002	22	AT3G47500	CDF3	cycling DOF factor 3	2	-4	0.97
Kaladp1154s0002	22	AT5G62430	CDF1	cycling DOF factor 1	24	-2	0.97
Kaladp1154s0002	22	AT5G39660	CDF2	cycling DOF factor 2	24	-2	0.87
Kaladp0674s0030	6	AT5G63160	BT1	BTB and TAZ domain protein 1			
Kaladp0674s0030	6	AT3G48360	BT2	BTB and TAZ domain protein 2	20	+10	0.94
Kaladp0878s0025	8	AT1G07050		CCT motif family protein	12	-2	0.92
Kaladp0009s0042	22						

Table 2. Clock-controlled stomata-related genes with rescheduled gene expression from different studies.

"Relationship" refers to the type of transcriptional regulation. "Stomata-related" refers to if the gene is either known as a stomata-related gene via publication in the literature (known), is annotated as a stomata-related (GO), or was identified in Moseley et al. 2018b as being stomata-related (New).

Clock TF	Relationship	Study	Target	Stomata-related	Gene Description
Kaladp0748s0043	Represses	Yang et al. 2017	Kaladp0059s0048	New	aquaporin pip1-2
Kaladp0748s0043	Represses	Moseley et al. 2018b	Kaladp0062s0167	New	receptor-like protein kinase haiku2
Kaladp0007s0017	Activates	Yang et al. 2017	Kaladp0011s0363	New	trehalose-phosphate synthase
Kaladp0007s0017	Activates	Moseley et al. 2018b	Kaladp0092s0084	Known/GO	calcium-dependent protein kinase 26
Kaladp0011s1342	Represses	Yang et al. 2017	Kaladp0040s0264	New	btb poz domain-containing protein npy2-like
Kaladp0011s1342	Represses	Yang et al. 2017	Kaladp0008s0539	New	mitogen-activated protein kinase
Kaladp0011s1342	Represses	Yang et al. 2017	Kaladp0033s0113	Known/GO	phototropin-2
Kaladp0011s1342	Represses	Moseley et al. 2018b	Kaladp0001s0016	New	catalase isozyme 1
Kaladp0011s1342	Represses	Moseley et al. 2018b	Kaladp0093s0030	New	[36] homeobox-leucine zipper protein anthocyaninless 2 isoform x1
Kaladp0059s0037 (ELF4)	Represses	Moseley et al. 2018b	Kaladp0062s0076	New	3-ketoacyl- synthase 19-like
Kaladp0047s0123 (LNK1)	Activates	Moseley et al. 2018b	Kaladp0008s0414	New	cyclic nucleotide-gated ion channel 15
Kaladp0060s0264 (LNK2)	Activates	Moseley et al. 2018b	Kaladp0024s0371	New	pectin lyase-like superfamily protein isoform 1
Kaladp0099s0129 (LNK2)	Activates	Moseley et al. 2018b	Kaladp0042s0353	Known	abscisic acid receptor pyl8-like
Kaladp0060s0264 (LNK2)	Activates	Moseley et al. 2018b	Kaladp0095s0634	GO	mitogen-activated protein kinase homolog mmk2
Kaladp0101s0041 (PRR7)	Represses	Moseley et al. 2018b	Kaladp0092s0115	New	pleiotropic drug resistance protein 1-like
Kaladp0055s0349 (RVE6)	Activates	Moseley et al. 2018b	Kaladp0043s0103	New	phospholipid-transporting atpase 3
Kaladp0055s0349 (RVE6)	Activates	Moseley et al. 2018b	Kaladp0090s0003	New	receptor-like protein kinase

10 Figure Legends

Figure 1. Models for circadian clock-related alterations of physiology at the molecular level.

(A) The core clock generates waves of transcription that propagate into connected (clock-regulated) gene regulatory networks (GRNs) **(B)** GRNs can be rewired to different core clock genes **(C)** Core clock genes can be rewired within the core clock network, carrying the GRNs they are connected to with them. Red circles are core clock genes. Green circles are clock-regulated genes. White and black bars indicate daytime (12-hour) and nighttime (12-hour), respectively.

Figure 2. Several candidate core clock transcription factors are predicted to regulate and be regulated by known core clock transcription factors.

Orange ovals represent candidate core clock transcription factors predicted to regulate and be regulated by known core clock transcription factors. Purple ovals represent known clock transcription factors. Blue ovals represent candidate core clock transcription factors that are predicted to only be regulated by known core clock transcription factors. Edges with a green arrow represent activation of gene expression. Edges with red lines represent repression of gene expression.

Figure 3. Seven candidate core clock transcription factors are predicted to play a role in the core clock network of *Kalanchoë fedtschenkoi*

(A) Predicted regulatory relationships between candidate core clock transcription factors and known core clock genes via the Local Edge Machine. Blue line is the probability that any known core clock transcription factor regulates a candidate clock transcription factor. Orange line is the probability that a candidate core clock transcription factor regulates any known core clock transcription factor. Grey circle is the probability of a regulatory relationship. **(B)** The seven candidate core clock transcription factors and their relationship with the core clock network in *K. fedtschenkoi*. Blue ovals represent core clock genes and red ovals represent candidate clock transcription factors. The numbers arrayed at the bottom of the image indicate the number of hours passed after first light. Green dotted lines represent activation of transcription and black dotted lines represent repression of transcription. White and black bars indicate daytime (12-hour) and nighttime (12-hour), respectively. EC: Evening complex, CTF1: Kaladp0748s0043, CTF2: Kaladp0007s0017, CTF3: Kaladp0011s1342,CTF4: Kaladp1154s0002, CTF5: Kaladp0674s0030, CTF6: Kaladp0878s0025, CTF7: Kaladp0009s0042.

Figure 4. *Kalanchoë fedtschenkoi* core clock transcription factors and stomata-related genes regulatory network.

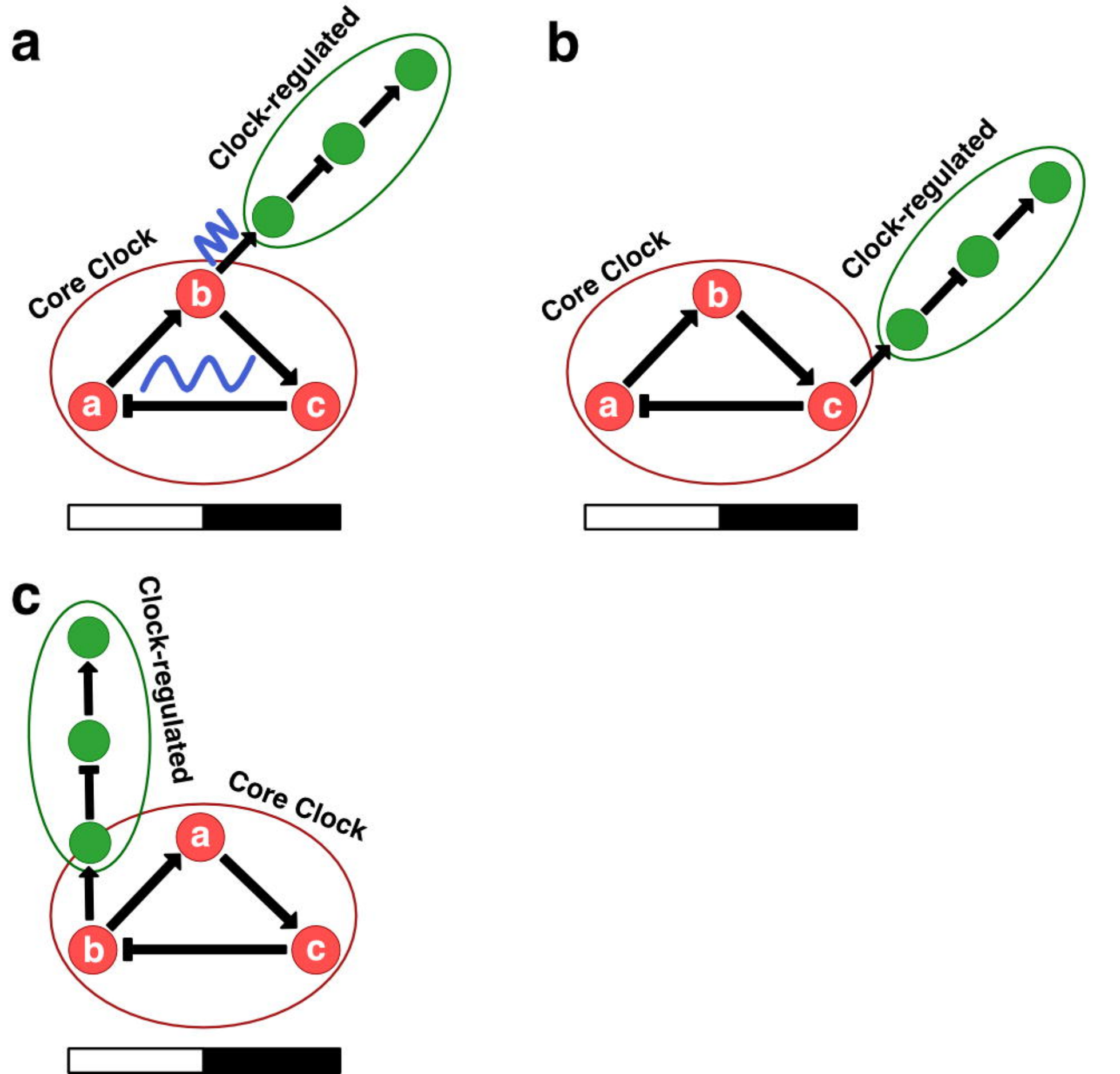
672 Orange ovals represent candidate core clock transcription factors predicted to regulate and be
673 regulated by known core clock transcription factors. Purple ovals represent known clock
674 transcription factors. Blue ovals represent stomata-related genes. Edges with a green arrow
675 represent activation of gene expression. Edges with red lines represent repression of gene
676 expression.

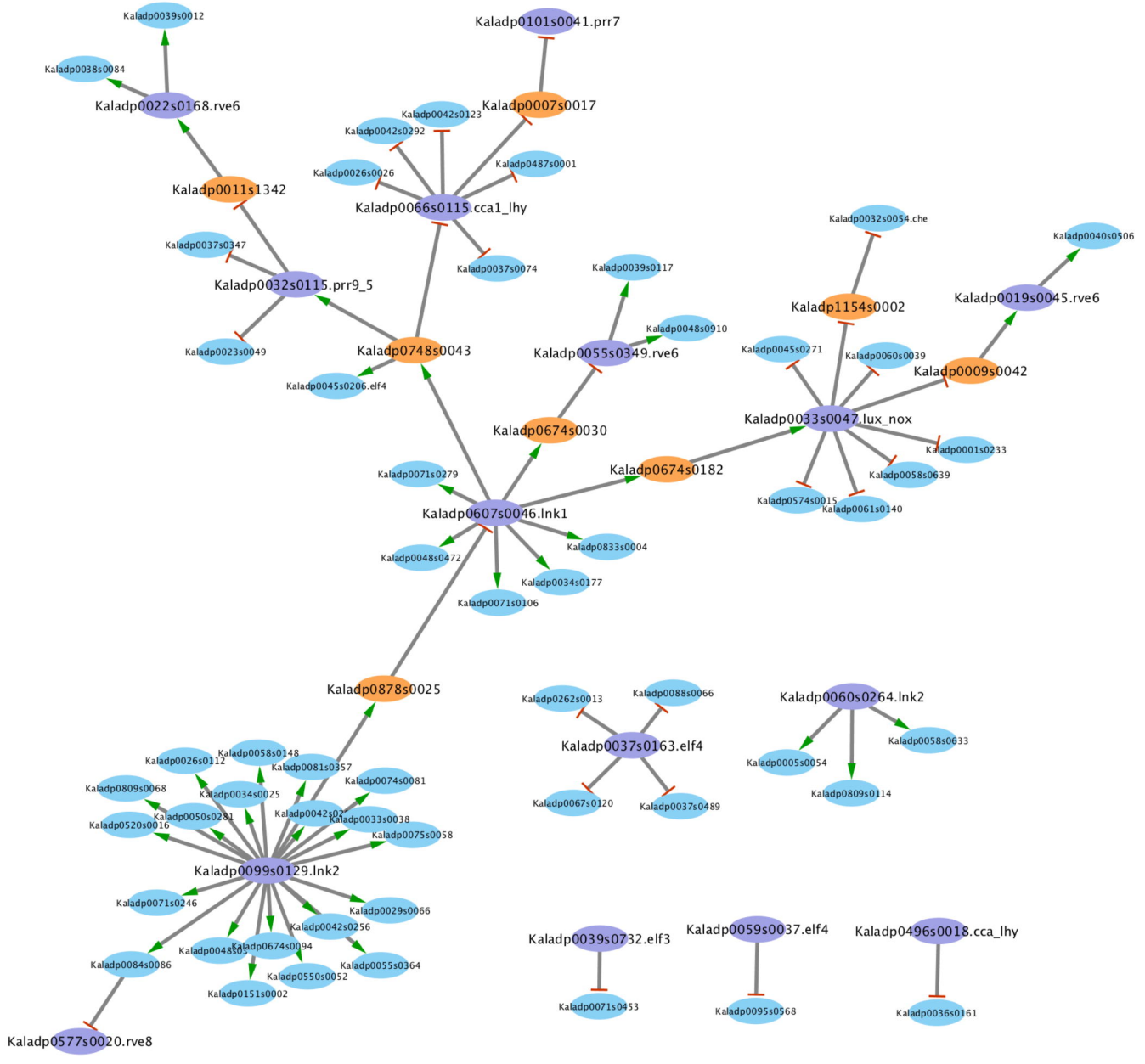
677 **Figure 5. Gene expression profiles of *Kalanchoë fedtschenkoi* known and candidate core**
678 **clock transcription factors and phase calls of their respective targets.**

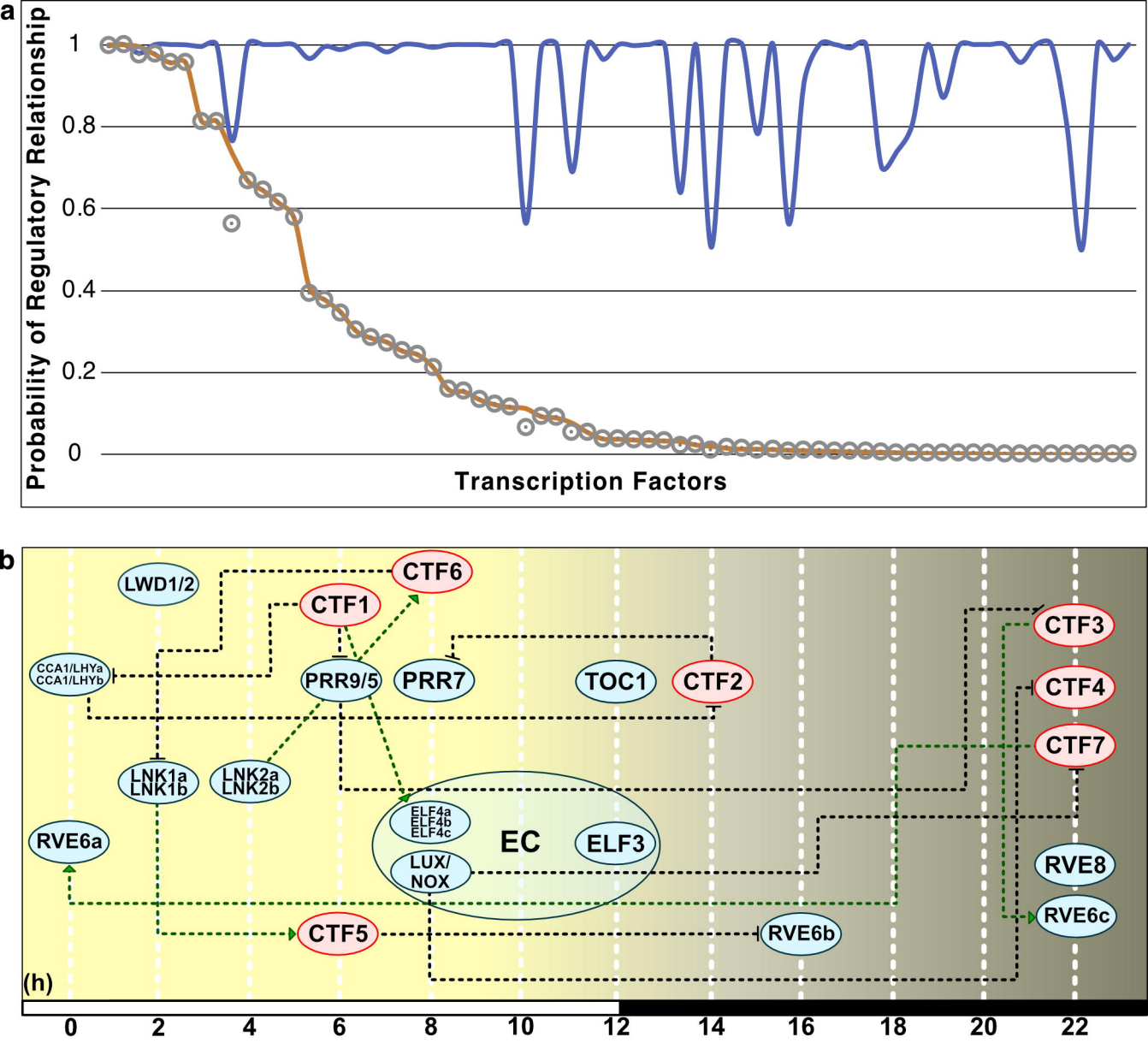
679 **a)** *Kalanchoë fedtschenkoi* known core clock genes. **b)** *Kalanchoë fedtschenkoi* candidate core
680 clock genes. The black lines represent the z-score standardized expression profile of the
681 respective gene. Red and green bars represent the number of predicted target genes phases to the
682 same time of day. Red bars signify that the target genes are repressed by the respective regulator
683 and green bars signify that the target genes are activated by the respective regulator. White and
684 black bars indicate daytime (12-hour) and nighttime (12-hour), respectively.

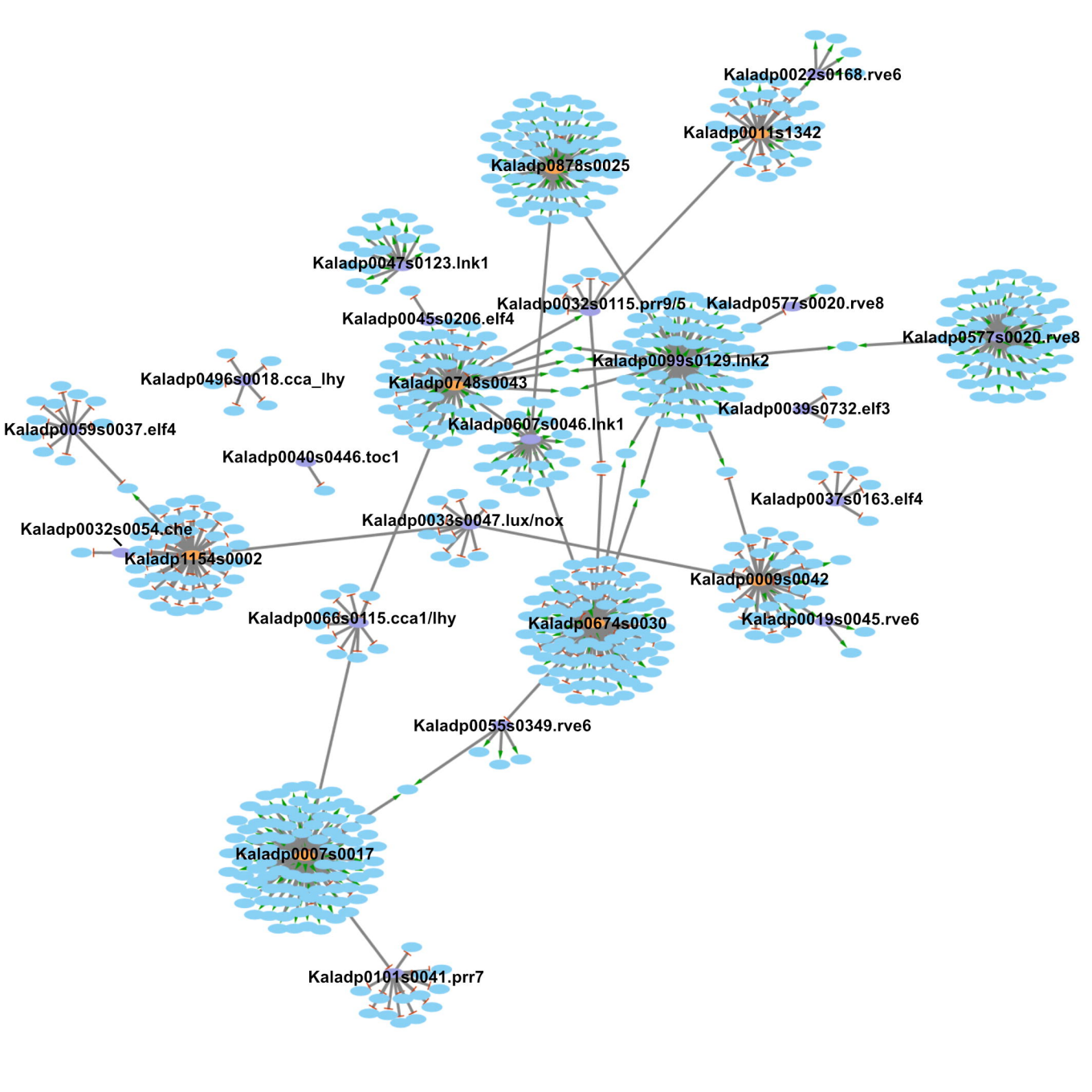
685 **Figure 6. Enriched functional groups within of clock-regulated stomata-related genes in**
686 ***Kalanchoë fedtschenkoi*.**

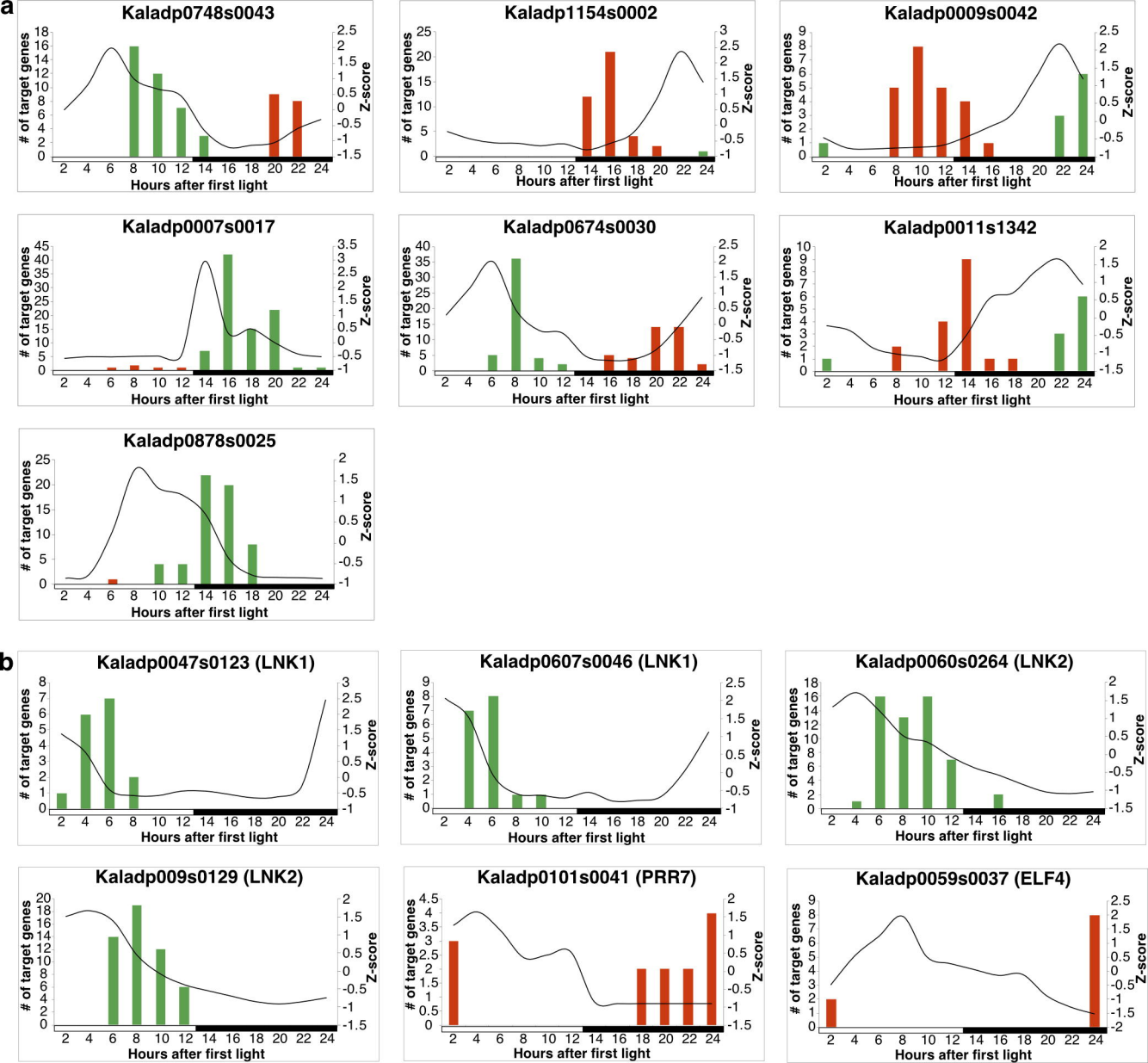
687 Enriched functional groups were determined in the Cytoscape application. ClueGO. Parameters
688 used in ClueGO are described in the Materials and Methods section.

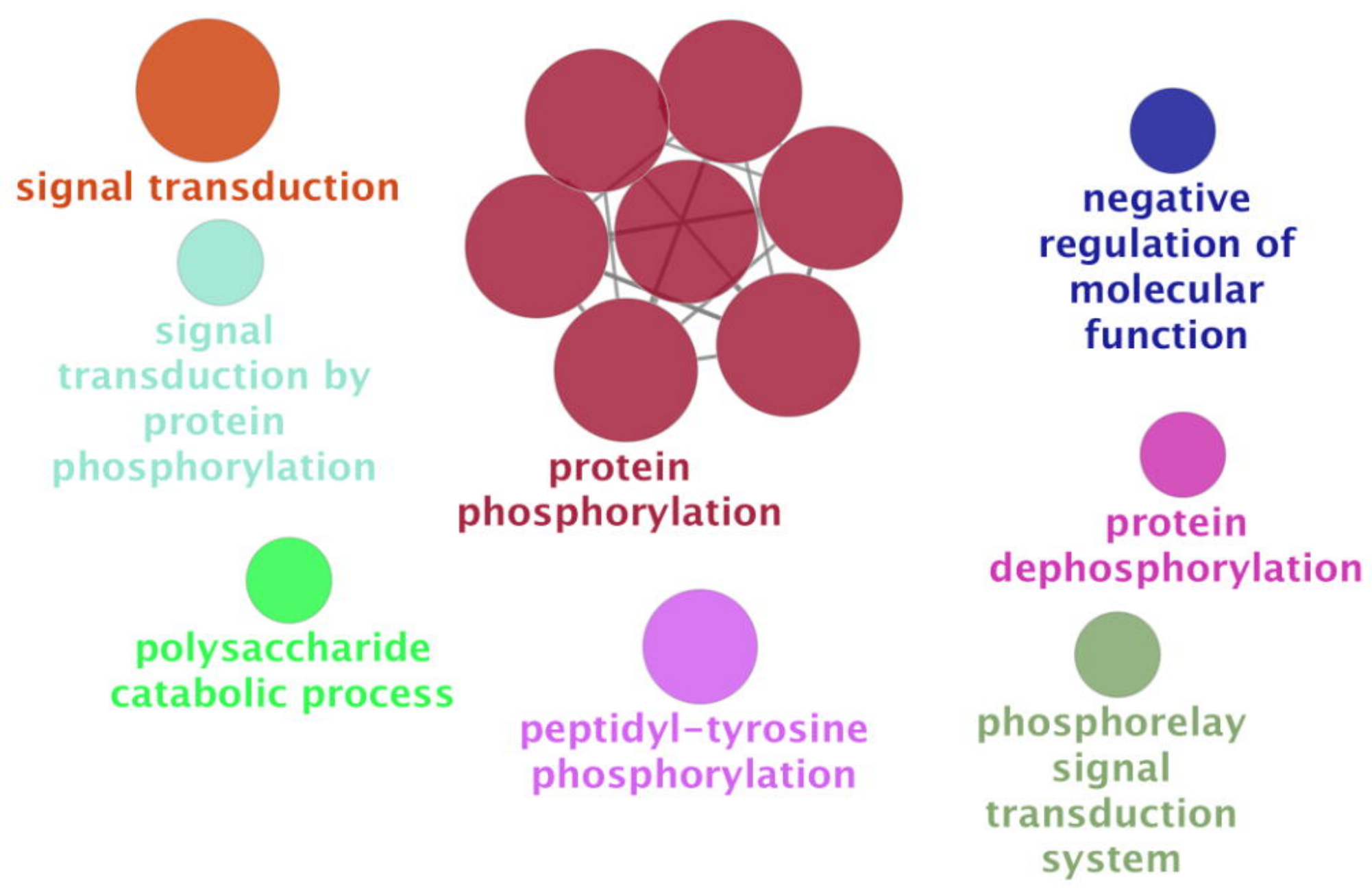

a
Clock-regulated
Core Clock
b
a
c
b
Clock-regulated
Core Clock
b
a
c
c
Clock-regulated
Core Clock
a
b
c


Kaladp0039s0012
Kaladp0101s0041.prr7
Kaladp0038s0084
Kaladp0022s0168.rve6
Kaladp0007s0017
Kaladp0042s0123
Kaladp0042s0292
Kaladp0026s0026
Kaladp0487s0001
Kaladp0011s1342
Kaladp0066s0115.cca1_lhy
Kaladp0037s0347
Kaladp0037s0074
Kaladp0032s0115.prr9_5
Kaladp0039s0117
Kaladp0032s0054.che
Kaladp0040s0506
Kaladp1154s0002
Kaladp0019s0045.rve6
Kaladp0048s0910
Kaladp0055s0349.rve6
Kaladp0023s0049
Kaladp0748s0043
Kaladp0045s0271
Kaladp0060s0039
Kaladp0009s0042
Kaladp0045s0206.elf4
Kaladp0033s0047.lux_nox
Kaladp0674s0030
Kaladp0001s0233
Kaladp0674s0182
Kaladp0058s0639
Kaladp0071s0279
Kaladp0574s0015
Kaladp0061s0140
Kaladp0607s0046.lnk1
Kaladp0833s0004
Kaladp0048s0472
Kaladp0034s0177
Kaladp0071s0106
Kaladp0878s0025
Kaladp0262s0013
Kaladp0088s0066
Kaladp0060s0264.lnk2
Kaladp0037s0163.elf4
Kaladp0058s0633
Kaladp0005s0054
Kaladp0809s0114
Kaladp0067s0120
Kaladp0037s0489
Kaladp0058s0148
Kaladp0026s0112
Kaladp0081s0357
Kaladp0809s0068
Kaladp0074s0081
Kaladp0034s0025
Kaladp0050s0281
Kaladp0042s02
Kaladp0033s0038
Kaladp0520s0016
Kaladp0075s0058
Kaladp0099s0129.lnk2
Kaladp0071s0246
Kaladp0029s0066
Kaladp0042s0256
Kaladp0048s03
Kaladp0674s0094
Kaladp0084s0086
Kaladp0055s0364
Kaladp0550s0052
Kaladp0151s0002
Kaladp0577s0020.rve8
Kaladp0039s0732.elf3
Kaladp0059s0037.elf4
Kaladp0496s0018.cca_lhy
Kaladp0071s0453
Kaladp0095s0568
Kaladp0036s0161


a
Probability of Regulatory Relationship
1
0.8
0.6
0.4
0.2
0
Transcription Factors
b
LWD1/2
CTF6
CTF1
CCA1/LHYa
CCA1/LHYb
PRR9/5
PRR7
TOC1
CTF2
CTF3
CTF4
CTF7
LNK1a
LNK1b
LNK2a
LNK2b
ELF4a
ELF4b
ELF4c
EC
ELF3
LUX/
NOX
RVE6a
RVE8
RVE6c
CTF5
RVE6b
(h)
0
2
4
6
8
10
12
14
16
18
20
22


Kaladp0022s0168.rve6
Kaladp0011s1342
Kaladp0878s0025
Kaladp0047s0123.lnk1
Kaladp0032s0115.prr9/5
Kaladp0577s0020.rve8
Kaladp0045s0206.elf4
Kaladp0577s0020.rve8
Kaladp0099s0129.lnk2
Kaladp0496s0018.cca_lhy
Kaladp0748s0043
Kaladp0039s0732.elf3
Kaladp0059s0037.elf4
Kaladp0607s0046.lnk1
Kaladp0040s0446.toc1
Kaladp0037s0163.elf4
Kaladp0033s0047.lux/nox
Kaladp0032s0054.che
Kaladp1154s0002
Kaladp0009s0042
Kaladp0066s0115.cca1/lhy
Kaladp0674s0030
Kaladp0019s0045.rve6
Kaladp0055s0349.rve6
Kaladp0007s0017
Kaladp0101s0041.prr7


a
Kaladp0748s0043
Kaladp1154s0002
Kaladp0009s0042
Kaladp0007s0017
Kaladp0674s0030
Kaladp0011s1342
Kaladp0878s0025
b
Kaladp0047s0123 (LNK1)
Kaladp0607s0046 (LNK1)
Kaladp0060s0264 (LNK2)
Kaladp009s0129 (LNK2)
Kaladp0101s0041 (PRR7)
Kaladp0059s0037 (ELF4)
# of target genes
Z-score
Hours after first light


signal transduction
signal transduction by protein phosphorylation
protein phosphorylation
negative regulation of molecular function
protein dephosphorylation
polysaccharide catabolic process
peptidyl-tyrosine phosphorylation
phosphorelay signal transduction system