Extension of a genetic network model by iterative experimentation and mathematical analysis

James CW Locke1,2,3, Megan M Southern1, László Kozma-Bognár1,4, Victoria Hibberd1, Paul E Brown1,3,4, Matthew S Turner2,3 and Andrew J Millar1,3,4,*

1 Department of Biological Sciences, University of Warwick, Coventry, UK, 2 Department of Physics, University of Warwick, Coventry, UK and 3 Interdisciplinary Programme for Cellular Regulation, University of Warwick, Coventry, UK.

* Corresponding author. Institute of Molecular Plant Sciences, University of Edinburgh, Rutherford Building, Mayfield Road, Edinburgh EH9 3JH, UK. Tel.: ++ 44 1316513325; Fax: ++ 44 1316505592; E-mail: andrew.millar@ed.ac.uk

Received 5.4.05; accepted 7.6.05

Circadian clocks involve feedback loops that generate rhythmic expression of key genes. Molecular genetic studies in the higher plant Arabidopsis thaliana have revealed a complex clock network. The first part of the network to be identified, a transcriptional feedback loop comprising TIMING OF CAB EXPRESSION 1 (TOC1), LATE ELONGATED HYPOCOTYL (LHY) and CIRCADIAN CLOCK ASSOCIATED 1 (CCA1), fails to account for significant experimental data. We develop an extended model that is based upon a wider range of data and accurately predicts additional experimental results. The model comprises interlocking feedback loops comparable to those identified experimentally in other circadian systems. We propose that each loop receives input signals from light, and that each loop includes a hypothetical component that had not been explicitly identified. Analysis of the model predicted the properties of these components, including an acute light induction at dawn that is rapidly repressed by LHY and CCA1. We found this unexpected regulation in RNA levels of the evening-expressed gene GIGANTEA (GI), supporting our proposed network and making GI a strong candidate for this component.

Molecular Systems Biology 28 June 2005; doi:10.1038/msb4100018

Subject Categories: metabolic & regulatory networks; plant biology

Keywords: biological rhythms; gene network; mathematical modelling; parameter estimation

Introduction

A circadian system that generates biological rhythms with a period of approximately 24 h is found in organisms ranging from cyanobacteria to mammals. The system is capable of sustained oscillations under constant environmental conditions and maintains synchrony with the environment by entraining to rhythmic cues of the day/night cycle, especially input signals from light. Circadian rhythms allow diverse biological processes to occur at times in the day/night cycle (phases) that confer a selective advantage: it might be important, for example, that a particular process occurs in anticipation of a light/dark transition. The molecular mechanism of the circadian clock has been studied in several model organisms. A shared feature of these systems appears to be that the rhythms are generated by the interactions of rhythmically expressed genes that form positive and negative feedback loops (Dunlap, 1999).

Computational models of these feedback loops have been developed for a variety of organisms including the fungus Neurospora crassa (Leloup et al., 1999; Ruoff et al., 2000, 2001), the fruitfly Drosophila melanogaster (Tytson et al., 1999; Ueda et al., 2001; Smolen et al., 2004) and the mouse (Forger and Peskin, 2003; Leloup and Goldbeter, 2003). These models have shown that, within defined parameter ranges, the regulatory networks proposed from experimental data are capable of reproducing the main characteristics of circadian rhythms. Simple models indicate that a single feedback loop is sufficient to generate robust 24 h oscillations (Leloup et al., 1999; Ruoff et al., 2000, 2001), although the experimental data show that a series of interlocked feedback loops are important for generating the observed circadian rhythms (Glossop et al., 1999; Lee et al., 2000). It is an open question why circadian systems have evolved a more complex structure. Recent mathematical studies proposed that interlocked feedback loops increase the flexibility of regulation during evolution (Rand et al., 2004) and enhance precision (Stelling et al., 2004).

In higher plants, the circadian system controls many processes, including leaf movement, photoperiodism, and photosynthesis. The first part of the clock mechanism in Arabidopsis to be identified was proposed to comprise a feedback loop, in which two partially redundant genes encoding similar DNA-binding proteins, LATE ELONGATED HYPOCOTYL (LHY) and CIRCADIAN CLOCK ASSOCIATED 1 (CCA1), repress the expression of their activator, TIMING OF CAB EXPRESSION 1 (TOC1) (Alabadi et al., 2001). We refer to this single loop as the LHY/CCA1–TOC1 network. Light can activate LHY and CCA1 expression, possibly by several...
mechanisms (Wang and Tobin, 1998; Martinez-Garcia et al., 2000; Kim et al., 2003), providing a potential pathway for light input to the clock. Several other rhythmically expressed genes have been associated with the Arabidopsis circadian system (reviewed in Eriksson and Millar, 2003). For example, EARLY FLOWERING 3 (ELF3) (McWatters et al., 2000; Covington et al., 2001), GIGANTEA (GI) (Fowler et al., 1999; Park et al., 1999) and EARLY FLOWERING 4 (ELF4) (Doyle et al., 2002) are genes expressed in the evening. Mutations in these genes strongly affect circadian rhythms and reduce LHY and CCA1 gene expression, but their functions have not been located in more detail within the LHY/CCA1–TOC1 network.

The LHY/CCA1–TOC1 network alone did not readily account for some aspects of circadian behaviour, such as the long delay between TOC1 transcription in the evening and LHY/CCA1 activation the following morning (Alabadi et al., 2001; Salome and McClung, 2004). Our previous differential equation model of the LHY/CCA1–TOC1 loop confirmed that this network failed to fit certain experimental data and quantitatively tested a range of its predicted behaviours (Locke et al., 2005). For example, we showed that the LHY/CCA1–TOC1 loop could not reproduce the short-period phenotype of plants that carry loss-of-function mutations in either LHY or CCA1 (Green and Tobin, 1999; Mizoguchi et al., 2002; Locke et al., 2005). The delay required for the model to fit appropriate phases of gene expression was estimated at 12 h between TOC1 transcription and LHY/CCA1 activation (Locke et al., 2005). There is no obvious mechanism for this delay, reinforcing the suggestion that TOC1 protein may activate LHY and CCA1 expression indirectly.

Here, we extend the LHY/CCA1–TOC1 network beyond the structures inferred solely from data, in new mathematical models that we use to direct further experimentation. To check the effects of each addition to the network, the outputs of the extended models are compared to published data and to the new experiments. The biochemical parameter values required in the model are constrained by the time-series data but have not been measured directly, so we made a global search of parameter space, in contrast to previous clock models. This reduces the possibility that problems with the model are due to a particular set of parameter values, allowing us to focus on the network structure. The fit of the model to experimental data is dramatically improved by the addition of two hypothetical components, X and Y, to the model. Their properties are predicted; X remains to be identified, whereas experimental analysis shows that GI has several of the properties predicted for Y. The model suggests further experiments: we expect that iterative application of modelling and experiment will facilitate a more quantitative understanding of the Arabidopsis circadian clock.

Results

Limitations of the LHY/CCA1–TOC1 network

Our previous simulations using the single-loop LHY/CCA1–TOC1 network (Supplementary Figure 1) showed that it was possible for this network to correctly reproduce the phases of TOC1 and LHY RNA accumulation in wild type (WT) under light–dark cycle (LD) 12:12. (In this and subsequent models, we use a single gene, LHY, to represent both CCA1 and LHY functions; see Supplementary text.) However, simulated TOC1 RNA levels remained high until LHY protein accumulated, rather than falling after dusk as observed (Mizoguchi et al., 2002). This was exaggerated by halving the LHY mRNA translation rate in the simulation (representing lhy or cca1 loss-of function mutants), which incorrectly predicted a long-period phenotype. Thus, there must be another factor responsible for reducing TOC1 expression, which is not modelled by this network (Locke et al., 2005).

Studies of a fluorescent protein, TOC1 fusion protein, suggest an additional limitation (Mas et al., 2003b). The TOC1 fusion was shown to be close to its minimum abundance before dawn under LD12:12, whereas according to the single-loop LHY/CCA1–TOC1 network, TOC1 should be activating LHY transcription maximally at that time (Locke et al., 2005). This suggests that either the active form of TOC1 is present at a far lower concentration than bulk TOC1 protein, perhaps in a complex, or that an additional, TOC1-dependent component is the direct activator of LHY and CCA1.

A third problem is that the LHY/CCA1–TOC1 network did not respond to day length (simulated gene expression profiles were identical in LD cycles with long and short photoperiods, data not shown), whereas it is clear experimentally that the clock has a later phase under longer photoperiods (Millar and Kay, 1996; Roden et al., 2002). This limitation occurs because light input to this network is modelled only by the activation of LHY expression at dawn, so the model is insensitive to light at the end of the photoperiod. Indeed, LHY and CCA1 expression fall to a low level before the end of a 12 h photoperiod (Kim et al., 2003), so another mechanism is required to mediate light input at the end of the day.

Model one—the LHY/CCA1–TOC1–X network

We extended the single-loop LHY/CCA1–TOC1 network by adding components that would address these limitations, as directed by the experimental data. After each addition, we tested network parameters until it became clear that the new network could not account for further experimental data. We identified optimal parameters for the most promising of the extended, single-loop models, which we term the LHY/CCA1–TOC1–X network (Figure 1). Firstly, light activation of TOC1 transcription was included to provide light input at the end of the day and, conversely, to reduce TOC1 activation immediately after lights-off. Secondly, an additional gene X was added to the network after TOC1, with nuclear X protein as the immediate activator of LHY instead of nuclear TOC1. Thirdly, as the F-box protein ZEITLUPE (ZTL) has been shown to degrade TOC1 protein more effectively during the night (Mas et al., 2003b), we added this factor into our network equations (see Supplementary text).

Figure 1 shows the simulated expression profiles for the LHY/CCA1–TOC1–X network using the optimal parameter set (Supplementary Table 1). TOC1 RNA levels peak at dusk in WT under LD12:12, and LHY RNA levels at dawn. The model allows TOC1 mRNA levels to drop before LHY levels rise, as observed in experiment. Including gene X within the model permits simulated TOC1 protein levels to fit well with the published data (Supplementary Figure 2). ztl mutants were...
unlikely to be a simple light activation of TOC1 characterising circadian rhythms in the We sought more direct evidence for this light input by et al. conditions show a damping, short-period oscillation (Alabadi of-function mutant. RNA data for TOC1 protein represses TOC1 transcription at dawn and activates TOC1 transcription throughout the day. TOC1 activates a putative gene X, which in turn activates LHY. Nuclear LHY protein represses TOC1 transcription. Right panel: Simulation of mRNA levels for the optimal parameter set. In all figures, filled box above the panel represent dark interval and open or no box represent light interval. LHY mRNA (dotted line) peaks at dawn in LD12:12 and TOC1 (solid line) falls after dusk, due to the loss of light activation. modelled by reducing the degradation rates of TOC1 protein in the cytoplasm and the nucleus by 50%. This results in a long-period phenotype, with a period of 32 h, similar to or longer than the period of ztl mutants (Mas et al., 2003b). A prediction of X mRNA and protein levels is also possible (Supplementary Figure 2): X mRNA peaks in the middle of the night under LD12:12 and nuclear X protein levels peak at dawn. Strong x mutants have the same predicted phenotype as the strongest phenotype of toc1 loss-of-function mutants, causing arrhythmia due to the lack of LHY activation (data not shown). The pattern of X mRNA accumulation and its mutant phenotype are similar to those of characterised genes such as ELF4 (Doyle et al., 2002). However, this model still incorrectly predicts a long period in the simulated cca1 single mutant (Supplementary Table 2) and the strong, LL activation of TOC1 transcription causes several problems, for example the model becomes arrhythmic under LD cycles with long photoperiods (data not shown).

Experimental characterisation of the cca1;lhy double mutant

The response of circadian phase to day length (Millar and Kay, 1996; Roden et al., 2002) strongly suggested that the circadian system receives at least one light input in addition to the activation of LHY and CCA1 expression, yet simulations with the LHY/CCA1–TOC1–X network indicated that this was unlikely to be a simple light activation of TOC1 transcription. We sought more direct evidence for this light input by characterising circadian rhythms in the cca1;lhy double loss-of-function mutant. RNA data for cca1;lhy mutants in constant conditions show a damping, short-period oscillation (Alabadi et al., 2002; Mizoguchi et al., 2002), which has been described as arrhythmia. We repeated these experiments using luciferase imaging (Figure 2). In the cca1;lhy mutant, promoter activity of CCA1 and of the clock output genes CCR2 and CAB2 showed an 18 h rhythm for at least three cycles in constant light (LL), which subsequently lost amplitude. The rhythm is more robust in LL but is also apparent in constant dark (DD) (Figure 2). The double mutant retains a regulatory network capable of supporting rhythmic gene expression. Reproducible entrainment of the double mutant by LD cycles was implicit in previous reports, suggesting that entrainment by light is still possible in the residual network (Alabadi et al., 2002; Mizoguchi et al., 2002) (Figure 2). To test this more stringently, we generated a phase transition curve (PTC) for the WT and double mutant (Figure 3). The PTC
shows the response of an oscillator to a resetting stimulus and is plotted as the phase to which the oscillator is set (‘new phase’), for each phase at which the resetting stimulus is applied (‘old phase’). In WT, light pulses induced phase delays during the early subjective night and phase advances during the late subjective night, whereas relatively small phase shifts were elicited during the subjective day. The WT showed a type 1 (weak) resetting pattern with less than 6 h maximal phase shifts, in contrast to the type 0 (strong) resetting observed in a previous report (Covington et al., 2001), probably due to the lower fluence of our light stimulus. In contrast, the double mutant showed type 0 resetting: irrespective of the phase of the light stimulus, the clock was reset to a narrow phase range (circadian time (CT) 20–23). Light input to a residual, light stimulus, the clock was reset to a narrow phase range. In contrast, the double mutant showed type 0 resetting; irrespective of the phase of the light stimulus, the clock was reset to a narrow phase range (circadian time (CT) 20–23). Light input to a residual, light stimulus, the clock was reset to a narrow phase range.

**Model two—the interlocked feedback loop network**

Removing LHY function from the single-loop models prevents any oscillation (data not shown), so none of these models can reproduce the entrainable, damped rhythms observed in cca1;lhy plants. We therefore developed an interlocked feedback loop network that is capable of oscillation in simulated cca1;lhy double mutants (Figure 4). A hypothetical gene Y activates TOC1 transcription and TOC1 protein represses Y transcription, forming a feedback loop. The proposal that TOC1 has a negative function as well as a positive one is novel. Light input into this loop occurs via transcriptional activation of Y rather than of TOC1; there is as yet no evidence of direct light activation of TOC1 (Makino et al., 2001). Light input to Y can both be through an acute response at dawn similar to that for LHY and as a constant activation term throughout the day. Y is also repressed by LHY, as this allowed the network to fit the WT as well as the cca1;lhy experimental data. LHY therefore acts as a powerful delaying factor in the early day, when it inhibits expression of both TOC1 and Y.

Optimal parameters for the interlocked feedback loop network (Supplementary Table 3) were identified (see Computational methods). The optimised model achieved a good fit to experimental results that were specifically required by the optimisation process, showing that the proposed network is sufficient to explain these data. Simulations of the WT and cca1;lhy mutant using the optimal model fit well to RNA expression profiles in DD and LD12:12 (Figures 5A and B). For the WT simulation (Figure 4), LHY mRNA peaks at dawn, TOC1 at dusk, and the oscillations follow a stable limit cycle with a period of 26 h in DD. TOC1 mRNA levels under LD cycles are shown to increase at dawn. This is due to the induction of Y by light activating TOC1 expression, overcoming the repression by LHY protein. The simulation of cca1;lhy gives a low-amplitude oscillation in DD with a 17 h period (Figure 4), as observed experimentally (Figure 2). Under LD12:12, TOC1 mRNA oscillates with an early peak phase in the double mutant, ~5 h after dawn, as specified in the optimisation. The rhythm of TOC1 expression in the double

**Figure 4** The interlocked feedback loop network. Left panel: Network diagram. Compared to Figure 1, TOC1 is activated by light indirectly via hypothetical gene Y. Y activates TOC1 transcription and both LHY and TOC1 repress Y transcription, forming a second feedback loop. Right panel: Simulation of LHY (dashed line) and TOC1 (solid line) mRNA levels for the optimal parameter set, representing WT (top) and cca1;lhy double mutant (bottom) in DD. Translation rate of LHY mRNA in simulated mutant is 1/1000 WT value. Period of WT in DD is 26 h and period of mutant is 17 h.

**Figure 5** Comparison of interlocked feedback loop simulations (dashed line) under LD to data (solid line). (A) TOC1 mRNA levels in WT plants entrained to LD12:12, left axis; TOC1 mRNA levels relative to UBQUITIN (UBQ) (Makino et al., 2000), right axis. (B) LHY mRNA levels in WT plants entrained to LD12:12, left axis; data from Kim et al. (2003), right axis. (C–E) TOC1 mRNA levels in WT (C), cca1;lhy mutant (D) and cca1 mutant (E) entrained to LD16:8; data from Mizoguchi et al. (2002). Translation rate of LHY in simulated cca1 is set to 1/2 WT value. Highest value of data and simulation is set to 1, for each panel. (F) TOC1 protein levels for WT simulation entrained to LD12:12; TOC1 fusion protein data from Mas et al. (2003b).
mutant also shows a higher amplitude than WT (Figure 4), which is observed experimentally (Mizoguchi et al., 2002) but was not specified during optimisation. Figures 5C and D show similar expression profiles for simulated and observed (Mizoguchi et al., 2002) TOC1 mRNA in the WT and cca1;lhy mutant under LD16:8 (note that normalisation of data and simulated values obscures the change in amplitude in this figure). TOC1 mRNA anticipates dawn in the simulation of the cca1;lhy double mutant, which has not been so clearly observed in published experimental data and points to an area for future experimentation.

**Analysis and validation of the interlocked feedback loop network**

The interlocked feedback loop model with the optimal parameters not only fits the above data but its behaviour is also robust to parameter changes. This is widely thought to be a realistic requirement for models of biological regulation, because effective parameter values may be poorly buffered in biology. Changes in the period and amplitude of TOC1 RNA oscillation under LL were examined after a 5% increase or decrease of each parameter value in turn (Supplementary Figure 3). The resulting change in period varied from 0 to 4%. As for previous clock models (Smolen et al., 2004; Locke et al., 2005), some parameters are more sensitive to change than others. The most sensitive parameters are those involved in TOC1 degradation, X translation and X nuclear transport. The period and amplitude of this model are much less sensitive to parameter changes than the single-loop LHY/CCA1–TOC1 model (data not shown), suggesting that some of the weaknesses of the single-loop model have been overcome.

Simulations using the optimal parameter set also fit well to several experimental results that were not specified in the optimisation, giving additional support for the proposed network structure. This is the first model that fits well to LL data for LHY and TOC1 mRNA levels. The WT period in LL is correctly shorter (25 h) than the period in DD (26 h; Supplementary Figure 4) although this effect is less than that observed experimentally. The rhythms in LL generally have a higher amplitude than in DD, as observed. The model correctly predicts the short-period phenotype of cca1 and lhy single mutants in LL and DD (Supplementary Table 2), and the early phase of TOC1 RNA expression in the single mutant under LD12:12 (Figure 5E). The single mutants were simulated by halving the LHY mRNA translation rate. Simulated overexpression of LHY produced arrhythmia with low levels of TOC1 mRNA (data not shown), as observed in plants that overexpress LHY or CCA1 (Schaffer et al., 1998; Wang and Tobin, 1998; Alabadi et al., 2001). Protein levels are also well fitted: simulated LHY protein levels (data not shown) peak 1–2 h after LHY mRNA levels, as observed (Kim et al., 2003).

Figure 5F compares simulated and measured (Mas et al., 2003b) TOC1 protein levels in WT, showing low levels at dawn in both cases. The optimal parameter set has minimised the light regulation of TOC1 degradation (<1% of total TOC1 degradation), indicating that light-regulated degradation (Mas et al., 2003b) is not required to fit these data. Simulation of ztl mutants by halving the total TOC1 degradation rate results in a 28 h period phenotype, again similar to that observed in ztl mutants (Mas et al., 2003b). A simulated toci mutant results in lower levels of LHY mRNA as expected from experiment (Alabadi et al., 2001), and simulated TOC1 overexpression is predicted to increase LHY mRNA levels. The observed decrease in LHY mRNA where TOC1 is overexpressed (Makino et al., 2002; Mas et al., 2003a; Somers et al., 2004) remains paradoxical, since one would expect overexpressing an activator of LHY to cause its levels to rise.

Simulations of the WT and cca1;lhy double mutant PTCs were performed, as shown in Figure 3. Both simulations are similar to our experimental data, with a type 1 PTC in the WT and a type 0 PTC in the double mutant. Increasing the light level in the WT simulation results in a type 0 PTC (data not shown), as previously observed (Covington et al., 2001). As expected, the entrained phase of the interlocked feedback model is photoperiod responsive (Supplementary Figure 5), with simulated mRNA levels peaking later under longer photoperiods, as observed (Roden et al., 2002; Yanovsky and Kay, 2002). Light input to Y allows the network to respond to light throughout the day. This network will therefore be a good starting point for models of the photoperiod sensor involved in flowering time. The photoperiod range of entrainment is approximately from 3:16 h light for a 24 h period, and the simulations remain entrained for an approximate period range of 22–30 h, where half the period is in light and half in dark. At the end of the ranges, entrainment produces a beat in the amplitude, although with little effect on phase. The balance of light input to LHY, Y and ZTL should now be examined in greater detail to determine how their contributions affect circadian entrainment.

**GIGANTEA is a candidate for Y**

The interlocked feedback model predicts a distinctive pattern of Y mRNA accumulation in the WT and double mutant (Figure 6). Y mRNA levels peak at the end of the day, but also increase transiently at dawn due to the acute light response of Y transcription. This early expression is quickly repressed by rising LHY protein levels, delaying the peak in Y mRNA level until after LHY protein is degraded at the end of the day. Y transcription is then repressed as TOC1 protein levels begin to rise during the night (Figure 4). In the cca1;lhy double mutant, however, the light activation of Y at dawn is de-repressed, resulting in a much stronger activation than in WT, and causing Y mRNA levels to peak soon after dawn. No gene with this expression pattern had been observed experimentally.

In order to identify Y, we analysed the transcript abundance of clock-affecting genes with peak RNA levels in the evening in WT and cca1;lhy double mutant seedlings. Tissue samples were harvested across the light–dark transitions in one LD cycle, followed by one cycle in LL. GI mRNA levels fitted very well to our predicted mRNA profiles for Y (Figure 6). GI was shown to be significantly but transiently light activated in the WT and had a very strong light response in the double mutant. The subsequent circadian peak also fitted closely to the prediction for Y mRNA, including the 12 h phase advance in the mutant relative to WT (Figure 6). The tentative identification of Y as GI allowed us to test whether Y in our model fitted additional, published results for GI; indeed, further data do
support this proposed function of GI. GI mRNA is at a low, arrhythmic level in plants that overexpress LHY (Fowler et al., 1999) or TOC1 (Makino et al., 2002): this agrees with model predictions (data not shown) and occurs because both LHY and TOC1 repress Y transcription (Figure 4). The sequence of the GI promoter includes several Evening Elements, the putative binding sites for LHY (Harmer et al., 2000). gi loss-of-function mutations result in low-amplitude circadian rhythms, with low levels of LHY and CCA1 RNA and either shorter or longer circadian periods (Fowler et al., 1999; Park et al., 1999) or, in some conditions, in arrhythmia (A Hall, personal communication). A simulated null mutation of Y indeed results in still lower LHY transcription and therefore in arrhythmia. If another gene in Arabidopsis can substitute for a fraction of Y function, then null mutants will avoid arrhythmia. Supplementary Figure 6 shows the oscillation of LHY mRNA levels in a simulated partial loss-of-function y mutant, where Y translation rate has been halved compared to the WT rate. As observed in gi mutants, the oscillations have reduced LHY expression and a low amplitude both in LD cycles and in LL (Mizoguchi et al., 2002).

**Discussion**

We use a joint, experimental and mathematical approach to understand the plant circadian clock as an example of a regulatory subnetwork that is not completely identified. We start from the first proposed feedback loop of the circadian clock mechanism in Arabidopsis, the LHY/CCA1–TOC1 network (Alabadi et al., 2001). Comparing model predictions with experimental results, we have progressively incorporated additional components and interactions identified by molecular genetics or inferred from physiological analysis. The final, interlocking loop model accounts for a greater range of data than the single-loop models, including the entrainable, short-period oscillations in the cca1;thy double mutant. In developing this model, we included two putative genes X and Y, and used experiments designed from the model predictions to identify GI as a candidate gene for Y. Additional components of the plant circadian clock mechanism almost certainly remain to be identified, but we believe that this model is a significant step forward in understanding of the timing mechanism.

The prediction of new components is a particularly beneficial outcome from formal modelling of a system that has not been completely identified by experiment. Mathematical models, in contrast to intuitive reasoning, can produce quantitative predictions of dynamic processes that allow detailed experimental design. This was important: the acute light activation of Y in WT was predicted to be very transient (peak 25 min after lights-on; Figure 6), allowing us to target our tissue sampling to the appropriate interval, whereas conventional sampling had obscured this induction of GI RNA (Mizoguchi et al., 2002). The interlocked feedback model now highlights the importance of GI as a component of light input to the clock, a role that had not previously been emphasised and should now be tested in greater detail. The activation of TOC1 by GI in an interlocked feedback loop is also a new proposal, which is consistent with the timing of peak GI expression before TOC1. Mutants that remove both the loops, such as the lhy;cca1;gi triple mutant, should now be tested to determine whether further oscillating subnetworks remain in their absence. A recent study has suggested the existence of a feedback loop between APRR9/APRR7 and LHY/CCA1 (Farre et al., 2005). Including this loop would not affect our conclusions on the residual network in the cca1;thy double mutant, which would lack this additional loop. As more data become available, it will be possible to determine how the PRR genes should be included into the network model. The component(s) that activate CCA1 and LHY at the end of the night remain to be identified: the model predicts the likely accumulation pattern of such a component, X. The level of detail in such predictions is obviously limited by the data upon which the model is based, so including statistical measures of uncertainty with the predictions will be increasingly important (Brown and Sethna, 2003).

Each model makes further, qualitative predictions that appear robust and readily testable. The constant activation of TOC1 by light reproducibly caused arrhythmia of the LHY/CCA1–TOC1–X model under long photoperiods or LL, for example, which is not observed in WT plants. This highlights the importance of rhythmic inhibition of the light input (Roenneberg and Merrow, 2002), which is a wide spread feature of clocks (Fleissner and Fleissner, 1992; Jewett et al., 2005 EMBO and Nature Publishing Group)
1999). It is reminiscent of the ELF3-dependent zeitnehmer function observed in Arabidopsis (McWatters et al., 2000; Covington et al., 2001). In the interlocked loop model, repression of Y by LHY and by TOC1 are sufficient to gate the light activation of Y, so we had no justification for further additions to this model. Clearly, such models should be interpreted with caution, because undiscovered components cannot be included explicitly. A model that accurately recapitulates the regulation of known components is very likely to have captured the relevant effects of the ‘hidden’ components. The model can advance understanding and make useful predictions but might not capture the real number or mechanism of the hidden components. For example, we model the direct activator of LHY and CCA1 as the product of a TOC1-activated gene, X, which could be a minor population of modified TOC1 protein or TOC1-dependent protein complex. We assume that Y mediates both the second light input and the additional feedback loop for parsimony, which is now supported by data on GI, although these functions could in principle be split among several components.

The importance of the light input pathways in our models was to be expected, because the plant circadian system is known to interact with multiple photoreceptor pathways in a complex fashion (reviewed in Fankhauser and Staiger, 2002; Millar, 2003). The tracking of multiple phases during entrainment is thought to require at least two light inputs to two feedback loops (Rand et al., 2004), for which we are present in our final model. The entrainment patterns of the Arabidopsis clock under different photoperiods (Millar and Kay, 1996) indicate that the phase of the clock does not simply track dawn. Therefore, the clock must receive light input(s) at times other than the dark–light transition. In our model, LHY allows light input at dawn, while input to Y and ZTL is potentially effective throughout the day. The known input photoreceptors could in future be explicitly included, providing quantitative estimates of their function for comparison to data from plant photobiology. Similarly, the models will help to reveal how the circadian output pathways allow the few genes of the clock to control over a thousand rhythmically regulated genes in the Arabidopsis genome (Harmer et al., 2000). However, the complexity of such biological networks is likely to limit the quantitative accuracy of early models, so the potential value of simplified experimental model systems that facilitate the link to mathematical analysis is clear. These will include synthetic gene networks in microbial hosts but also ‘reduced’ systems: we have recently characterised circadian rhythms in seedlings without light exposure, in which both the complexity of the circadian system and the number of clock-controlled target genes are greatly reduced (A Hall et al., unpublished results).

Materials and methods

Plant materials and growth conditions

Wassilewskija (Ws) WT and cca1-1; lhy-2l (termed cca1; lhy) double mutants in the Ws background (Hall et al., 2003) were used in all experiments. Luciferase reporter gene fusions containing the promoter region of CCA1 (CCA1:LUC+), CHLOROPHYLL A/B-BINDING PROTEIN2 (LHCBL1.1)(CAB2:LUC+), and COLD AND CIRCADIAN REGULATED 2 (CCR2:LUC+) were introduced into Ws and mutant plants by Agrobacterium-mediated transformation, essentially as described (Hall et al., 2003). For each genotype and reporter, three independent transgenic lines were tested in each experiment; all gave very similar results. Light sources were as described (Hall et al., 2003).

Seedlings for luminescence analysis were grown under 12 h light:12 h dark cycles (LD12:12), as described (Hall et al., 2003). Seedlings for RNA analysis were grown under LD12:12 comprising 13–20 μmol m–2 s–1 red light for 6 days, followed by constant 13–20 μmol m–2 s–1 red light for 3 days. Samples of ~150 μl packed volume of seedlings were harvested into RNA later buffer (Ambion, Huntingdon, UK) to stabilise RNAs, starting in the last cycle of entrainment.

Luminescence and rhythm analysis

Luminescence of individual seedlings was measured with an automated luminometer (Doyle et al., 2002). Rhythmic data were analysed using the fast Fourier transform nonlinear least squares procedure (Plautz et al., 1997) through the Biological Rhythms Analysis Software System, available online (Brown, 2004a). Variance-weighted mean periods and standard errors are presented. To create PTCs, seedlings expressing the CCR2:LUC+ reporter were grown and entrained as above, and then transferred to DD at the predicted time of lights-off. Luminescence signals were monitored for 5 days in DD. After 24 h in DD, separate populations of seedlings were treated with 15 μmol m–2 s–1 red light for 1 h and returned to DD at 3 h intervals. The free running period and phase of the control (nontreated WT and mutant) plants was used to calculate the circadian time of the light treatments (‘old phase’). The time of the next peak of CCR2:LUC+ expression was determined in the treated plants and circadian time of the ‘new phase’ set at the light pulse was estimated using the cognate period value.

RNA analysis

Seedlings were homogenised in RLT buffer (Qiagen, Crawley, UK) using a MixerMill MM300 at a frequency of 30 s–1 for 3 min which a 5 mm stainless steel cone ball (Retsch, Leeds, UK). Total RNA was isolated using a Plant RNeasy kit and RNase-free DNase (Qiagen, Crawley, UK) according to the manufacturer’s instructions. A 1 μg portion of total RNA was reverse-transcribed using the RevertAid cDNA kit (Fermentas, Helena Biosciences, Sunderland, UK) with random hexamer primers, according to the manufacturer’s instructions. GI sequence abundance in each cDNA sample was assayed by quantitative PCR in an ABI PRISM 7700 using ABI SYBRgreen PCR Mix (Applied Biosystems, Warrington, UK) in a final volume of 15 μl. GI sequence abundance was normalised relative to ACTIN2 (ACT2), using a cDNA dilution series for each primer set in each experiment. The following primers were used:

GI forward primer AATTCAGACCCGCGCTATTGT, GI reverse primer GTTGTTCTCTGCTGCGAGAACCT; ACT2 forward primer CAGTGTCTGGATCGGAGGAT, ACT2 reverse primer TGACAACTCGATGACCTGA, each at 300 nM

Each RNA sample was assayed in triplicate. Data shown are a representative trace from two independent biological replicates that gave very similar results.

Computational methods

As there is too little data to discriminate LHY from CCA1 regulation and function, we combine them in the single model component ‘LHY/CCA1’ in order to simplify our models (see Supplementary text); for brevity, we refer to this joint component as LHY, as in our previous work. Our method of parameter estimation uses a cost function, which is based on reproducible, qualitative features of experimental data, to score the performance of a model with a test parameter set. The cost function is minimised by the optimisation procedure described (Locke et al., 2005). A low cost (indicating a good fit) is obtained for parameter sets that allow the model of WT plants to be entrained in LD12:12.
cycles, with LHY RNA levels that peak at dawn, TOC1 RNA levels that peak at dusk and oscillations with a period greater than 24 h in DD. As there is only limited, noisy experimental data for the mRNA oscillation of TOC1 and LHY in DD, it is difficult to verify that the TOC1 and LHY mRNA levels converge to a stable limit cycle. The cost function only requires that LHY and TOC1 mRNA levels oscillate with slow damping in DD, giving a reasonable score if the size of oscillation has dropped by 25% over 300 h (Strayer et al., 2000; Kim et al., 2003). Developing the model based on LD and DD data allowed subsequent testing of the model by comparison to the larger amount of experimental data available from LL conditions.

The interlocked feedback loop network proposed here was scored both as a model of WT and of the cca1;lhy mutant. The double mutant was simulated by reducing the translation rate of LHY to 1/1000th of its WT value. This simulated mutation led to arrhythmia in all the single-loop models (data not shown). Additional terms were introduced to the cost function to score models of the double mutant, specifying entrainment under LD12:12 with peak TOC1 expression 5 h after dawn and oscillations with a period of 18 h or less in DD. To enable TOC1 activation sufficiently early in the day in the double mutant, we required that Y transcription peaked sharply at dawn in the double mutant.

The 20 parameter sets with the lowest costs (which allowed the model to best fit the specified criteria) all simulated similar gene expression profiles in WT and cca1;lhy backgrounds (data not shown). An optimal parameter set was chosen from these 20 by comparing the simulated rhythms to experimental data that were not included in the cost function (see Results).

The equations were solved using MATLAB (Mathworks, Cambridge, UK). Parameter optimisation was carried out (Locke et al., 2005) by compiling MATLAB code into C and running the code on a task farm computer consisting of 62 × 2.6 GHz Xeon CPUs. We have developed a user-friendly interface, Circadian Modelling, to allow simulations using this and other circadian models, without MATLAB. This software and files for our final model are available online (Brown, 2004b).

Supplementary information
Supplementary information are available at the Molecular Systems Biology website (www.nature.com/msb).

Acknowledgements
We are grateful to OE Biriring-Akman, D Salazar, JA Langdale, CD Westbrook and DA Rand for useful discussions, and to N Shariff for technical assistance. JCWL was supported by a postgraduate studentship from the Gatsby Charitable Foundation; MMS was supported by a postgraduate studentship from BBiSC; LKB was supported by an EMBO postdoctoral fellowship; experimental work was funded by grants G15231 and G19886 from BBiSC to AJM. Computer facilities were provided by the Centre for Scientific Computing at the University of Warwick.

References


© 2005 EMBO and Nature Publishing Group
Roden LC, Song HR, Jackson S, Morris K, Carre IA (2002) Floral responses to photoperiod are correlated with the timing of rhythmic expression relative to dawn and dusk in Arabidopsis. Proc Natl Acad Sci USA 99: 13313–13318