COMMUNICATIONS BIOLOGY

ARTICLE

Waveforms of molecular oscillations reveal circadian timekeeping mechanisms

Hang-Hyun Jo [1,2,3], Yeon Jeong Kim[4], Jae Kyoung Kim [5], Mathias Foo[6], David E. Somers[4] & Pan-Jun Kim[7,8,9,10]

Circadian clocks play a pivotal role in orchestrating numerous physiological and developmental events. Waveform shapes of the oscillations of protein abundances can be informative about the underlying biochemical processes of circadian clocks. We derive a mathematical framework where waveforms do reveal hidden biochemical mechanisms of circadian timekeeping. We find that the cost of synthesizing proteins with particular waveforms can be substantially reduced by rhythmic protein half-lives over time, as supported by previous plant and mammalian data, as well as our own seedling experiment. We also find that previously enigmatic, cyclic expression of positive arm components within the mammalian and insect clocks allows both a broad range of peak time differences between protein waveforms and the symmetries of the waveforms about the peak times. Such various peak-time differences may facilitate tissue-specific or developmental stage-specific multicellular processes. Our waveform-guided approach can be extended to various biological oscillators, including cell-cycle and synthetic genetic oscillators.

[1] Asia Pacific Center for Theoretical Physics, Pohang, Gyeongbuk 37673, Republic of Korea. [2] Department of Physics, Pohang University of Science and Technology, Pohang, Gyeongbuk 37673, Republic of Korea. [3] Department of Computer Science, Aalto University, FI-00076 Espoo, Finland. [4] Department of Molecular Genetics, The Ohio State University, Columbus, OH 43210, USA. [5] Department of Mathematical Sciences, Korea Advanced Institute of Science and Technology, Daejeon 34141, Republic of Korea. [6] School of Mechanical, Aerospace and Automotive Engineering, Coventry University, Coventry CV1 5FB, UK. [7] Department of Biology, Hong Kong Baptist University, Kowloon, Hong Kong. [8] Center for Quantitative Systems Biology, Hong Kong Baptist University, Kowloon, Hong Kong. [9] Institute of Computational and Theoretical Studies, Hong Kong Baptist University, Kowloon, Hong Kong. [10] Abdus Salam International Centre for Theoretical Physics, 34151 Trieste, Italy. Correspondence and requests for materials should be addressed to P.-J.K. (email: panjunkim@hkbu.edu.hk)

A variety of light-sensing organisms feature circadian clocks, which generate endogenous molecular oscillations with ~24 h periodicity and thereby control numerous physiological and behavioral events[1–4]. Despite the identification of biochemical mechanisms of circadian timekeeping in various organisms[5–9], our understanding of a design principle of these clock mechanisms is far from complete. For example, the mammalian clock protein BMAL1 exhibits the abundance oscillations[10], but these oscillations are not empirically required for the generation of circadian rhythms per se, leaving their biological roles still unclear[11–13]. As another example, the plant circadian system involves post-translational regulations, such as the degradation of PSEUDO RESPONSE REGULATOR 5 (PRR5) protein by ZEITLUPE (ZTL) protein[14,15]; however, a previous mathematical modeling suggests that such post-translational interactions may not be strictly required for the formation of the rhythms of any core clock components[16], raising a question about the fundamental role of these interactions.

The temporal trajectory of mRNA or protein concentration exhibiting a circadian rhythm can be characterized by its shape or waveform. A waveform of a protein expression profile, apart from its few characteristic quantities (period, amplitude, and peak phase)[17–21], has long been underappreciated, but recently recognized for its potential importance to clock function[16,22,23]. A cuspidate waveform, which shows a notable acuteness around its peak phase, was speculated to confer high-resolution timing of downstream biological events around the peak phase[16]. In addition, according to plant–clock experiments, precise changes in the waveform of *GIGANTEA* (*GI*) expression were sufficient to alter hypocotyl growth as a downstream phenotype[23]. Moreover, a specific circadian waveform seems crucial for the molecular arithmetic processes involved in daily starch degradation[24]. Although not in the circadian context, there are interesting reports that modifying the waveform shape of neuro-stimulating signals changes the efficiency of entraining the neural spiking activities[25]. Nevertheless, the reverse yet complementary view of the waveforms as a window to the inner biochemical mechanisms of circadian clocks has not yet been taken into consideration for systematic investigation.

Here, we report that the waveforms of clock protein profiles can serve as an information source of previously underexplored, biochemical mechanisms of circadian timekeeping. These mechanisms can be exemplified by the above PRR5–ZTL interaction and BMAL1 abundance oscillation. Interestingly, our waveform analysis predicts the considerable benefit of rhythmic regulation of protein degradation in reducing the biosynthetic cost of the waveform formation. Our mathematical framework is supported by previous, as well as our new, experimental data. This study can be extended to time-course data from various biological oscillators, such as cell cycle systems and synthetic genetic oscillators.

Results

Relationship between waveforms and cost. In a circadian system, the dynamics of protein production governs the protein concentration profile $x(t)$ over time and thereby its waveform. This dynamics can often be described by the following equation:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = g(t) - r(t)x(t), \qquad (1)$$

where $g(t)$ and $r(t)$ denote protein synthesis and degradation rates, respectively, as depicted in Fig. 1a. $g(t)$ is proportional to an mRNA concentration and a translation rate. Protein degradation with a rate $r(t)$ is driven by post-translational mechanisms. An oscillatory waveform of $x(t)$ satisfies $x(t) = x(t + T)$ with $T = 24$ h in diurnal light and dark cycles or $T \approx 24$ h in constant light or darkness. We stress that to maintain $x(t)$'s rhythmicity, $g(t)$ or $r(t)$ should not remain constant but change over time. We will consider the relationships between $x(t)$, $g(t)$, $r(t)$, and later, the cost c of protein production defined as

$$c \equiv \frac{\Delta x}{T} = \langle g(t) \rangle = \langle r(t)x(t) \rangle, \qquad (2)$$

where Δx denotes the amount of proteins synthesized over the period T, and $\langle \cdot \rangle$ represents a time average, e.g., $\langle g(t) \rangle \equiv (1/T) \int_0^T g(t)\mathrm{d}t$. The equalities $\Delta x/T = \langle g(t) \rangle$ and $\langle g(t) \rangle = \langle r(t)x(t) \rangle$ are derived from Eq. (1) and $x(t) = x(t + T)$. In other words, the cost c is defined as an average protein amount synthesized per time, which is equal to an average protein amount degraded per time. Because the circadian protein levels are periodic over time, the proteins must be synthesized as much as they
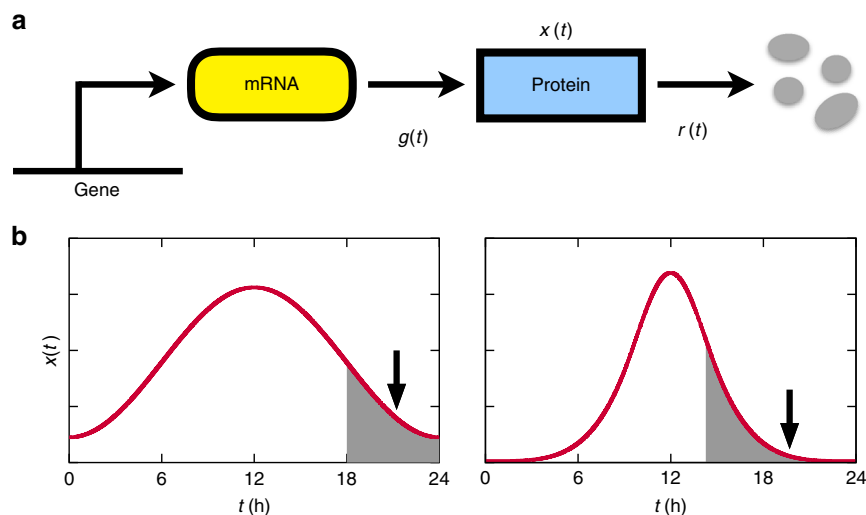

Fig. 1 Schematic diagrams of protein synthesis and turnover, and the resulting protein profiles in the circadian system. **a** Proteins are synthesized through mRNA-to-protein translation, and destined for degradation. **b** Cyclic protein abundances are represented by waveforms. For each waveform, the arrow indicates the point when $R(t) = r_{min}$ in Eq. (4), and the shaded area corresponds to the interval between the steepest decline and the trough. The right waveform ($r_{min} \approx 0.69\,\mathrm{h}^{-1}$) has larger r_{min} than the left waveform ($r_{min} \approx 0.30\,\mathrm{h}^{-1}$). For the definition of each notation in **a**, **b**, refer to Eq. (1)

are degraded. We will show step by step that the biosynthetic cost c of a protein waveform helps us decipher circadian degradation mechanisms, mainly through the examples from the plant circadian system. Then, we will focus on other cases such as the mammalian system.

In the case of the plant *Arabidopsis thaliana*, more than 20 clock genes have been discovered, and many of their mRNAs undergo high-amplitude cycling in their abundance[26,27]. This mRNA-level oscillation is a result of transcriptional control by other clock gene products or by light signals. In the core plant clock, the protein synthesis rate $g(t)$, which is largely proportional to the transcript concentration, would likely exhibit similar oscillatory patterns. On the other hand, the characteristics of the degradation rate $r(t)$ remain rather elusive for plant clock proteins, with only a limited number of experimental reports [28–32]. Given the clearly time-dependent nature of the protein synthesis rate, the degradation rate may not have to be also time-dependent, as demonstrated by the previous mathematical modeling[16]. Existing experimental data, nonetheless, indicate that plant clock proteins often seem to have time-specific or phase-specific degradation rates[28–31], raising a question about the beneficial effect of such rhythmic regulation of protein stability. One study suggests that rhythmic degradation rates allow nontrivial phase differences between transcript and protein profiles[18]. However, given that the phases of transcript profiles have relatively little functional significance, this previous study is unlikely to be about biologically beneficial effects of the rhythmic degradation rates.

We begin with the following observation: because $x(t) \geq 0$, $r(t) \geq 0$, and $g(t) \geq 0$, Eq. (1) leads to

$$r(t) \geq R(t) \equiv \max\left\{ -\frac{x'(t)}{x(t)}, 0 \right\}. \qquad (3)$$

Note that the above inequality is always satisfied with arbitrary $g(t) \geq 0$. In other words, regardless of any specific form of a transcript profile, the protein waveform $x(t)$ imposes a stringent constraint on the protein degradation rate $r(t)$, through a lower bound $R(t)$ in Eq. (3). Therefore, a protein waveform itself can be informative about the degradation rate.

Can waveforms indicate the effect of time-specific or phase-specific degradation rates observed in empirical data? In order to address this issue, we start with a contradictory scenario that the degradation rate $r(t)$ is constant over time, i.e., $r(t) = r$, and examine its consequence. From Eq. (3),

$$r \geq r_{\min} \equiv \max_t R(t). \qquad (4)$$

Here, $r_{\min}$, the strict lower bound of the degradation rate r, is essentially determined only at a single time point $t = t_R$ with $t_R \equiv \arg\max_t R(t)$ ($0 < t_R \leq T$; throughout this work, time t in a periodic function $f(t) = f(t + T)$ is represented by a unique value within the range $0 < t \leq T$, unless specified). Because $R(t) \equiv \max\{-x'(t)/x(t), 0\}$, t_R in practice would be a point that approaches the trough of $x(t)$ after the $x(t)$'s steepest decline (t_R is placed between t_a and t_b, where $t_a \equiv \arg\max_t \{-x'(t)\}$ and $t_b \equiv \arg\min_t x(t)$, as shown in Fig. 1b). It is surprising that only such a single time point, which will be henceforth referred to as a single hotspot, plays a critical role in determining a range of the constant degradation rate r. Typically, the sharper a waveform $x(t)$ is, the larger is $r_{\min}$ at the hotspot (Fig. 1b).

For each plant clock protein, we can calculate the lower bound of its degradation rate, $r_{\min}$. Figures 2a and 3a exhibit the empirical PRR7 and PRR5 protein profiles in equal length light–dark (12L:12D) cycles[26]. Here, time points in light–dark cycles are counted from dawn (zeitgeber time). Using each

protein profile $x(t)$, we obtain $R(t)$ in Eq. (3), and then by Eq. (4), $r_{\min} \approx 0.88\,\mathrm{h}^{-1}$ for PRR7 ($t_R \approx 21\,\mathrm{h}$) and $r_{\min} \approx 1.69\,\mathrm{h}^{-1}$ for PRR5 ($t_R \approx 22.3\,\mathrm{h}$), as in Figs. 2b and 3b. It means that if the degradation rates are constant over time, the PRR7 and PRR5 half-lives at any given time points cannot be longer than ~47 and ~25 min, respectively. Provided that there are some erroneous data points in the experimental profiles, the PRR7 and PRR5 half-lives might be up to ~13 and ~51 min longer than the above, respectively (Methods). In any cases, these half-lives appear to be very short, compared to other documented protein half-lives[33,34]. As previously mentioned, such a large degradation rate over the entire course of a day is attributed to only a single hotspot $t = t_R$, under the assumption that the degradation rate is constant over time.

Next, we demonstrate that such a constant and large degradation rate can incur too large a cost of the protein production. In Eq. (2), the cost c of protein production is defined as an average protein amount synthesized per time, which is equal to an average protein amount degraded per time. For a constant degradation rate $r(t) = r$, one obtains from Eqs. (2) and (4)

$$c = r\langle x(t) \rangle \geq c_g \equiv r_{\min}\langle x(t) \rangle. \qquad (5)$$

Therefore, given the protein profile $x(t)$, the lower bound of the cost c (i.e., c_g) is directly proportional to $r_{\min}$. PRR7 or PRR5, which exhibits large $r_{\min}$, would pay an accordingly high production cost if the degradation rate is constant. More specifically, from $cT = \Delta x \geq r_{\min}T\langle x(t) \rangle$, PRR7 and PRR5 must be synthesized per day at least ~21 and ~41 times more than actual protein level $\langle x(t) \rangle$s, respectively. In other words, these protein syntheses are far excessive compared to the actual protein abundance levels.

Time-dependent degradation rates and cost reduction. The above excessive cost of protein production can be effectively alleviated by time-varying degradation rates. If the degradation rate $r(t)$ is no longer constant, $r(t)$ at $t \neq t_R$ is allowed to be smaller than $r_{\min}$, as far as Eq. (3) is satisfied. This fact leads to the possibility that the cost c can be lower than $c_g = r_{\min}\langle x(t) \rangle$. Hence, the cost can be reduced below the case of a constant degradation rate. A time-dependent degradation rate is enabled in nature by rhythmic post-translational regulation, such as PRR5 degradation by ZTL in the plant clock. Both PRR5 and ZTL levels oscillate over time, and this ZTL oscillation possibly contributes to the rhythmic degradation rate of PRR5. Including PRR5, plant clock proteins often seem to have phase-specific half-lives. These experimental data allow us to evaluate our hypothesis that rhythmic degradation rates help reduce protein production costs.

Before the calculation of the protein production costs to examine our hypothesis, we stress that all experimental degradation rates of the plant PRR7 and PRR5 proteins and of the mouse PERIOD2 (PER2) protein[28–30,35] are found to satisfy the fundamental relation $r(t) \geq R(t)$ in Eq. (3) (see Figs. 2b, 3b, and 4b). PER2 is an essential component of the mammalian clock[10,35–38], and its synthesis and turnover dynamics approximately follows Eq. (1), thereby satisfying Eq. (3). To further test the validity of Eq. (3), we performed a cycloheximide (CHX) experiment and measured the PRR7 half-life at a time point that lacks preexisting half-life data (Fig. 2c, d and Methods). Again, the PRR7 half-life at this time point ($t = 18\,\mathrm{h}$) from our own experiment is in good agreement with Eq. (3) (Fig. 2b, c). Integration of these new and previous experimental data offers a rough estimate of protein production costs, as in the following paragraphs.

Calculation of protein production cost c requires information on both degradation rate $r(t)$ and waveform $x(t)$ over time, as $c =$
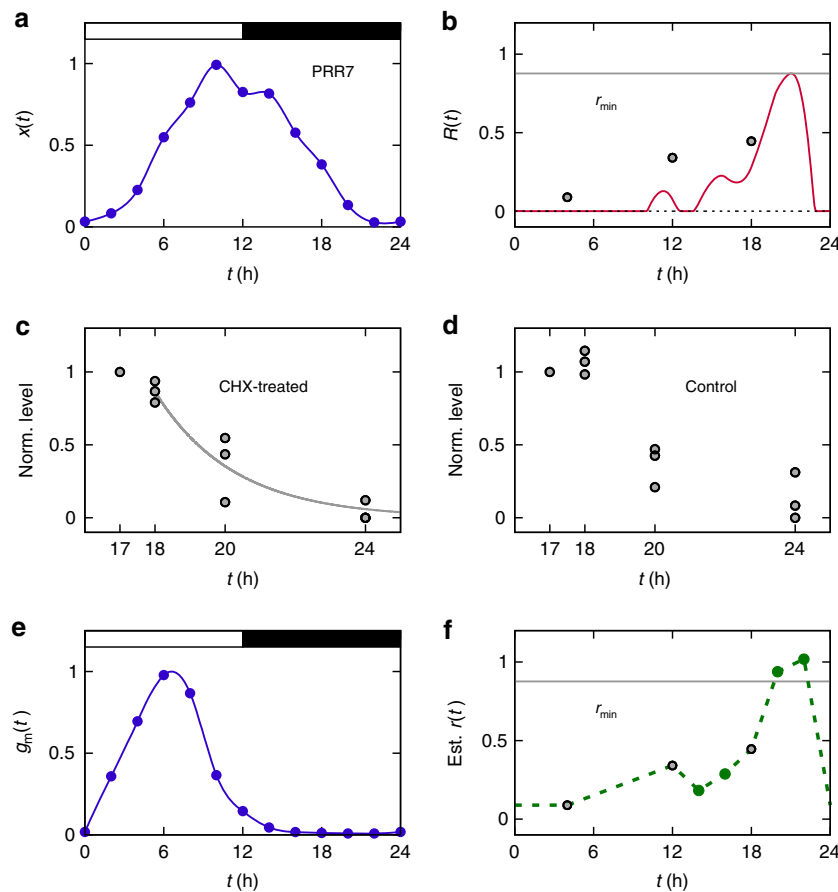
Fig. 2 PRR7 in the plant clock. For the definition of each notation, refer to Eqs. (1), (3), (4) and (18). **a** Existing experimental data of PRR7 protein levels (x (t), filled circles; normalized by the peak level of their spline curve)[28]. **b** $R(t)$ (red solid line; calculated from $x(t)$ in **a**), r_{min} (gray solid line), and experimental $r(t)$ values (circles). The vertical axis unit is h^{-1}. The value of $r(t)$ at $t = 18$ h is from our own experimental data in **c**. The rest $r(t)$ values in **b** are from previous experimental data[30]. In agreement with Eq. (3), there exists no $r(t)$ smaller than $R(t)$. **c** Our experimental measurement of PRR7 levels after CHX treatment at $t = 17$ h. **d** Similar to **c**, but without CHX treatment. In **c**, **d**, PRR7 levels are normalized to the levels at $t = 17$ h. Data points were obtained from three biological repeats. In **c**, considering a lag time for the full effect of CHX, an exponential fit (gray solid line) was made from $t = 18$ h, and then $r(t) \approx$ $0.45 \pm 0.11\,h^{-1}$ (avg. $\pm$ s.d.) at $t = 18$ h in **b** was obtained (this standard deviation of $r(t)$ does not much change the cost reduction in Table 1, because it leads to $(c_g - c)/c_g \approx 0.68$–$0.73$); an exponential fit from $t = 17$ h also supports Eq. (3) (Supplementary Fig. 2). Control PRR7 levels at and after $t = 18$ h in **d**, when averaged over three repeats at each time point and then rescaled together, are almost identical to $x(t)$ in **a**. **e** Existing experimental data of *PRR7* mRNA levels ($g_m(t)$, filled circles; normalized by the peak level of their spline curve)[27]. **f** Estimated $r(t)$ over time (green dashed line; green circles for direct calculation from experimental $r(t)$, $x(t)$, and $g_m(t)$ using Eqs. (18)–(20) with constant k), along with r_{min} in **b**. The vertical axis unit is h^{-1}. All experimental data here pertain to 12L:12D cycles, and white and black segments in **a**, **e** correspond to light and dark intervals, respectively

$\langle r(t)x(t) \rangle$ from Eq. (2). Because the degradation rate of each plant clock protein is only known for at most a few time points as presented above, we infer the rest degradation rates from those scarce experimental data. For this purpose, we interpolate and extrapolate the experimental degradation rate $r(t)$s based on the formula from Eq. (1): $r(t) \approx [g(t) - x'(t)]/x(t)$. Here, the protein synthesis rate $g(t)$ can be written as $g(t) = k(t)g_m(t)$, where $g_m(t)$ is an mRNA concentration and $k(t)$ is an mRNA-to-protein translation rate. We discard the temporal variation of $k(t)$ and take an approximation $k(t) \approx k$. Note that experimental data of both protein and mRNA profiles, $x(t)$ and $g_m(t)$, are available enough for a wide range of time in the cases of PRR7 and PRR5 (Figs. 2a, e and 3a, c). Using these data, one can estimate k and therefore the protein degradation rate every time (see Methods). Accordingly, Figs. 2f and 3d show the estimated degradation rates of PRR7 and PRR5. Alternatively, considering the time-varying nature of $k(t)$ does not much affect our main results (Methods). In addition, we estimate the degradation rate $r(t)$ of PER2. Experimental degradation rates of PER2 cover a relatively wide

range of time and are thus informative enough to envisage the overall trend of $r(t)$. Therefore, only based on these experimental degradation rates and $R(t)$, without mRNA profile data, we can make a rough estimate of the PER2 degradation rate over the entire circadian period, as demonstrated in Fig. 4c (Methods).

The estimated, phase-specific degradation rates of clock proteins in Figs. 2f, 3d and 4c show the characteristic curves that peak around the hotspots ($t \approx t_R$) and decline elsewhere. These patterns are the hallmarks of the rhythmic degradation rates that can reduce the protein production costs below the cases of constant degradation rates; except for the hotspots that must have large degradation rates ($\geq r_{min}$) by Eq. (3), if degradation rates remain small for most time, proteins do not have to be much synthesized to compensate for the degradation (Eq. (2)) and hence the production costs will become reduced.

Using the above degradation rate curves of several clock proteins, we now compute the actual protein production cost c by $c = \langle r(t)x(t) \rangle$ in Eq. (2). Compared to the cases of constant degradation rates, the PRR7, PRR5, and PER2 production costs
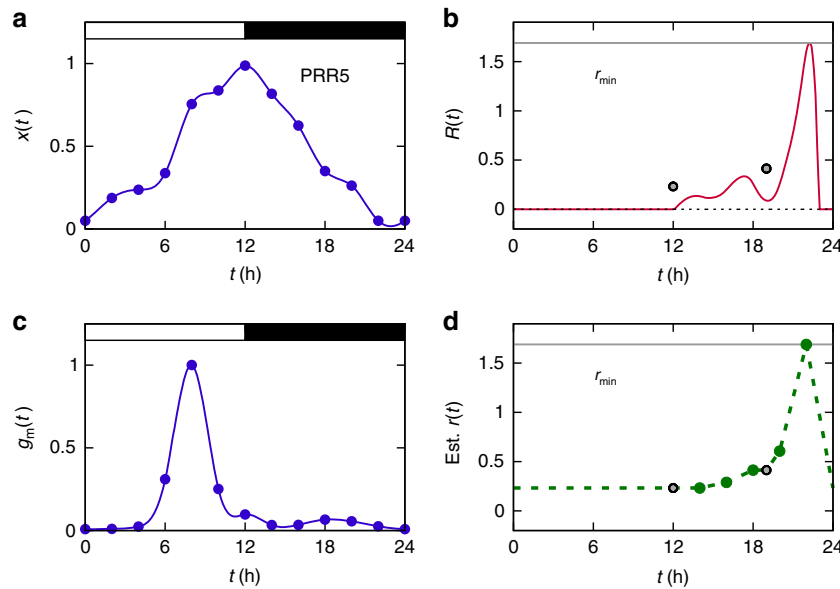
Fig. 3 PRR5 in the plant clock. For the definition of each notation, refer to Eqs. (1), (3), (4) and (18). **a** Existing experimental data of PRR5 protein levels (x(t), filled circles; normalized by the peak level of their spline curve)[26]. **b** $R(t)$ (red solid line; calculated from $x(t)$ in **a**), r_{min} (gray solid line), and empirical r(t) values (circles). The vertical axis unit is h^{-1}. The $r(t)$ values are from previous experimental data[29, 30]. In agreement with Eq. (3), there exists no $r(t)$ smaller than $R(t)$. **c** Existing experimental data of *PRR5* mRNA levels ($g_m(t)$, filled circles; normalized by the peak level of their spline curve)[27]. **d** Estimated r(t) over time (green dashed line; green circles for direct calculation from experimental $r(t)$, $x(t)$, and $g_m(t)$ using Eqs. (18)–(20) with constant k), along with r_{min} in **b**. The vertical axis unit is h^{-1}. All experimental data here pertain to 12L:12D cycles, except for $r(t)$ at $t = 19$ h in **b**, which was collected from a different light condition due to the scarcity of experimental data (Supplementary Methods). White and black segments in **a**, **c** correspond to light and dark intervals, respectively
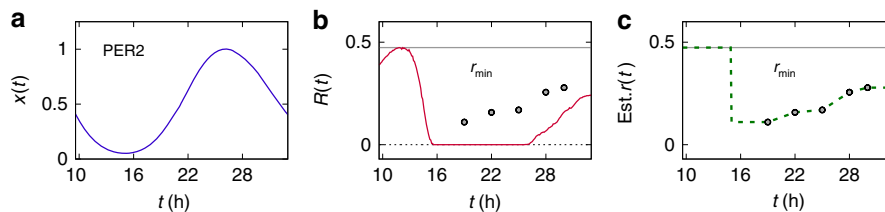

Fig. 4 PER2 in the mammalian clock. For the definition of each notation, refer to Eqs. (1), (3) and (4). **a** Existing experimental data of PER2 protein levels ($x(t)$, normalized by the peak level; moving window average of experimental data)[35]. **b** $R(t)$ (red solid line; calculated from $x(t)$ in **a**), r_{min} (gray solid line), and empirical $r(t)$ values (circles). The vertical axis unit is h^{-1}. The $r(t)$ values are from previous experimental data[35]. In agreement with Eq. (3), there exists no $r(t)$ smaller than $R(t)$. **c** Estimated $r(t)$ over time (green dashed line; green circles for experimental $r(t)$ data in **b**), along with r_{min} in **b**. The vertical axis unit is h^{-1}

indeed decrease by at least ~70%, ~83%, and ~52%, respectively, as summarized in Table 1. If we consider a possible deviation of the degradation rate in Fig. 2c, the PRR7 production cost decreases by ~68% to ~73% (Fig. 2). Interestingly, in the case of alga *Ostreococcus tauri*, rhythmic protein degradation is known to be very crucial for circadian timekeeping[39]. For its clock proteins CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) and TIMING OF CAB EXPRESSION 1 (TOC1), the full time series of experimental $r(t)$ is available[39], and our analysis suggests that the rhythmic $r(t)$ saves ~30% and ~41% of the CCA1 and TOC1 production costs, respectively (Methods). These results well support our hypothesis that rhythmic control of clock protein half-lives is beneficial to the cost reduction of protein production. This cost saving effect would be valid even if other benefits from the rhythmic half-lives are not clear. We thus predict the statistical tendency that the sharper a waveform is at the hotspot (i.e., the larger is r_{min}, and therefore is c_g), the more likely a protein half-life is to be phase-specific.

Table 1 Estimated values of r_{min}, $c_g = r_{min}\langle x(t)\rangle$, and $c = \langle r(t)x(t)\rangle$ as well as cost reduction for PRR7, PRR5, and PER2

Protein	r_{min} (h^{-1})	c_g (h^{-1})	c (h^{-1})	Cost reduction (%)
PRR7	0.88	0.40	0.12	~70
PRR5	1.69	0.77	0.13	~83
PER2	0.47	0.23	0.11	~52

For the definitions of c_g and c, refer to Eqs. (2) and (5). The cost reduction due to the time-specific or phase-specific $r(t)$ is defined as $(c_g - c)/c_g$. We here assume constant k in Eq. (18). We treat $x(t)$ as dimensionless through the normalization of $x(t)$ by its peak value (Figs. 2a, 3a and 4a), and thus units of r_{min}, c_g, and c in the table are h^{-1}. The cost reduction itself is not a quantity affected by the normalization of $x(t)$, and hence there is no loss of generality in its values

Enigmatic elements of animal circadian systems. Thus far, we have investigated circadian dynamics driven by protein synthesis and degradation in Eq. (1). We now discuss another class of circadian dynamics with Eq. (8) below, which is crucial for
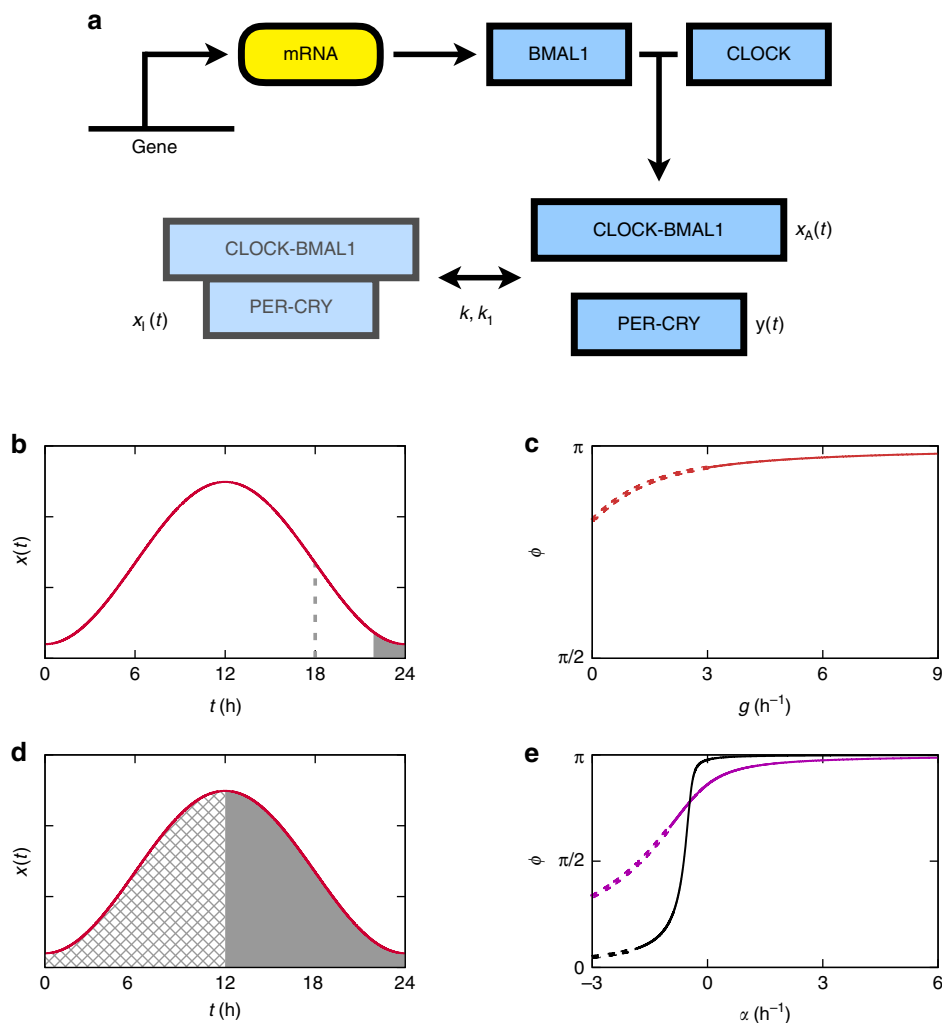
Fig. 5 Protein activity dynamics of the mammalian circadian system. **a** CLOCK and BMAL1 proteins, as well as PER and CRY proteins, heterodimerize with each other. The activities of CLOCK–BMAL1 complexes are inhibited by PER–CRY complexes through their interactions in the nucleus. For the definition of each notation, refer to Eq. (8) and its predecessor equations. **b–e** Possible phase relationship between active CLOCK–BMAL1, and PER–CRY that is not binding to CLOCK–BMAL1. **b** Regarding Eq. (9), when $g_A(t)$ is constant, the shaded area corresponds to a range of $y(t)$'s peak time, i.e., t_y's range in Eq. (12). For comparison, the dashed line indicates the time of $x(t)$'s steepest decline. **c** Phase difference ϕ between $x(t)$ and $y(t)$ as a function of g, when $g_A(t) = g$ and $x(t)$ is modeled by Eq. (13) with $L = 3\,\mathrm{h}^{-1}$ and $h_0 = 2$. Dotted is an infeasible solution with $g < g_{min} = L$ (Eq. (11)). **d** Regarding Eq. (9), when $g_A(t)$ varies over time as in Eq. (14), the left and right shaded areas correspond to the ranges of $y(t)$'s peak time for $\beta\tau > 1$ and for $\beta\tau < 1$, respectively (Methods). **e** Phase difference ϕ between $x(t)$ and $y(t)$ as a function of α, when $g_A(t)$ varies over time as in Eq. (14) with $\beta = 0.5\,\mathrm{h}^{-1}$ (violet) or $\beta = 0.95\,\mathrm{h}^{-1}$ (black), and $\tau = 1\,\mathrm{h}$, and $x(t)$ is modeled by Eq. (13) with $L = 3\,\mathrm{h}^{-1}$ and $h_0 = 2$. Dotted is an infeasible solution with $\alpha < \alpha_{min}$. Full details are described in Methods

mammals and insects, but does not follow the underlying mechanism of Eq. (1).

The core part of the mammalian clock harbors a transcriptional/post-translational negative feedback loop[10,35,36], which involves transcription factors, CLOCK and BMAL1 proteins. CLOCK–BMAL1 heterodimers activate the transcription of *Per* and *Cryptochrome* (*Cry*) genes, and the encoded PER and CRY proteins form PER–CRY complexes that are translocated to the nucleus. In the nucleus, they interact with CLOCK–BMAL1 complexes to inhibit the CLOCK–BMAL1 transcriptional activities. These positive (CLOCK and BMAL1) and negative (PER and CRY) arms constitute a negative feedback loop.

In the following equations, $x_A(t)$ and $x_I(t)$ represent the concentrations of active and inactive CLOCK–BMAL1 complexes in the nucleus, respectively, and $y(t)$ represents the concentration of nuclear PER–CRY complexes that are not binding to

CLOCK–BMAL1 complexes:

$$\frac{dx_A(t)}{dt} = \tilde{\alpha}(t) + k_1 x_I(t) - ky(t)x_A(t) - r_1 x_A(t), \qquad (6)$$

$$\frac{dx_I(t)}{dt} = ky(t)x_A(t) - k_1 x_I(t) - r_2 x_I(t). \qquad (7)$$

Here, $\tilde{\alpha}(t)$ is a rate of CLOCK–BMAL1 translocation from the cytoplasm to the nucleus, k_1 and k are, respectively, dissociation and association rate constants of two complexes, CLOCK–BMAL1 and PER–CRY, and r_1 and r_2 correspond to the sums of degradation rates and the rates of translocation to the cytoplasm. Employing another variable $x_n(t) \equiv x_A(t) + x_I(t)$ for the total CLOCK–BMAL1 concentration, the upper equation can

be rewritten as

$$\frac{dx_A(t)}{dt} = g_A(t) - [r_0 + ky(t)]x_A(t), \tag{8}$$

where $g_A(t) \equiv k_1 x_n(t) + \tilde{\alpha}(t)$ and $r_0 \equiv k_1 + r_1$.

Equation (8) represents a class of circadian dynamics distinguished from our previous case, Eq. (1). Equation (8) for the core mammalian clock captures the dynamics of active CLOCK–BMAL1 complexes (i.e., CLOCK–BMAL1 that is not binding to PER–CRY) in the nucleus, as depicted in Fig. 5a, and is applied to the insect clock as well. A fundamental difference between Eqs. (1) and (8) is as follows: in Eq. (1), $g(t)$ exhibits high-amplitude oscillation as evident from the transcript profiles of many plant clock genes and mammalian *Per* genes, and hence $g(t)$ is a main driving force of $x(t)$'s oscillation. In contrast, in Eq. (8), $x_A(t)$'s oscillation is largely driven by $y(t)$'s oscillation, rather than by $g_A(t)$'s. Compared to PER2 levels ($\propto y(t)$; Fig. 4a), BMAL1 levels are only weakly oscillating over time[10], and correspondingly, $g_A(t)$ would be only weakly oscillating. In fact, cyclic BMAL1 expression is not even required for mammalian circadian rhythmicity, as the mutant with constitutive BMAL1 expression still exhibits circadian rhythms[11–13].

Given the apparently minor role of BMAL1's abundance oscillation in circadian rhythmicity, what beneficial effects on the clock might follow from this BMAL1 oscillation?

Diverse phase differences between clock components. To the above enigmatic presence of BMAL1's abundance oscillation in the mammalian clock, the effect of protein production cost c in our previous analysis is not straightforwardly relevant. Unlike $x(t)$ in Eq. (1), $x_A(t)$ involves the only active, not the total, molecules. In other words, $x_A(t)$ is mainly driven by the relatively costless, post-translational conversion of inactive to active molecular forms, devoid of severe biosynthetic cost problems in the previous analysis. Again, we suggest that the clue for the effect of the BMAL1 oscillation can be found from waveforms, especially $x_A(t)$ and $y(t)$ from the CLOCK–BMAL1 and PER–CRY complexes. As will be shown later, such BMAL1 oscillation confers at least two advantages on the circadian system: one is a wide range of a peak time difference between the two clock components, active CLOCK–BMAL1, and PER–CRY that is not binding to CLOCK–BMAL1. The other advantage is the symmetry of the waveforms of these components. For the sake of convenience, we will henceforth drop subscript As from $x_A(t)$ and $x'_A(t)$, and simply write them as $x(t)$ and $x'(t)$. Equation (8) can be rewritten as

$$ky(t) = \frac{g_A(t) - x'(t)}{x(t)} - r_0. \tag{9}$$

If we assume that $g_A(t)$ ($\propto$ BMAL1 level) is completely constant over time, i.e., $g_A(t) = g$, waveforms $y(t)$ and $x(t)$ in Eq. (9) are substantially constrained by each other. Specifically,

$$ky(t) = \frac{g - x'(t)}{x(t)} - r_0, \tag{10}$$

$$g \geq g_{\min} \equiv \max_t x'(t), \tag{11}$$

where the inequality of the lower relation comes from $x(t) \geq 0$ and $r_0 + ky(t) \geq 0$. We will use a notation $t_f \equiv \arg\max_t f(t)$ ($0 < t_f \leq T$) for any given periodic function $f(t)$ ($f(t) = f(t+T)$) when t_f is uniquely determined by $0 < t_f \leq T$. For example, t_x denotes the

peak time of $x(t)$ during $0 < t \leq T$. From Eq. (10),

$$t_{\frac{x'}{x}} \leq t_y \leq t_{\frac{1}{x}}. \tag{12}$$

Note that $t_{\frac{x'}{x}}$ is identical to the hotspot t_R. In other words, t_y is even closer to $x(t)$'s trough time than t_R. This range of t_y is illustrated in Fig. 5b. Generally, a peak time difference between $y(t)$ and $x(t)$ takes such a narrow range that $y(t)$'s peak time is almost the same as $x(t)$'s trough time. Therefore, if $g_A(t)$ stays constant over time, waveforms $x(t)$ and $y(t)$ are only allowed to have a near anti-phase relationship.

To exemplify the above point, we consider the case with a sinusoidal wave $x'(t) = L\sin(\omega t)$ where L is a constant and $\omega \equiv 2\pi/T$. In this case,

$$x(t) = -\frac{L}{\omega}\cos(\omega t) + \frac{L}{\omega} + h_0, \tag{13}$$

with an additional constant h_0. In the subsequent analyses, we treat $x(t)$ in Eq. (13) as dimensionless, without loss of generality. We define a phase difference between $x(t)$ and $y(t)$ in Eq. (10) as $\phi \equiv |\omega(t_y - t_x)|$. Using Eq. (10), $\phi = \pi - 2\tan^{-1}\left[\left(\sqrt{C^2 - L^2 + g^2} - g\right)/(C + L)\right]$, where $C \equiv h_0\omega + L$ and $g \geq g_{\min} = L$ from Eq. (11). This exact solution of ϕ is indeed very close to π, as shown in Fig. 5c. This tendency is consistent with our generic result that constant $g_A(t)$ forces $x(t)$ and $y(t)$ into a near anti-phase relationship.

In contrast, if $g_A(t)$ is no longer constant but cycles over time, as observed with cyclic BMAL1 expression in nature, then waveforms $y(t)$ and $x(t)$ in Eq. (9) are not much constrained by each other, and their peak time difference (or phase difference) can be flexible depending on $g_A(t)$'s oscillatory form. Because there is a lack of compelling experimental data on the waveform of $g_A(t)$, we start with the following assumption:

$$g_A(t) \approx \alpha + \beta x(t + \tau), \tag{14}$$

where α, β, and τ are constants, and $\beta \geq 0$ and $\tau \geq 0$. From Eq. (9), $r_0 + ky(t) \geq 0$, and $g_A(t) \geq 0$, α should satisfy

$$\alpha \geq \alpha_{\min} \equiv \max\left\{\max_t[x'(t) - \beta x(t+\tau)], \max_t[-\beta x(t)]\right\}. \tag{15}$$

Next, we show that $x(t)$ and $y(t)$ can have almost any in-phase to anti-phase relationship, covering a wide range of the phase difference. If $\tau \ll T$,

$$x(t + \tau) \approx x(t) + \tau x'(t). \tag{16}$$

Combined with Eq. (9), it leads to

$$ky(t) \approx \frac{\alpha - (1 - \beta\tau)x'(t)}{x(t)} + \beta - r_0. \tag{17}$$

Depending on signs of α and $1 - \beta\tau$ in Eq. (17), $y(t)$ is now allowed to peak anytime of a day relative to $x(t)$'s peak time, as proven in Methods.

This result is illustrated in Fig. 5d. Together, if BMAL1 level ($\propto g_A(t)$ in Eqs. (8) and (9)) is not constant but varies over time, it confers much freedom on the waveform $y(t)$ of PER–CRY that is not binding to CLOCK–BMAL1, and thus allows various phase differences between those unbinding CLOCK–BMAL1 and PER–CRY complexes ($x(t)$ and $y(t)$) through the adjustment of parameters α, β, and τ. As in Fig. 5d, the unbinding CLOCK–BMAL1 and PER–CRY complexes can take almost any in-phase to anti-phase relationship. This result is in sharp contrast to the case with constant $g_A(t)$, where the unbinding CLOCK–BMAL1 and PER–CRY complexes have a

predominantly anti-phase-like relationship. Our predictions can be verified by experimental techniques, such as co-immunoprecipitation assays, measuring the time series of CLOCK–BMAL1 and PER–CRY levels across different tissues or developmental stages, while excluding the levels of inactive CLOCK–BMAL1.

These potentially diverse phase differences, conferred by cyclic expression of positive arm components, may help in the coordination of tissue-specific or developmental stage-specific clock events in complex multicellular organisms, such as mammals and insects[20,40]. Interestingly, the fungus *Neurospora crassa*, a relatively simple species, shows almost constant levels of *white collar-1* (*wc-1*) expression[41], and thus would have almost constant $g_A(t)$. Therefore, we expect that the fungal clock may have an only anti-phase-like relationship between its core components, nuclear WC-1 and FREQUENCY (FRQ) proteins (see Supplementary Discussion).

To illustrate the above diverse phase differences conferred by BMAL1 cycling, we revisit the case with a sinusoidal wave $x(t)$ in Eq. (13) and consider the oscillation of $g_A(t)$ in Eq. (14). From Eq. (9), we obtain the exact solution of the phase difference between $x(t)$ and $y(t)$, as plotted in Fig. 5e. This exact solution is in good agreement with our generic results based on the approximation Eqs. (16) and (17).

Symmetries of waveforms. Another advantage of cyclic expression of positive arm components in mammalian and insect clocks is in the symmetry of waveforms about the peak phases. Previous experimental data from the mammalian clock indicate the existence of such symmetry that ascending and descending phases span almost the same time intervals[42], while its phenotypic significance still remains unknown.

To demonstrate the effect of BMAL1 cycling on the waveform symmetry, we first assume the contradictory scenario that $g_A(t)$ is constant over time with $g_A(t) = g$. Therefore, $ky(t) = [g - x'(t)]/x(t) - r_0$ from Eq. (10). In this case, waveforms $x(t)$ and $y(t)$ from CLOCK–BMAL1 and PER–CRY complexes cannot easily satisfy both symmetric relations $x(t_x - t) \approx x(t_x + t)$ and $y(t_y - t) \approx y(t_y + t)$ at the same time, because $x'(t)$ term in $ky(t) = [g - x'(t)]/x(t) - r_0$ breaks the symmetry of either $x(t)$ or $y(t)$ waveform unless $g \gg \max_t |x'(t)|$ to diminish the effect of $x'(t)$.

In contrast, if BMAL1 level ($\propto g_A(t)$) is not constant but varies over time, both unbinding CLOCK–BMAL1 and PER–CRY profiles ($x(t)$ and $y(t)$) are allowed to have symmetric waveforms relatively easily. For example, if $g_A(t) \approx \alpha + \beta x(t + \tau)$ and $x(t + \tau) \approx x(t) + \tau x'(t)$ with $\tau \ll T$ as in Eqs. (14) and (16), then $ky(t) \approx [\alpha - (1 - \beta\tau)x'(t)]/x(t) + \beta - r_0$ in Eq. (17). Therefore, both $x(t)$ and $y(t)$ waveforms can be approximately symmetric at the same time, as long as $|\alpha/(1 - \beta\tau)| \gg \max_t |x'(t)|$ for the diminished effect of $x'(t)$. This condition can be satisfied more easily than the previous one.

This waveform symmetry, along with the above phase difference between two core components (unbinding CLOCK–BMAL1 and PER–CRY complexes), shows that the waveforms are useful to understand the effect of the enigmatic oscillation in BMAL1 expression.

Discussion

In this study, we have revealed that protein waveforms are informative about the underlying mechanisms of circadian clockwork.

A sharp waveform at the hotspot time point (i.e., with large r_{min}) implies rhythmic post-translational regulation that yields a phase-specific protein half-life; otherwise, too large costs of protein syntheses can be incurred for those waveforms. Such

rhythmic degradation rates are observed in plant and mammalian circadian clocks, and can substantially reduce the protein production costs, as demonstrated in Table 1. If more experimental data become available, our waveform-cost analysis can be extended to other clock proteins. For example, the orphan nuclear receptor REV-ERBα in the mammalian clock may have a phase-specific half-life, driven by the rhythmic activity of glycogen synthase kinase-3β that regulates the REV-ERBα stability[43–45]. Hence, if the half-lives measured at multiple specific time points become available, REV-ERBα will be a good target candidate for our cost analysis, aided by the existing REV-ERBα expression profiles[46].

On the other hand, regarding any possible extra costs that may be incurred by rhythmic degradation rates, we note that the cost c of a given protein is not the concept to include the production cost of its proteolytic factor. Yet, the half-life can exhibit a rhythmic pattern by the proteolytic factor's oscillation, and thus one may suggest that the cost c should be extended to the proteolytic factor's production cost. This extra cost from the proteolytic factor production, however, is not always relevant and needs cautious analyses in the future. For example, if the proteolytic factor has not only evolved for the degradation of a particular protein but also for other functions, then the cost of the proteolytic factor production shall not be covered by the cost c in question. This is because such a proteolytic factor continues to be produced for multiple purposes, not exclusively for the degradation of that particular protein.

In this study, we also suggest that seemingly dispensable, cyclic expression of certain clock proteins in mammals and insects may allow both a broad range of phase differences between clock components and the symmetries of the waveforms. The various phase differences may be important for tissue-specific or developmental stage-specific clock coordination in complex multicellular organisms, such as mammals and insects. As previously mentioned, fungi do not show such cyclic expression of the corresponding components, and their relatively simple organismal forms may not necessitate as widely ranging phase differences as in the cases of mammals and insects.

Our waveform-guided approach is well supported by experimental data (Figs. 2b, 3b and 4b), and provides insights into circadian mechanisms of evolutionarily distant organisms[6–8]. Furthermore, we envisage that the concepts presented in this study can be applied beyond circadian dynamics, such as to time-course data from cell cycle systems and synthetic genetic oscillators[47–49].

Methods

Experimental measurement of the PRR7 half-life. We describe the details of our experimental methods for the measurement of the PRR7 degradation rate, of which data are available in Fig. 2c, d and Supplementary Figs. 1 and 2. For CHX assays, *PRR7pro::FLAG-PRR7-GFP* seedlings were grown on MS media with 3% sucrose and 1% agar under 12L:12D cycles (white fluorescent light; 30–40 µmol m^{-2} s^{-1}) at 22 °C for 14 days. Seedlings were transferred to MS liquid media with 100 µM CHX or mock (ethanol) at ZT17 in darkness. The tissues were kept in the dark under slow shaking and collected at 0, 1, 3, 5, and 7 h post treatment.

For immunoblots, the tissue was ground in liquid nitrogen and extracted in protein extraction buffer (50 mM Tris–Cl, pH 7.5, 150 mM NaCl, 0.5% Nonidet P-40, 1 mM EDTA, 3 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, 5 µg ml^{-1} leupeptin, 1 µg ml^{-1} aprotinin, 1 µg ml^{-1} pepstatin, 5 µg ml^{-1} antipain, 5 µg ml^{-1} chymostatin, 50 µM MG132, 50 µM MG115, 50 µM ALLN). Total proteins were separated using an 8% SDS–PAGE gel (acrylamide:bisacrylamide, 37.5:1), immunoblotted and probed with anti-GFP antibody (Abcam, ab6556) and polyclonal anti-ADK antibody (gift from Dr. David Bisaro) diluted to 1:4000 and 1:15,000, respectively, followed by anti-rabbit IgG conjugated with horseradish peroxidase (GE healthcare, NA934). Chemiluminescent detection was performed using SuperSignal^{TM} West Pico Chemiluminescent Substrate (Thermo Scientific, 34080). The FLAG-PRR7-GFP protein signals were calculated by ImageJ software (NIH, version 1.8.0) from three biological repeats, and were normalized to their corresponding ADK (adenosine kinase) signal intensities individually.

All unique biological materials used in this study (*PRR7pro::FLAG-PRR7-GFP*) are available from the authors upon request.

Analysis of data from the plant and mammalian clocks. By writing the protein synthesis rate $g(t)$ as $g(t) = k(t)g_m(t)$ and by assuming the roughly constant $k(t)$, i.e., $k(t) \approx k$, Eq. (1) can be written as

$$\frac{dx(t)}{dt} \approx kg_m(t) - r(t)x(t), \qquad (18)$$

which leads to

$$k \approx \frac{x'(t_i) + r(t_i)x(t_i)}{g_m(t_i)}, \qquad (19)$$

where t_i corresponds to each time point $t = t_i$ with experimentally available degradation rate $r(t)$. In the case of PRR7, we used the protein degradation rates at $t_i = 4$, 12, and 18 h (Fig. 2 and Supplementary Fig. 3). The first two degradation rates were obtained from the protein abundance data in Fig. 7b of Farre et al.[28], while the last degradation rate was from our own experimental data in Fig. 2c. We also obtained the experimental data of the mRNA and protein profiles from Fig. 5d of Flis et al.[27] and Fig. 5a of Nakamichi et al.[26], respectively. Both datasets have 2-h sampling intervals under 12L:12D cycles. These mRNA and protein levels were normalized by the peak levels of their splines, and adopted for $g_m(t)$ and $x(t)$ in Eq. (18), respectively. From Eq. (5), $c_g \approx 0.40\,h^{-1}$. Using $r(t_i)$, $x(t_i)$, and $g_m(t_i)$, we obtained k from Eq. (19). To be precise, although we treat k as a constant, different t_is can have different k values calculated from Eq. (19). For simplicity of our analysis, we discarded such differences and took the average of k over t_i. Using this k, we inferred $r(t)$ for the rest of time ($t \neq t_i$) by the following formula from Eq. (18):

$$r(t) \approx \frac{kg_m(t) - x'(t)}{x(t)}. \qquad (20)$$

Because experimental protein and mRNA levels have 2-h sampling intervals, we inferred degradation rate $r(t)$ every 2 h, except for $t = 4$, 12, and 18 h for which we used experimentally known $r(t)$ values. The overall $r(t)$ profile exhibits two peaks at $20\,h \leq t \leq 22\,h$ and at $2\,h \leq t \leq 10\,h$. The former peak is a natural consequence of large $R(t)$ around that time (red solid line in Supplementary Fig. 3a), while the latter may be an artifact from unconsidered biological factors. To reduce the effect of such possible artifact, we replace every $r(t) > \max_{20h \leq t \leq 22h} r(t)$ by $\max_{20h \leq t \leq 22h} r(t)$, because $\max_{20h \leq t \leq 22h} r(t) \approx 1.02\,h^{-1}$ and the real degradation rate is unlikely to be larger than $1.02\,h^{-1}$. We also replace every $r(t) < \min\{r(t = 4\,h), r(t = 12\,h), r(t = 18\,h)\}$ by $\min\{r(t = 4\,h), r(t = 12\,h), r(t = 18\,h)\}$, and therefore the lower bound of $r(t)$ is set to the minimum value of experimental $r(t)$ values. In such a way, the difference between c and c_g is reduced (Eqs. (2) and (5)), leading to a conservative estimate of that difference. The resulting $r(t)$ is presented in Supplementary Fig. 3a. Because $r(t)$ at $2\,h \leq t \leq 10\,h$ is improbably deviated from the overall trend of experimental $r(t)$ values, we correct this part by linear interpolation and extrapolation of the experimental $r(t = 4\,h)$ and $r(t = 12\,h)$ values, as shown in Fig. 2f. Consequently, $c \approx 0.30c_g$ with $r(t)$ in Fig. 2f and $c \approx 0.67c_g$ with $r(t)$ in Supplementary Fig. 3a. In other words, whether correcting $r(t)$ at $2\,h \leq t \leq 10\,h$ or not, the actual cost of PRR7 waveform maintenance would be at most one-third to two-thirds of the assumed cost in the case of a constant degradation rate.

Thus far, we have adopted the experimental protein levels for $x(t)$. However, we suppose that experimental protein levels, when low around a trough phase, can be susceptible to measurement errors. Such potentially inaccurate data, if these data underestimate the protein levels around the trough phase, can lead to the overestimation of r_{min} in Eq. (4) and c_g in Eq. (5), and thereby exaggerate a difference between c_g and c. To mitigate these possibly erroneous effects, we consider a new $x(t)$ whose values at $t = 0$, 22, and 24 h are replaced by that of $x(t = 2\,h)$, as plotted in Supplementary Fig. 3b. With this smoothed $x(t)$, we obtain $r_{min} \approx 0.69\,h^{-1}$, which is smaller than $r_{min} \approx 0.88\,h^{-1}$ from the original $x(t)$. Likewise, new $c_g \approx 0.32\,h^{-1}$ and $c \approx 0.12\,h^{-1}$. Here, c is calculated from the newly estimated $r(t)$ in Supplementary Fig. 3c. On the other hand, without a correction for $2\,h \leq t \leq 10\,h$ as in Supplementary Fig. 3d, $c \approx 0.22\,h^{-1}$. Still, the cost of PRR7 waveform maintenance is at most one-third to two-thirds of the assumed cost in the case of a constant degradation rate. These results are summarized in Supplementary Table 1.

In the case of PRR5, we used experimental protein degradation rates at $t_i = 12$ and 19 h (Fig. 3 and Supplementary Fig. 4) from the protein abundance data in Fig. 7c of Baudry et al.[29]. We obtained the experimental data of the mRNA and protein profiles from Fig. 5d of Flis et al.[27] and Fig. 5a of Nakamichi et al.[26], respectively. Both datasets have 2-h sampling intervals under 12L:12D cycles. These mRNA and protein levels were normalized by the peak levels of their splines, and adopted for $g_m(t)$ and $x(t)$ in Eq. (18), respectively. Following a similar procedure to the case with PRR7, we obtained $c_g \approx 0.77\,h^{-1}$, and inferred the degradation rate $r(t)$ every 2 h, except for $t = 12$ and 19 h for which we used experimentally known $r(t)$ values. When calculating c based on this inferred $r(t)$, we replace every $r(t) > r_{min}$ by r_{min}, because the real degradation rate is unlikely to be larger than $r_{min} \approx 1.69\,h^{-1}$. We also replace every $r(t) < \min\{r(t = 12\,h), r(t = 19\,h)\}$ by $\min\{r(t = 12\,h), r(t = 19$

h)}, and therefore the lower bound of $r(t)$ is set to the minimum value of experimental $r(t)$ values. In such a way, the difference between c and c_g is reduced, leading to a conservative estimate of that difference. The resulting $r(t)$ is presented in Supplementary Fig. 4a. Because $r(t)$ at $6\,h \leq t \leq 10\,h$ is improbably deviated from the overall trend of experimental $r(t)$ values, we correct this part by linear extrapolation of the experimental $r(t = 12\,h)$ value, as shown in Fig. 3d. Consequently, $c \approx 0.17c_g$ with $r(t)$ in Fig. 3d and $c \approx 0.34c_g$ with $r(t)$ in Supplementary Fig. 4a. In other words, whether correcting $r(t)$ at $6\,h \leq t \leq 10\,h$ or not, the actual cost of PRR5 waveform maintenance would be at most one-sixth to one-third of the assumed cost in the case of a constant degradation rate.

To mitigate the aforementioned, possibly erroneous effects from low protein levels around a trough phase, we consider new $x(t)$ whose values at $t = 0$, 22, and 24 h are increased as in Supplementary Fig. 4b. With this smoothed $x(t)$, we obtain $r_{min} \approx 0.55\,h^{-1}$, which is smaller than $r_{min} \approx 1.69\,h^{-1}$ from the original $x(t)$. Likewise, new $c_g \approx 0.26\,h^{-1}$ and $c \approx 0.13\,h^{-1}$. Here, c is calculated from the newly estimated $r(t)$ in Supplementary Fig. 4c. On the other hand, without a correction for $6\,h \leq t \leq 10\,h$ as in Supplementary Fig. 4d, $c \approx 0.17\,h^{-1}$. Still, the cost of PRR5 waveform maintenance is at most one-half to two-thirds of the assumed cost in the case of a constant degradation rate. These results are summarized in Supplementary Table 1.

For PRR7 and PRR5, the time-varying nature of $k(t)$ in $g(t) = k(t)g_m(t)$ can be considered as an alternative to the above possibility $k(t) \approx k$. Because of a lack of data on the genuine form of $k(t)$ for these proteins, we tried a sinusoidal approximation $k(t) \approx \max\{a\sin(2\pi t/T - \phi) + b, \epsilon_k\}$, where a, b, and ϕ are constants that fit the function $a\sin(2\pi t/T - \phi) + b$ to the right-hand side of Eq. (19) and ϵ_k is a small positive constant to ensure $k(t) > 0$ (ϵ_k was set to the minimum value of the right-hand side of Eq. (19)). In the PRR5 case, a, b, and ϕ were underdetermined, and thus a and b were obtained for each value of ϕ in the range $0 \leq \phi \leq \pi/2$ (which does not involve any loss of generality for the PRR5 data). Applying such $k(t)$ to Eq. (20), instead of k therein, and repeating all the above procedures (Supplementary Fig. 5) did not much change our results: under this assumption of the time-varying $k(t)$, the estimated rhythmic degradation rates led to ~73% reduction of the PRR7 production cost and ~83–84% reduction of the PRR5 production cost (cf. Table 1 for the case $k(t) \approx k$).

In the case of the mouse PER2 protein, we obtained the time-course abundance data of the CHX-untreated control in Fig. 1a of Zhou et al.[35], and adopted this protein profile for $x(t)$. The original profile covers ~45-h-long data with 0.1-h resolution. Therefore, we considered the data at $9.6\,h \leq t \leq 33\,h$ for one circadian period ($T = 23.4\,h$), and smoothed them with a moving window average (3-h window). These data were normalized by their peak level, and the resulting $x(t)$ appears in Fig. 4a. $R(t)$ derived from this $x(t)$ is very noisy, and therefore smoothed with a moving window average (1-h window). For experimental protein degradation rates, we used the instantaneous half-lives after 0.5 h since CHX treatment at $t = 19$, 22, 25, 28, and 30 h in Supplementary Fig. 1a of Zhou et al.[35].

Full details of the PRR7, PRR5, and PER2 data collection are provided in Supplementary Methods and Supplementary Table 1.

Analysis of data from the algal clock. In the case of CCA1 and TOC1 proteins in the *Ostreococcus* circadian system, we obtained the full time-course degradation rate $r(t)$ and protein level $x(t)$ data from Fig. 1a, b of van Ooijen et al.[39], respectively (12L:12D-cycle condition). We did not perform any normalization of $x(t)$, and the unit of $x(t)$ here follows that of van Ooijen et al.[39] (molecules cell^{-1}). Because $x(t)$'s sampling resolution was rather low (4-h sampling interval), we did not apply $r(t)$ and $x(t)$ to Eq. (3) wherein the specific form of $R(t)$ could be sensitive to the $x(t)$'s sampling resolution. For the calculation of c_g, we estimated r_{min} as $r_{min} \approx \min\{\max_t r(t), \max_t R(t)\}$, with regards to possibly inaccurate $R(t)$ from the low sampling resolution of $x(t)$. For the calculation of c, we adopted $r(t)x(t)$ in Fig. 1c of van Ooijen et al.[39]. As a result, for CCA1 and TOC1, $r_{min} \approx 0.25$ and 0.28 h^{-1}, $c_g \approx 60.7$ and 19.7 molecules cell^{-1} h^{-1}, and $c \approx 42.5$ and 11.6 molecules cell^{-1} h^{-1}, respectively. In other words, the cost of CCA1 and TOC1 production is about two-thirds of the assumed cost in the case of constant degradation rates.

Effects of oscillating BMAL1 expression. If $\tau \ll T$ in Eqs. (14) and (17) can be used to calculate a phase difference between $x(t)$ and $y(t)$. Without loss of generality, let $x(t)$ be the lowest at $t = T$, i.e., $t_1 = T$. Depending on signs of α and $1 - \beta\tau$ in Eq. (17), we consider the following four cases:

1. If $\alpha > 0$ and $\beta\tau < 1$, $y(t)$ in Eq. (17) is described essentially in the same way as Eq. (10), while extra constants in Eq. (17) do not affect the way to determine a phase difference between $x(t)$ and $y(t)$. Therefore, t_y still follows

$$t_{\frac{x'}{x}} \leq t_y \leq t_{\frac{1}{x}} = T, \qquad (21)$$

and the phase difference between $x(t)$ and $y(t)$ is determined in a similar way to the case with constant $g_A(t)$ (i.e., $g_A(t) = g$). $y(t)$ in this case will be called $y_1(t)$.

2. If $\alpha > 0$ and $\beta\tau > 1$, $y(t)$ is determined in a similar way to $y_1(t)$, but with the flipped sign of $x'(t)$. Therefore,

$$0 \leq t_y \leq t_{\frac{x'}{x}}. \qquad (22)$$

$y(t)$ in this case will be called $y_2(t)$.

3. If $\alpha < 0$ and $\beta\tau < 1$, $y(t)$ is described in a similar way to $-y_2(t)$. Therefore,

$$t_x \leq t_y \leq t_{-\frac{x'}{x}}. \qquad (23)$$

4. If $\alpha < 0$ and $\beta\tau > 1$, $y(t)$ is described in a similar way to $-y_1(t)$. Therefore,

$$t_{\frac{x'}{x}} \leq t_y \leq t_x. \qquad (24)$$

In addition, both $x(t)$ and $y(t)$ can have symmetric waveforms as long as $|\alpha/(1 - \beta\tau)| \gg \max_t |x'(t)|$ (for example, this condition can be satisfied when $\beta\tau \approx 1$).

Besides the case of Eq. (14) with $\tau \ll T$, we analyze the case with $\tau \sim T/2$. In this case, $\tau = T/2 + \epsilon$ with $|\epsilon| \ll T$, and $x(t + \tau)$ in Eq. (14) can be approximated as

$$x\left(t + \frac{T}{2} + \epsilon\right) \approx x\left(t + \frac{T}{2}\right) + \epsilon x'\left(t + \frac{T}{2}\right). \qquad (25)$$

We further assume the waveform that $x(t + T/2) \approx J - x(t)$, where J is a constant satisfying $J \approx (2/T)\int_0^T x(t)\mathrm{d}t$. From Eq. (9),

$$ky(t) \approx \frac{\alpha + \beta J - (1 + \beta\epsilon)x'(t)}{x(t)} - (r_0 + \beta). \qquad (26)$$

This equation takes a similar form to Eq. (17). By dividing four different categories of $y(t)$ depending on signs of $\alpha + \beta J$ and $1 + \beta\epsilon$, it is straightforward to obtain similar results to our previous analysis of a phase difference between $x(t)$ and $y(t)$ when $\tau \ll T$.

In addition, both $x(t)$ and $y(t)$ can have symmetric waveforms as long as $|(\alpha + \beta J)/(1 + \beta\epsilon)| \gg \max_t |x'(t)|$ (for example, this condition can be satisfied when $\beta\epsilon \approx -1$).

To illustrate the diverse phase differences conferred by BMAL1 cycling, we study the case with a sinusoidal wave $x(t)$ in Eq. (13) and consider the oscillation of $g_A(t)$ in Eq. (14). From Eq. (9), t_y is obtained as

$$\omega t_y = 2\pi n + 2\tan^{-1}\left\{ \frac{\omega\alpha + C\beta[1 - \cos(\omega\tau)]}{(C+L)[\omega - \beta\sin(\omega\tau)]} \right.$$
$$\left. - \frac{\sqrt{(\omega\alpha + C\beta[1 - \cos(\omega\tau)])^2 + (C^2 - L^2)[\omega - \beta\sin(\omega\tau)]^2}}{(C+L)[\omega - \beta\sin(\omega\tau)]} \right\} \qquad (27)$$

with $C = h_0\omega + L$ and an integer n that satisfies $0 < \omega t_y \leq 2\pi$. From Eq. (15), α satisfies

$$\alpha \geq \alpha_{\min} = \max\left\{ L\sqrt{1 - \frac{2\beta}{\omega}\sin(\omega\tau) + \frac{\beta^2}{\omega^2}} - \frac{\beta L}{\omega}, 0 \right\} - \beta h_0. \qquad (28)$$

Equation (27) and $t_x = T/2$ give rise to the exact solution of the phase difference ϕ between $x(t)$ and $y(t)$ ($\phi = |\omega(t_x - t_y)|$), as plotted in Fig. 5e. This exact solution is in good agreement with our generic results based on the approximation Eqs. (16) and (17).

Code availability. Source codes for analyzing data in the manuscript have been deposited into the public repositories GitHub and Zenodo.

Data availability
All relevant data are available in Methods, Figs. 2–4, Supplementary Methods, Supplementary Figs. 1–5, and Supplementary Data 1.

References
1. Nagel, D. H. & Kay, S. A. Complexity in the wiring and regulation of plant circadian networks. *Curr. Biol.* **22**, R648–R657 (2012).
2. Sehgal, A., Price, J. L., Man, B. & Young, M. W. Loss of circadian behavioral rhythms and *per* RNA oscillations in the *Drosophila* mutant *timeless*. *Science* **263**, 1603–1606 (1994).
3. Brody, S. & Harris, S. Circadian rhythms in *Neurospora*: spatial differences in pyridine nucleotide levels. *Science* **180**, 498–500 (1973).
4. Gachon, F., Nagoshi, E., Brown, S., Ripperger, J. & Schibler, U. The mammalian circadian timing system: from gene expression to physiology. *Chromosoma* **113**, 103–112 (2004).
5. Fei, C., Cao, Y., Ouyang, Q. & Tu, Y. Design principles for enhancing phase sensitivity and suppressing phase fluctuations simultaneously in biochemical oscillatory systems. *Nat. Commun.* **9**, 1434 (2018).
6. Hsu, P. Y. & Harmer, S. L. Wheels within wheels: the plant circadian system. *Trends Plant Sci.* **19**, 240–249 (2014).
7. Hurley, J., Loros, J. J. & Dunlap, J. C. Dissecting the mechanisms of the clock in Neurospora. *Methods Enzymol.* **551**, 29–52 (2015).
8. Mendoza-Viveros, L. et al. Molecular modulators of the circadian clock: lessons from flies and mice. *Cell Mol. Life Sci.* **74**, 1035–1059 (2017).
9. Sancar, C., Sancar, G., Ha, N., Cesbron, F. & Brunner, M. Dawn- and dusk-phased circadian transcription rhythms coordinate anabolic and catabolic functions in *Neurospora*. *BMC Biol.* **13**, 17 (2015).
10. Lee, C., Etchegaray, J. P., Cagampang, F. R., Loudon, A. S. & Reppert, S. M. Posttranslational mechanisms regulate the mammalian circadian clock. *Cell* **107**, 855–867 (2001).
11. Liu, A. C. et al. Redundant function of REV-ERBα and β and non-essential role for Bmal1 cycling in transcriptional regulation of intracellular circadian rhythms. *PLoS Genet.* **4**, e1000023 (2008).
12. Preitner, N. et al. The orphan nuclear receptor REV-ERBα controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell* **110**, 251–260 (2002).
13. McDearmon, E. L. et al. Dissecting the functions of the mammalian clock protein BMAL1 by tissue-specific rescue in mice. *Science* **314**, 1304–1308 (2006).
14. Kiba, T., Henriques, R., Sakakibara, H. & Chua, N.-H. Targeted degradation of PSEUDO-RESPONSE REGULATOR5 by an SCF^{ZTL} complex regulates clock function and photomorphogenesis in *Arabidopsis thaliana*. *Plant Cell* **19**, 2516–2530 (2007).
15. Fujiwara, S. et al. Post-translational regulation of the *Arabidopsis* circadian clock through selective proteolysis and phosphorylation of pseudo-response regulator proteins. *J. Biol. Chem.* **283**, 23073–23083 (2008).
16. Foo, M., Somers, D. E. & Kim, P.-J. Kernel architecture of the genetic circuitry of the *Arabidopsis* circadian system. *PLoS Comput. Biol.* **12**, e1004748 (2016).
17. Hatakeyama, T. S. & Kaneko, K. Reciprocity between robustness of period and plasticity of phase in biological clocks. *Phys. Rev. Lett.* **115**, 218101 (2015).
18. Lück, S., Thurley, K., Thaben, P. F. & Westermark, P. O. Rhythmic degradation explains and unifies circadian transcriptome and proteome data. *Cell Rep.* **9**, 741–751 (2014).
19. Patton, A. P., Chesham, J. E. & Hastings, M. H. Combined pharmacological and genetic manipulations unlock unprecedented temporal elasticity and reveal phase-specific modulation of the molecular circadian clock of the mouse suprachiasmatic nucleus. *J. Neurosci.* **36**, 9326–9341 (2016).
20. Korenčič, A. et al. Timing of circadian genes in mammalian tissues. *Sci. Rep.* **4**, 5782 (2014).
21. Yoo, S.-H. et al. Competing E3 ubiquitin ligases govern circadian periodicity by degradation of CRY in nucleus and cytoplasm. *Cell* **152**, 1091–1105 (2013).
22. Kim, J. K. & Forger, D. B. On the existence and uniqueness of biological clock models matching experimental data. *SIAM J. Appl. Math.* **72**, 1842–1855 (2012).
23. de Montaigu, A. et al. Natural diversity in daily rhythms of gene expression contributes to phenotypic variation. *Proc. Natl Acad. Sci. USA* **112**, 905–910 (2015).
24. Scialdone, A. et al. *Arabidopsis* plants perform arithmetic division to prevent starvation at night. *eLife* **2**, e00669 (2013).
25. Cole, S. R. & Voytek, B. Brain oscillations and the importance of waveform shape. *Trends Cogn. Sci.* **21**, 137–149 (2017).
26. Nakamichi, N. et al. PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the *Arabidopsis* circadian clock. *Plant Cell* **22**, 594–605 (2010).
27. Flis, A. et al. Defining the robust behaviour of the plant clock gene circuit with absolute RNA timeseries and open infrastructure. *Open Biol.* **5**, 150042 (2015).
28. Farré, E. M. & Kay, S. A. PRR7 protein levels are regulated by light and the circadian clock in *Arabidopsis*. *Plant J.* **52**, 548–560 (2007).
29. Baudry, A. et al. F-Box proteins FKF1 and LKP2 act in concert with ZEITLUPE to control *Arabidopsis* clock progression. *Plant Cell* **22**, 606–622 (2010).
30. Wang, L., Fujiwara, S. & Somers, D. E. PRR5 regulates phosphorylation, nuclear import and subnuclear localization of TOC1 in the *Arabidopsis* circadian clock. *EMBO J.* **29**, 1903–1915 (2010).
31. Kim, W.-Y., Geng, R. & Somers, D. E. Circadian phase-specific degradation of the F-box protein ZTL is mediated by the proteasome. *Proc. Natl Acad. Sci. USA* **100**, 4933–4938 (2003).

32. Más, P., Kim, W.-Y., Somers, D. E. & Kay, S. A. Targeted degradation of TOC1 by ZTL modulates circadian function in *Arabidopsis thaliana*. *Nature* **426**, 567–570 (2003).

33. Schwanhäusser, B. et al. Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011).

34. Christiano, R., Nagaraj, N., Fröhlich, F. & Walther, T. C. Global proteome turnover analyses of the yeasts *S. cerevisiae* and *S. pombe*. *Cell Rep.* **9**, 1959–1965 (2014).

35. Zhou, M., Kim, J. K., Eng, G. W., Forger, D. B. & Virshup, D. M. A Period2 phosphoswitch regulates and temperature compensates circadian period. *Mol. Cell* **60**, 77–88 (2015).

36. Leloup, J.-C. & Goldbeter, A. Toward a detailed computational model for the mammalian circadian clock. *Proc. Natl Acad. Sci. USA* **100**, 7051–7056 (2003).

37. Kim, J. K. & Forger, D. B. A mechanism for robust circadian timekeeping via stoichiometric balance. *Mol. Syst. Biol.* **8**, 630 (2012).

38. Vanselow, K. et al. Differential effects of PER2 phosphorylation: molecular basis for the human familial advanced sleep phase syndrome (FASPS). *Genes Dev.* **20**, 2660–2672 (2006).

39. van Ooijen, G., Dixon, L. E., Troein, C. & Millar, A. J. Proteasome function is required for biological timing throughout the twenty-four hour cycle. *Curr. Biol.* **21**, 869–875 (2011).

40. Yagita, K. et al. Development of the circadian oscillator during differentiation of mouse embryonic stem cells in vitro. *Proc. Natl Acad. Sci. USA* **107**, 3846–3851 (2010).

41. Merrow, M. et al. Circadian regulation of the light input pathway in *Neurospora crassa*. *EMBO J.* **20**, 307–315 (2001).

42. Meng, Q.-J. et al. Setting clock speed in mammals: The CK1ε *tau* mutation in mice accelerates circadian pacemakers by selectively destabilizing PERIOD proteins. *Neuron* **58**, 78–88 (2008).

43. Yin, L., Wang, J., Klein, P. S. & Lazar, M. A. Nuclear receptor Rev-erbα is a critical lithium-sensitive component of the circadian clock. *Science* **311**, 1002–1005 (2006).

44. Iitaka, C., Miyazaki, K., Akaike, T. & Ishida, N. A role for glycogen synthase kinase-3β in the mammalian circadian clock. *J. Biol. Chem.* **280**, 29397–29402 (2005).

45. Besing, R. C. et al. Circadian rhythmicity of active GSK3 isoforms modulates molecular clock gene rhythms in the suprachiasmatic nucleus. *J. Biol. Rhythms* **30**, 155–160 (2015).

46. Narumi, R. et al. Mass spectrometry-based absolute quantification reveals rhythmic variation of mouse circadian clock proteins. *Proc. Natl Acad. Sci. USA* **113**, E3461–E3467 (2016).

47. Chen, Y., Kim, J. K., Hirning, A. J., Josić, K. & Bennett, M. R. Emergent genetic oscillations in a synthetic microbial consortium. *Science* **349**, 986–989 (2015).

48. Tyson, J. J. & Novák, B. Models in biology: lessons from modeling regulation of the eukaryotic cell cycle. *BMC Biol.* **13**, 46 (2015).

49. O'Keeffe, K. P., Hong, H. & Strogatz, S. H. Oscillators that sync and swarm. *Nat. Commun.* **8**, 1504 (2017).

Author contributions

H.-H.J., Y.J.K., J.K.K., D.E.S., and P.-J.K. designed the research. H.-H.J., Y.J.K., M.F., and P.-J.K. performed the research. H.-H.J., Y.J.K., J.K.K., D.E.S., and P.-J.K. analyzed the data. H.-H.J., Y.J.K., J.K.K., D.E.S., and P.-J.K. wrote the manuscript.

Additional information

Supplementary information accompanies this paper at https://doi.org/10.1038/s42003-018-0217-1.