Critical role of deadenylation in regulating poly(A) rhythms and circadian gene expression

Xiangyu Yao[1,2], Shihoko Kojima[1,3], Jing Chen[1,3*]

1. Department of Biological Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061, USA
2. Genetics, Bioinformatics, and Computational Biology program, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061, USA
3. Fralin Life Sciences Institute, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061, USA

* Corresponding author. Email: chenjing@vt.edu

Abstract

The mammalian circadian clock is deeply rooted in rhythmic regulation of gene expression. Rhythmic transcriptional control mediated by the circadian transcription factors is thought to be the main driver of mammalian circadian gene expression. However, mounting evidence has demonstrated the importance of rhythmic post-transcriptional controls, and it remains unclear how the transcriptional and post-transcriptional mechanisms collectively control rhythmic gene expression. A recent study discovered rhythmicity in poly(A) tail length in mouse liver and its strong correlation with protein expression rhythms. To understand the role of rhythmic poly(A) regulation in circadian gene expression, we constructed a parsimonious model that depicts rhythmic control imposed upon basic mRNA expression and poly(A) regulation processes, including transcription, deadenylation, polyadenylation, and degradation. The model results reveal the rhythmicity in deadenylation as the strongest contributor to the rhythmicity in poly(A) tail length and the rhythmicity in the abundance of the mRNA subpopulation with long poly(A) tails (a rough proxy for mRNA translatability). In line with this finding, the model further shows that the experimentally observed distinct peak phases in the expression of deadenylases, regardless of other rhythmic controls, can robustly group the rhythmic mRNAs by their peak phases in poly(A) tail length and in abundance of the subpopulation with long poly(A) tails. This provides a potential mechanism to synchronize the phases of target gene expression regulated by the same deadenylases. Our findings highlight the critical role of rhythmic deadenylation in regulating poly(A) rhythms and circadian gene expression.

Keywords: circadian, biorhythm, mammalian, poly(A) tail, post-transcriptional regulation, gene expression, mathematical modeling

Author Summary

The biological circadian clock regulates various bodily functions such that they anticipate and respond to the day-and-night cycle. To achieve this, the circadian clock controls rhythmic gene expression, and these genes ultimately drive the rhythmicity of downstream biological processes. As a mechanism of driving circadian gene expression, rhythmic transcriptional control has attracted the central focus. However, mounting evidence has also demonstrated the importance of rhythmic post-transcriptional controls. Here we use mathematical modeling to investigate how transcriptional and post-transcriptional rhythms coordinately control rhythmic gene expression. We have particularly focused on rhythmic regulation of the length of poly(A) tail, a nearly universal feature of mRNAs that controls mRNA stability and translation. Our model reveals that the rhythmicity of deadenylation, the process that shortens the poly(A) tail, is the dominant contributor to the rhythmicity in poly(A) tail length and mRNA translatability. Particularly, the phase of deadenylation nearly overrides the other rhythmic processes in controlling the phases of poly(A) tail length and mRNA translatability. Our finding highlights the critical role of rhythmic deadenylation in circadian gene expression control.

Introduction

Rhythmic control of gene expression is a hallmark of the circadian system. The daily rhythms in biochemistry, physiology and behavior ultimately stem from rhythmic gene expression in each cell [1, 2]. In mammals, approximately 3-15% of mRNAs are rhythmically expressed with a ~24 hr period in any given tissue [3-5]. The rhythmicity originates from a cell-autonomous circadian clock machinery, which consists of a set of core clock genes interlocked by transcription-translation feedback loops [6-8]. Many core clock genes encode transcription factors and interact with their respective target enhancers to exert rhythmic transcriptional control over circadian mRNA expression [6, 9].

While rhythmic transcriptional control has been extensively studied, rhythmic control of gene expression also occurs beyond transcription [10-12]. Recent genome-wide analyses and mathematical modeling particularly highlight the role of post-transcriptional regulations in driving rhythmic mRNA expression [13-17]. Post-transcriptional regulations target various processes, such as splicing, nuclear export, cellular translocation, dormancy and degradation of RNA [18]. Many post-transcriptional processes are under circadian control; these post-transcriptional processes, in turn, affect the phase and amplitude of mRNA level [10, 19-21]. Ultimately, rhythmic transcription and post-transcriptional processes couple with each other and jointly determine the gene expression rhythm. For example, rhythmic RNA transcription and degradation jointly determine the rhythmicity in mRNA level [16]. As yet, it remains unclear how the rhythmicities in other post-transcriptional processes affect the gene expression rhythm.

One of the post-transcriptional regulations that impact rhythmic gene expression is the regulation of poly(A) tail length. The tracts of adenosines at the 3’ end of nearly all eukaryotic mRNAs are critical for controlling stability and translatability of the mRNAs [22-24]. Hundreds

of mRNAs were recently discovered to exhibit robust circadian rhythms in their poly(A) tail lengths in mouse liver [25]. Interestingly, the rhythmicity in poly(A) tail length is closely correlated with the rhythmicity in the corresponding protein level, indicating that rhythmic poly(A) regulation plays an important role in driving rhythmic protein expression [25]. Similar daily fluctuations in poly(A) tail length also occur in mouse brain [26, 27]. In addition, the amplitude of mRNA rhythmicity increases in the absence of *Nocturnin*, a deadenylase (enzyme that removes poly(A) tails from mRNAs) which is rhythmically expressed in different mouse tissues [28, 29]. These data collectively support the importance of poly(A) tail rhythmicity in regulating circadian gene expression.

In this work, we built a mathematical model that describes mRNA dynamics under the regulation of rhythmic transcription, polyadenylation, deadenylation and degradation. We used the model to systematically examine how rhythmic transcription and poly(A) tail regulation generates rhythmicities in poly(A) tail length and mRNA abundance. Our results highlight the rhythmicity in deadenylation as the strongest determinant for the rhythmicities in the poly(A) tail length and in the abundance of mRNAs with long poly(A) tails. The latter can be regarded as a rough proxy for mRNA translatability because the poly(A) tail is known to regulate mRNA translation initiation [30-34]. As a corollary of this general finding, the model further shows that the experimentally observed distinct peak phases in the expression of deadenylases [25] can robustly group the rhythmic mRNAs by their peak phases in poly(A) tail length and in abundance of long-tailed mRNAs, regardless of other rhythmic controls. This result suggests that rhythmic deadenylation can potentially synchronize the rhythmicity in target gene expression regulated by the same deadenylases. Finally, we used the model to predict factors or combination of factors (e.g., amplitudes of or phase differences between specific processes) that can explain

the different classes of rhythmic characteristics found in mRNAs with rhythmic poly(A) tail length [25].

Results

Model

In a typical RNA expression process, RNAs are first transcribed in the nucleus and acquire long poly(A) tails as a result of nuclear polyadenylation [35]. After being exported into the cytoplasm, mRNAs undergo deadenylation and are ultimately degraded [36]. Cytoplasmic polyadenylation, as another important post-transcriptional regulation, elongates the poly(A) tail to promote mRNA stability and translatability [37]. Although cytoplasmic polyadenylation is typically associated with translational control in oocyte maturation, early embryo development and synaptic plasticity [37-40], it is recently suggested to play a role in circadian gene expression in mouse liver [25]. Furthermore, the expression level of *Gld2*, a poly(A) polymerase responsible for cytoplasmic polyadenylation, exhibits circadian rhythmicity in mouse liver [25]. In light of these biological facts, in the model we incorporated polyadenylation, together with transcription, deadenylation and degradation, to capture the major processes that dynamically regulate poly(A) tail length and mRNA abundance (**Fig. 1A**). Instead of explicitly tracking the exact length of poly(A) tails, the model divides the mRNA population into a long-tailed fraction and a short-tailed fraction, which mimics the experimental condition (long-tailed >~100nt, short-tailed <~100nt, [25]). Herein we use the ratio between the abundances of long-tailed and short-tailed mRNAs as the metric for poly(A) tail length, as was done in the experimental study [25].

For the sake of simplicity, we made the following assumptions in the model based on experimental evidence. First, degradation only occurs to the short-tailed mRNAs, because the

poly(A) tail of an mRNA must be shortened to 10~15 nt before the mRNA is degraded [41-44]. Second, transcription and nuclear polyadenylation are lumped together, because transcription is followed by nuclear polyadenylation in general [45] and the poly(A) polymerases responsible for nuclear polyadenylation are not rhythmically expressed [25]. In our model, therefore, the transcription process directly leads to a long-tailed mRNA, the downstream cytoplasmic deadenylation and polyadenylation further mediate conversion between the long-tailed and short-tailed mRNAs, and degradation consumes the short-tailed mRNA. The ordinary differential equations (ODEs) that govern the temporal dynamics of long-tailed (L) and short-tailed (S) mRNAs read as Eqs. (1) and (2).

$$\text{Long-tailed mRNA:}\quad \frac{dL}{dt} = \underbrace{\kappa_{\text{trsc}}(t)}_{\text{transcription}} - \underbrace{\kappa_{\text{deA}}(t)L}_{\text{deadenylation}} + \underbrace{\kappa_{\text{polyA}}(t)S}_{\text{polyadenylation}} \tag{1}$$

$$\text{Short-tailed mRNA:}\quad \frac{dS}{dt} = \underbrace{\kappa_{\text{deA}}(t)L}_{\text{deadenylation}} - \underbrace{\kappa_{\text{polyA}}(t)S}_{\text{polyadenylation}} - \underbrace{\kappa_{\text{dgrd}}(t)S}_{\text{degradation}} \tag{2}$$

To capture the circadian rhythmicity in the four RNA metabolism processes in Eqs. (1) and (2), each reaction rate term $\kappa(t)$ is represented by a sinusoid function like Eq.(3).

$$\kappa(t) = k\left(1 + A\cos\left(\omega(t-\varphi)\right)\right) \tag{3}$$

where k denotes the mean rate, A the relative amplitude, and φ the peak phase, of the process labeled by the subscript. The angular frequency, ω, equals $2\pi/(24\text{hr})$. ω is fixed, while the other parameters vary. The subscript of a parameter indicates the process it describes (e.g., k_{deA} stands for the mean deadenylation rate).

In this work we focus on how rhythmicities in the four processes affect the rhythmicities in total mRNA abundance and poly(A) tail length, because total mRNA abundance and poly(A) tail length were quantified in the previous circadian transcriptome study [13, 25]. Furthermore, we take the rhythmicity of long-tailed mRNA abundance as a rough proxy for mRNA

translatability, because poly(A) tail facilitates translation initiation [30-34]. From each simulation result based on Eqs. (1) and (2), i.e., time trajectories $L(t)$ and $S(t)$, we extracted the peak phases, relative amplitudes and mean levels of total mRNA abundance ($L(t) + S(t)$), poly(A) length metric ($L(t)/S(t)$), and long-tailed mRNA abundance ($L(t)$) (see Methods). These quantities were subject to further analysis, as elaborated in the following Results sections. For the rest of the paper, we will refer to these quantities, e.g., the peak phase of L/S ratio, generally as the "output quantities", unless any specific quantity is referred to.

Rhythmic deadenylation is the strongest contributor to rhythmicities in poly(A) tail length and long-tailed mRNA abundance

To investigate the effects of rhythmicities in transcription, degradation, polyadenylation and deadenylation on the output quantities, we ran numeric simulations of the model (Eqs. (1) and (2)) with random parameter values for the mean rates, relative amplitudes, and phases of each process (**Table 1 and Fig. S1**). Only the mean rate of transcription was omitted, because it only affects the overall abundance of mRNAs, but not the output rhythmicity, i.e., phases and relative amplitudes of mRNA abundance and poly(A) tail length (see Supplementary Materials).

Our model results reveal that the peak phase of deadenylation is the strongest contributor to the peak phase of L/S ratio (poly(A) length metric), followed by the peak phase of polyadenylation. Specifically, the scatter plots of the simulation demonstrate a strong correlation between the peak phases of L/S ratio and deadenylation, with a 10 ± 1.5 hr lag between the two (**Fig. 2A**). The peak phase of L/S ratio also depends on the peak phase of polyadenylation, although with much weaker dependence compared to that of deadenylation (**Fig. 2A**). In contrast, the peak phase of L/S ratio depends very little on the peak phases of transcription and

degradation (**Fig. 2A**).

To systematically quantify the impacts of input parameters on output quantities, we performed variance-based sensitivity analysis using the Sobol's method [46, 47] (**Fig. 1B**, Methods). Based on simulation results from a large number of random parameter sets spanning the global parameter space (**Fig. S1, Table 1**), the Sobol's method quantifies the sensitivity of an output quantity to an input parameter in terms of how much the parameter, due to the variation in its value, contributes to the variation in the output quantity. Specifically, the sensitivity is reported as the single (S) and total (T) Sobol indices, which represent the contribution of the parameter alone and the contribution of the parameter together with its (nonlinear) interactions with the other parameters, respectively (see Methods).

The estimated Sobol indices (**Fig. 2B**) confirm the findings from the scatter plots (**Fig. 2A**). For example, among all the input parameters, the peak phase of deadenylation has the largest Sobol indices with respect to the peak phase of L/S ratio. The values of the Sobol indices indicate that variance in the peak phase of deadenylation alone contributes to ~40% of variance in the peak phase of L/S ratio (longest "S" bar in **Fig. 2B**). When the interactions of deadenylation with other processes are counted, this contribution increases to ~75% (longest "T" bar in **Fig. 2B**). Additionally, the Sobol indices also indicate that the relative amplitude of deadenylation has the strongest impact on the relative amplitude of L/S ratio (**Fig. S2**). In comparison, the mean level of L/S ratio, a quantity not related to rhythmicity, depends nearly equally on the mean rates of deadenylation and polyadenylation (**Fig. S2**). These results collectively demonstrate the rhythmicity in deadenylation as the strongest contributor to the rhythmicity in poly(A) tail length.

Our model results also show a significant impact of rhythmic deadenylation and

polyadenylation on the rhythmicity of L+S (total mRNA abundance). Although the peak phases of transcription and degradation strongly influence the peak phase of L+S, as expected (**Figs. 2C, D**), the Sobol indices indicate a weaker, yet substantial impact from the peak phases of deadenylation and polyadenylation on the peak phase of L+S (**Figs. 2D**). These impacts can be understood from the regulation of mRNA stability by poly(A) tail length, which is reflected in the model by the assumption that degradation is restricted to the short-tailed mRNAs (**Fig. 1A**, Eqs. (1) and (2)).

We further used the model to examine the effects of the four processes on the rhythmicity of mRNA translatability, using L (long-tailed mRNA abundance) as a proxy. Although L is a quantity directly related to both L+S level and L/S ratio, the Sobol indices show that the peak phase of L relies most heavily on the peak phase of deadenylation, followed by that of polyadenylation (**Fig. 2F**). Consistently, the scatter plot shows a strong correlation between the peak phases of L and deadenylation, with an approximately 10 hr lag between the two (**Fig. 2E**). This is a relationship highly similar to that observed between the peak phases of L/S ratio and deadenylation (**Fig. 2A**). Furthermore, the relative amplitudes of deadenylation and polyadenylation are also among the strongest contributors to the relative amplitude of L (**Fig. S2**). Overall, the rhythmicities in deadenylation and polyadenylation make stronger impact on the rhythmicity of long-tailed mRNA abundance than the rhythmicities in transcription and degradation. This finding provides an explanation for the observed close correlation between the rhythmicities of poly(A) tail length and protein expression [25].

Cytoplasmic polyadenylation requires specific *cis*-elements in the 3’ untranslated region (UTR) of an mRNA to recruit the molecular machinery that elongates poly(A) tails [38]. However, such *cis*-elements do not necessarily exist in all mRNAs. Therefore, we also removed

the polyadenylation term in our model and conducted the same global sensitivity analysis. The results demonstrate similar impacts of the rhythmicity of transcription, deadenylation and degradation on the rhythmicity of L/S ratio, L+S and L (**Fig. S3**) as those found from the model with cytoplasmic polyadenylation (**Figs. 2, 3, S2**). Particularly, rhythmic deadenylation remains the strongest contributor to the rhythmicity of L/S ratio and L. Hence, our conclusion stays the same for mRNAs without the *cis*-elements that mediate cytoplasmic polyadenylation.

Taken together, these model results underscore the importance of rhythmic poly(A) regulation in circadian gene expression, especially its impact on the rhythmicity of poly(A) tail length, total mRNA abundance, and abundance of the long-tailed subpopulation. Importantly, deadenylation emerges as the strongest contributor to the rhythmicity of poly(A) tail length and long-tailed mRNA abundance.

Rhythmic deadenylation can robustly group genes by their poly(A) tail rhythms

The rhythmicities in transcription, deadenylation, polyadenylation and degradation of mRNAs are ultimately controlled by the rhythmicities in the abundance and activity of the molecules mediating these processes, e.g., transcription factors, deadenylases and poly(A) polymerases. Interestingly, although the core clock machinery includes several transcription factors with different peak phases, the peak phases of nascent RNA synthesis (indicated by intron abundance) are strongly concentrated around ZT 15 in mouse liver [13]. Additionally, a cytoplasmic poly(A) polymerase, *Gld2*, is rhythmically expressed with peak phase around ZT 3.5 (Zeitgeber time, where ZT 0 is defined as the time [hr] of lights on and ZT 12 is defined as the time of lights off) [25]. Meanwhile, five deadenylases are also rhythmically expressed, with *Ccr4e/Angel1* peaking around ZT 2, *Ccr4a/Cnot6*, *Ccr4b/Cnot6l*, *Caf1a/Cnot7/pop2* and *Parn*

peaking around ZT 5, and *Ccr4c/Nocturnin* peaking around ZT 13 [25]. These data indicate that deadenylases assume a more diverse rhythmic expression pattern than poly(A) polymerases and nascent RNA transcription.

Intrigued by the above observation, we used our model to explore the potential consequence of having several distinct peak phases in deadenylases. In four separate *in silico* experiments, we set transcription, degradation, deadenylation or polyadenylation, respectively, to peak at three narrow windows centered around ZT 0, 8, and 16 (chosen to represent distinct time windows in general), while setting the peak phases of the other three processes to distribute evenly around the clock (**Fig. 3**, ii-v). Our results demonstrate that, when deadenylation peaks in three narrow windows, the peak phases of L/S ratio and L are strongly grouped into three distinct windows as well (**Fig. 3**, iv). In contrast, when transcription (**Fig. 3**, ii), degradation (**Fig. 3**, iii) or polyadenylation (**Fig. 3**, v) peaks in three narrow windows, the resulting peak phases of L/S ratio and L do not show any distinct grouping, and are comparable to the control case with evenly distributed peak phases for all processes (**Fig. 3**, i). To test the effect of the actual rhythmic patterns observed in nascent RNA transcription and expression of deadenylases and polyadenylases, we set the distribution of peak phases centered around ZT 15 for transcription [13], narrow peak phase window centered around ZT 3.5 for polyadenylation, and narrow peak phase windows around ZT 2, ZT 5 and ZT 13 for deadenylation [25]. The simulations results demonstrate that the peak phases of both L/S ratio and L are strongly clustered into three distinct time windows (**Fig. 3**, vi). These results corroborate with the findings above about the strong impact of rhythmic deadenylation on the rhythmicities of L/S ratio and L (**Fig. 2B, F**). Note that the mean rates and relative amplitudes of all four processes assumed random variables in the model simulations (**Table 1, Fig. S1**). Therefore, our results indicate that multiple peak phases in

deadenylation, but not other processes, can robustly group the peak phases of poly(A) tail length and mRNA translatability (~ long-tailed mRNA abundance) into distinct time windows, regardless of variations in the mean rates or rhythmicities of other processes.

Factors that explain different classes of mRNAs with rhythmic poly(A) tail length

In the previous transcriptome-wide study [25], the mRNAs with poly(A) tail rhythmicity (PAR mRNAs) were grouped into three classes, based on their rhythmicities in pre-mRNA and total mRNA. The rhythmicity in pre-mRNA essentially reflects the rhythmicity in transcription. The Class I mRNAs are rhythmic not only in poly(A) tail length, but also in pre-mRNA and total mRNA (**Fig. 4A**). The Class II mRNAs are rhythmic in poly(A) tail length and pre-mRNA, but not in total mRNA (**Fig. 4A**). The Class III mRNAs are only rhythmic in poly(A) tail length, but not the other two (**Fig. 5A**). Differences in mRNA half-lives were suggested to explain why the rhythmic patterns of pre-mRNA, total mRNA, and poly(A) tail length are different between these classes [25]. Here we leverage our model to systematically identify factors that can lead to the combinatorial rhythmic patterns in these classes.

We first attempted to identify the model parameters that contribute most to the distinction between Class I and Class II. Because the only difference between Class I and II is whether total mRNA abundance is rhythmic or not, we focused on identifying model parameters that contribute most to the relative amplitude of L+S. The Sobol indices reveal the mean degradation rate as the strongest contributor to the amplitude of L+S (**Fig. 4B**). We then ran model simulations using random parameter sets (sampled from the distributions given in **Table 1** and **Fig. S1**) and identified the ones that exhibit the characteristics of Class I or Class II (**Fig. 4A**). Out of all the random parameter sets, the mean degradation rates in the Class II parameter sets

are overall smaller than those in the Class I parameter sets (**Fig. 4C**). This finding corroborates with the experimental observation that the average half-life (i.e., reciprocal of degradation rate) of Class II mRNAs is longer than that of Class I mRNAs [25].

The total Sobol indices also indicate that the peak phases of transcription and degradation as the second and third strongest contributors to the amplitude of L+S, respectively (**Fig. 4B**). However, the corresponding single indices are diminishingly small (**Fig. 4B**). The huge contrast between the total and single indices indicates that these two parameters exert strong impacts through interactions with other parameters. Because such huge contrasts between total and single indices do not exist in any other parameters, we speculated that the interactions likely happen between the two parameters themselves. Indeed, the Class I, but not the Class II, parameter sets, are strongly enriched with antiphasic rhythms between transcription and degradation (**Fig. 4D**). This finding is consistent with the effect of antiphasic coupling between rhythmic transcription and degradation predicted by a previous modeling study [16].

The Sobol indices also reveal that the relative amplitudes of transcription and degradation rates and the mean deadenylation rate are potentially important contributors to the amplitude of L+S (**Fig. 4B**). Indeed, the Class I parameter sets tend to have stronger amplitudes in transcription and degradation rates (**Fig. 4E, F**), again, consistent with the previous modeling study [16]. Interestingly, unlike the Class I parameter sets (**Figs. 4D-F**, purple), the Class II parameter sets exhibit nearly even distributions of transcription-degradation phase difference, transcription amplitude and degradation amplitude (**Figs. 4D-F**, green). Therefore, generation of significant rhythmicity in L+S (Class I) requires sufficient phase difference and amplitudes in transcription and degradation and sufficiently high amplitudes of transcription and degradation, simultaneously. If any of these conditions are not satisfied, total mRNA abundance would not

have significant rhythmicity, and the mRNA would belong to Class II. Lastly, the mean deadenylation rates in the Class I parameter sets tend to be larger than those in the Class II parameter sets (**Fig. 4G**). This is related to the above finding about mRNA half-lives, because deadenylation promotes degradation and hence increasing the mean deadenylation rate has a similar effect on mRNA turnover as increasing the mean degradation rate.

Class III is distinct from Class I and Class II, since it does not have rhythmic transcription (**Fig. 5A**). Because rhythmicity of transcription serves as input to our model, we cannot use the model to identify the origin of lack of transcriptional rhythmicity. However, we are interested in understanding why all PAR mRNAs without transcriptional rhythmicity also lack rhythmicity in L+S, and only exhibit rhythmicity in L/S ratio [25]. For the convenience of discussion, we use "Class IV" to refer to a hypothetical group of PAR mRNAs that would exhibit rhythmicity in total mRNAs and poly(A) tails, but not in pre-RNA (**Fig. 5A**); this group of mRNAs are not found in the experiments [25]. We used the model to identify model parameters that could contribute to the difference between Class III and the hypothetical Class IV. Because both Class III and Class IV do not have rhythmic transcription, we ran model simulations with non-rhythmic transcription (i.e., setting the relative amplitude of transcription to zero, while keeping the other parameters sampled from the same distributions as before (**Table 1, Fig. S1**)). Out of the random parameter sets, we identified those that fit the characteristics of Class III or Class IV (**Fig. 5A**). We also calculated the Sobol indices for this model.

When the model does not have rhythmic transcription, the Sobol indices again reveal the mean degradation rate as the strongest contributor to the relative amplitude of L+S (**Fig. 5B**). Consistently, the Class IV parameter sets require much larger mRNA degradation rate, i.e., much shorter half-life of mRNA, than the Class III parameter sets, to sustain rhythmic total mRNA

(**Fig. 5C**). Therefore, the absence of Class IV mRNAs in the experiment is most likely due to the long half-lives of the mRNAs without rhythmic transcription. Indeed, Class III has the longest average mRNA half-life measured among all mRNAs that are rhythmically expressed [25].

We also identified a few additional factors that would distinguish Class III from Class IV. The second strongest factor affecting the amplitude of L+S is the relative amplitude of degradation rate, based on the Sobol indices (**Fig. 5B)**. The Class IV parameter sets have markedly higher amplitudes of degradation rate than Class III (**Fig. 5D**). The phases of all three rhythmic processes, i.e., degradation, deadenylation and polyadenylation, are also potentially important contributors, because their total Sobol indices are substantial (**Fig. 5B**). Again, the huge contrast between the total and single indices for these phase parameters, but not the other parameters, suggests that they exert impacts through interactions among themselves. We hence examined the distribution of pairwise differences between the three phase parameters. To achieve the rhythmic characteristics of the hypothetical Class IV, the peak phases of deadenylation and degradation need to be close to each other, but opposite to that of polyadenylation (**Fig. 5E**). This can be understood from the fact that both deadenylation and degradation promote mRNA turnover while polyadenylation inhibits it. Unlike the Class IV parameter sets, no distinct patterns are found in the amplitude of degradation rate or the phase differences in the Class III parameter sets (**Figs. 5D, E**). Similar to the discussion above for Class I and Class II, these results indicate that the Class IV characteristics requires both sufficiently large amplitude in degradation and sufficient differences of the polyadenylation phase from the deadenylation and degradation phases. The missing of Class IV from the experiment suggests that mRNAs without transcriptional rhythmicity do not satisfy these conditions at the same time.

Overall, our model suggests that besides mRNA half-life, relative amplitudes and phase difference between transcriptional and post-transcriptional processes also contribute to the rhythmic characteristics that distinguishes the three observed classes of PAR mRNAs (**Figs. 4, 5**). These results highlight that rhythmic transcriptional and post-transcriptional processes collectively determine the rhythmicity in mRNA expression and poly(A) tail length. It will be of future interests to test if the factors predicted by the model are indeed correlated with different rhythmic characteristics.

Discussion

Many post-transcriptional steps are involved in regulating circadian gene expression [10], among which rhythmic regulation of poly(A) tail length was recently discovered and characterized at the genome-wide scale [25]. Notably, the rhythmicity in poly(A) tail length can be both the cause and result of rhythmic gene expression. On the one hand, the rhythmicity in poly(A) tail length can control the rhythmicity in mRNA degradation [16], because mRNA stability is related to the poly(A) tail length [24, 30-32]. On the other hand, the rhythmicity in poly(A) tail length itself stems from the rhythmic control of post-transcriptional processes, such as polyadenylation and deadenylation. At the mechanistic basis, it is the dynamic coupling between multiple rhythmic processes, i.e., transcription, polyadenylation, deadenylation and degradation, that determines the rhythmic patterns in both poly(A) tail length and mRNA abundance.

In this work, we developed a parsimonious mathematical model (**Fig. 1**) to quantitatively evaluate how rhythmic inputs from the transcription, degradation, polyadenylation and deadenylation processes collectively determine the rhythmic outputs in mRNA abundance,

poly(A) tail length and long-tailed mRNA abundance. Our model results and global sensitivity analyses reveal rhythmic deadenylation as the strongest factor in controlling the peak phases and amplitudes of rhythmic poly(A) tail length and long-tailed mRNA abundance (**Figs. 2, 3**). This finding highlights the crucial role of rhythmic poly(A) regulation in circadian gene expression. Our model also suggests how three classes of rhythmic characteristics observed in PAR mRNAs [25] arise from the dynamic features of the four processes, as well as the coupling among their rhythmicities (**Figs. 4, 5**).

The importance of dynamic coupling between rhythmic transcription and post-transcriptional processes was demonstrated by a previous modeling study by Lück et al. [16]. That work particularly highlights that rhythmic turnover is necessary for achieving >6 hr peak phase difference between transcription and mRNA abundance. In comparison, our study explicitly considers the effects of poly(A) regulation, a common intermediate process in the mRNA decay pathway, on the rhythmicity of both mRNA abundance and poly(A) tail length. Because deadenylation is necessary for mRNA degradation and polyadenylation opposes it, rhythmic deadenylation and polyadenylation, unsurprisingly, affect the rhythmicity of total mRNA abundance at a level comparable to rhythmic degradation (**Figs. 2, S2**). However, when poly(A) tail length and its effect on mRNA translatability are considered, rhythmic deadenylation emerges as the most important step affecting their rhythmicity (**Figs. 2, S2**). Of course, our model has not included other mRNA decay pathways that do not depend on poly(A) regulation, such as endonuclease cleavage of mRNA followed by 5’-3’ decay [42]. For any mRNA decayed through these pathways, which are less common, their rhythmicity obviously would not depend on the rhythmicity in poly(A) regulation.

Based on the finding of rhythmic deadenylation as the strongest contributor to

rhythmicity of poly(A) tail length and long-tailed mRNA abundance, we further discovered that rhythmic deadenylation is capable of synchronizing the target circadian gene expression post-transcriptionally. According to the model results, three distinct peak phases in deadenylation activity, as those observed in mouse liver [25], can robustly divide the mRNAs into three distinct groups by their peak phases of poly(A) tail length and long-tailed mRNA abundance; this grouping effect by deadenylation phases happens regardless of the rhythmicity in other processes (**Fig. 3**). This finding suggests a potential mechanism to synchronize the expression of genes controlled by the same deadenylases and hence foster synergy among these genes around the clock. This synchronization potential is unique to rhythmic deadenylation, but not the other rhythmic processes (**Fig. 3**).

The capability of deadenylation to synchronize circadian gene expression further poses two interesting questions. First, could deadenylation help synchronize circadian gene expression among different cells and entrain the cell-autonomous clock to the systemic rhythms? Recent studies suggest that rhythmic feeding or other systemic rhythmic cues control the rhythmic expression of several deadenylases, including *Parn*, *Pan2* [17] and *Nocturnin* [48], through clock-independent pathways. Given the findings by our model, such systemically driven rhythmicity in deadenylases could dictate the rhythmicity of poly(A) length and mRNA translatability (~long-tailed mRNA abundance). This could help synchronize circadian gene expression in cells influenced by the same systemic signals. Second, could deadenylases play a role in tissue-specific circadian gene expression? Rhythmic gene expression is known to vary tremendously from tissue to tissue: different tissues not only share very few rhythmically expressed genes beyond the core clock genes, but also display different peak times for some genes [5, 49, 50]. It is puzzling how the rhythmicity in gene expression varies so much across

different tissues while the cellular clock machineries are the same and are presumably synchronized throughout the organism. Most previous studies on the mechanisms of tissue-specific circadian gene expression have focused on tissue-specific transcriptional control, such as rhythmic fluctuations in chromatin structure and interactions between core clock transcription factors and tissue-specific transcription factors [51, 52]. In light of the findings from our work, differential expression patterns of deadenylases in different tissues [53] could serve as an additional mechanism to mediate tissue-specific circadian gene expression. These two interesting questions await future studies to answer.

Circadian gene expression is a critical, yet highly complex process. Expressing the right gene at the right time and the right place requires coordinated control at various gene expression steps, as well as across different cells and tissues. Systems-level study of the coupling between different rhythmic processes is necessary to gain comprehensive understanding of circadian gene expression control, and more importantly, the ability to make positive use of circadian rhythm in treatments of diseases. As our work demonstrates the significant impact of rhythmic poly(A) regulation and its coupling with rhythmic mRNA transcription and degradation on circadian gene expression, it will be of great future interest to examine how coupling between the rhythmicities of transcription, poly(A) regulation and other post-transcriptional, translational and post-translational processes influences circadian gene expression.

Methods

Model simulation and extraction of phase, amplitude and mean from simulation results

For any given parameter set, Eqs. (1) and (2) were simulated using the ordinary differential equation (ODE) solver, `ode45`, in MATLAB. For a simulated time trajectory

$\{L(t), S(t)\}$, the peak phases, relative amplitudes and mean levels of $L(t) + S(t), L(t)/S(t)$ and $L(t)$ were analyzed. First, the time trajectory for the output quantity of interest, e.g., $L(t)/S(t)$, was calculated from $\{L(t), S(t)\}$. Then a 48-hr window after 700 hrs (sufficiently long to pass the initial transient) were extracted from the simulated trajectory for data analysis. The trajectories typically have irregular time spacing (due to automatic time stepping in the ode45 solver) and often have insufficient time resolution for accurate determination of the peak phase. To improve the accuracy of the estimated peak phase, the 48-hr trajectory were interpolated upon 500 equally spaced time points spanning the 48 hrs. The peak phase was evaluated from the time for the maximum value, t_{max}, i.e., $\text{peak phase} = mod(t_{max} + 700, 24)$ (hr). The mean value was estimated by taking the average of the 500 interpolated data values within the 48-hr window. The relative amplitude was estimated by taking the maximum and minimum values within the 48-hr window and calculating $(\max - \min)/(2 \times \text{mean})$. An output quantity was considered rhythmic if its relative amplitude is equal to or greater than 0.2.

Parameter sampling

We performed global parameter sensitivity analysis [54] on the model to analyze the general contribution of each parameter to any specific output quantity (i.e., peak phase, relative amplitude and mean of $L(t) + S(t), L(t)/S(t)$ and $L(t)$). To perform such global sensitivity analysis, one needs to simulate the model with randomly chosen parameter values that maximally represent the parameter space. In this study we drew random parameter values from the distributions listed in **Table 1** and plotted in **Fig. S1**. The peak phases and relative amplitudes were sampled form uniform distributions of their possible ranges by definition (**Table 1**). The mean reaction rates were sampled from log-normal distributions suggested by previous genomic

scale measurements (see sources indicated in **Table 1**). It is noteworthy that we set the mean transcription rate as constant, as it only causes proportional changes in $L(t)$ and $S(t)$, and does not affect the rhythmic patterns of any quantity (see Supplementary Materials). To improve the accuracy of the global sensitivity analysis for models with many parameters, one needs sampled parameter values that well represent the parameter space. To this end, we used the sampling method of Latin hypercube [55], which is known to ensure good representation of the high-dimensional parameter space.

Table 1. Parameter distribution for sampling.

Parameter	Symbol	Distribution	Source
Mean rate of transcription	k_{trsc}	1 (constant)	Has no effect on rhythmic patterns (see Supplementary Materials)
Mean rate of degradation	k_{dgrd}	$log_{10}(k/\mathrm{hr}^{-1}) \sim \mathcal{N}(-1.10, 0.23^2)$ (Log-normal)	Fitting with half-life distribution measured in [56]
Mean rate of deadenylation	k_{deA}	$log_{10}(k/\mathrm{hr}^{-1}) \sim \mathcal{N}(-0.48, 0.23^2)$ (Log-normal)	Mean value of deadenylation rate estimated from [57]; deviation same as mRNA degradation
Mean rate of polyadenylation	k_{polyA}	$log_{10}(k/\mathrm{hr}^{-1}) \sim \mathcal{N}(-0.48, 0.23^2)$ (Log-normal)	Same as deadenylation
Relative amplitudes	$A_{\mathrm{trsc}}, A_{\mathrm{dgrd}}, A_{\mathrm{deA}}, A_{\mathrm{polyA}}$	$A \sim U(0,1)$ (Uniform)	
Peak phases	$\varphi_{\mathrm{trsc}}, \varphi_{\mathrm{dgrd}}, \varphi_{\mathrm{deA}}, \varphi_{\mathrm{polyA}}$	$\varphi \sim U(0,24)$ (Uniform)	

Sobol's method of global sensitivity analysis

To evaluate the impact of each model parameter (e.g., phase of deadenylation) on each model output (e.g., relative amplitude of L/S ratio), we used a variance-based global parameter sensitivity analysis method, the Sobol indices [47]. For each pair of parameter X_i and output Y_j in the model, the first-order, or single Sobol index, S_{ij}, characterizes the contribution of variance in X_i alone to the total variance in Y_j (Eq.(4)). The total-effect, or total index, S_{Tij}, characterizes the contribution of variance in X_i, as well as the variance caused by its coupling with other parameters, to the total variance in Y_j (Eq.(5)). The larger these indices are, the more sensitive Y_j is to X_i, or the more impact X_i has on Y_j.

$$S_{ij} = \frac{\mathrm{Var}_{X_i}\left(\mathrm{E}_{X_{\sim i}}\left(Y_j \mid X_i\right)\right)}{\mathrm{Var}\left(Y_j\right)} \tag{4}$$

$$S_{Tij} = \frac{\mathrm{E}_{X_{\sim i}}\left(\mathrm{Var}_{X_i}\left(Y_j \mid X_{\sim i}\right)\right)}{\mathrm{Var}\left(Y_j\right)} \tag{5}$$

where the subscript $\sim i$ indicates all indices except for i.

We followed the specific algorithms given in [46] and [58] for evaluating the single (Eq.(4)) and total indices (Eq.(5)), respectively. The details of implementation are explained below.

1) Sample from the distributions given in **Table 1** two independent groups of N parameter sets ($N = 100{,}000$ in this study):

$$A = \begin{bmatrix} A_{1,1} & \cdots & A_{1,k} \\ \vdots & \ddots & \vdots \\ A_{N,1} & \cdots & A_{N,k} \end{bmatrix}, B = \begin{bmatrix} B_{1,1} & \cdots & B_{1,k} \\ \vdots & \ddots & \vdots \\ B_{N,1} & \cdots & B_{N,k} \end{bmatrix} \tag{6}$$

Each row in A and B represents one set of k parameters. $k = 11$ for the model with cytoplasmic polyadenylation. $k = 8$ for the model without cytoplasmic polyadenylation.

$k = 9$ for the model without transcriptional rhythmicity.

2) Construct k hybrid groups of parameter sets. The i-th group, $A_B^{(i)}$, has the i-th column equal to the i-th column of B, and the remaining columns copied from A, where $i = 1, \ldots, k$.

$$A_B^{(i)} = \begin{bmatrix} A_{1,1} & \cdots & B_{1,i} & \cdots & A_{1,k} \\ A_{2,1} & \cdots & B_{2,i} & \cdots & A_{2,k} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{N,2} & \cdots & B_{N,i} & \cdots & A_{N,k} \end{bmatrix} \tag{7}$$

3) Estimate the total variance in the output.

$$V_j = Var\left(Y_j\right) = \frac{1}{2N}\sum_{n=1}^{N}\left\{\left[f_j\left(A_{(n)}\right) - \overline{f_j\left(A_{(n)}\right)}\right]^2 + \left[f_j\left(B_{(n)}\right) - \overline{f_j\left(B_{(n)}\right)}\right]^2\right\} \tag{8}$$

where f_j denotes the j-th output quantity from the circadian gene expression model (Eqs. (1) and (2)). $A_{(n)}$ and $B_{(n)}$ denote the n-th parameter set (row) in Groups A and B, respectively. The bars on top denote the average of output quantities over N parameter sets.

4) Estimate the single and total Sobol indices, using Eqs. (9) and (10) [46, 58].

$$S_{ij} \approx \frac{1}{N}\sum_{n=1}^{N} f_j\left(B_{(n)}\right)\left[f_j\left(\left(A_B^{(i)}\right)_{(n)}\right) - f_j\left(A_{(n)}\right)\right] \Big/ V_j \tag{9}$$

$$S_{Tij} \approx \frac{1}{2N}\sum_{n=1}^{N}\left[f_j\left(\left(A_B^{(i)}\right)_{(n)}\right) - f_j\left(A_{(n)}\right)\right]^2 \Big/ V_j \tag{10}$$

where $\left(A_B^{(i)}\right)_{(n)}$ denotes the n-th parameter set (row) in the i-th hybrid group, and the other notations follow those described above.

Acknowledgement

We thank Dr. Xi Chen (Virginia Tech) for helpful discussion of the Sobol method.

References

1. Mohawk JA, Green CB, Takahashi JS. Central and peripheral circadian clocks in mammals. Annu Rev Neurosci. 2012;35:445-62. doi: 10.1146/annurev-neuro-060909-153128. PubMed PMID: 22483041; PubMed Central PMCID: PMCPMC3710582.
2. Takahashi JS, Hong HK, Ko CH, McDearmon EL. The genetics of mammalian circadian order and disorder: implications for physiology and disease. Nature reviews Genetics. 2008;9(10):764-75. doi: 10.1038/nrg2430. PubMed PMID: 18802415; PubMed Central PMCID: PMCPMC3758473.
3. Duffield GE. DNA microarray analyses of circadian timing: the genomic basis of biological time. J Neuroendocrinol. 2003;15(10):991-1002. PubMed PMID: 12969245.
4. Kojima S, Green CB. Circadian genomics reveal a role for post-transcriptional regulation in mammals. Biochemistry. 2015;54(2):124-33. doi: 10.1021/bi500707c. PubMed PMID: 25303020; PubMed Central PMCID: PMCPMC4302021.
5. Zhang R, Lahens NF, Ballance HI, Hughes ME, Hogenesch JB. A circadian gene expression atlas in mammals: Implications for biology and medicine. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(45):16219-24. doi: 10.1073/pnas.1408886111. PubMed PMID: WOS:000344526800081.
6. Takahashi JS. Molecular components of the circadian clock in mammals. Diabetes Obes Metab. 2015;17 Suppl 1:6-11. Epub 2015/09/04. doi: 10.1111/dom.12514. PubMed PMID: 26332962; PubMed Central PMCID: PMCPMC4560116.
7. Zhang EE, Kay SA. Clocks not winding down: unravelling circadian networks. Nature reviews Molecular cell biology. 2010;11(11):764-76. doi: 10.1038/nrm2995. PubMed PMID: 20966970.
8. Partch CL, Green CB, Takahashi JS. Molecular architecture of the mammalian circadian clock. Trends in cell biology. 2014;24(2):90-9. doi: 10.1016/j.tcb.2013.07.002. PubMed PMID: 23916625; PubMed Central PMCID: PMCPMC3946763.
9. Ueda HR, Hayashi S, Chen W, Sano M, Machida M, Shigeyoshi Y, et al. System-level identification of transcriptional circuits underlying mammalian circadian clocks. Nature genetics. 2005;37(2):187-92. doi: 10.1038/ng1504. PubMed PMID: 15665827.
10. Kojima S, Shingle DL, Green CB. Post-transcriptional control of circadian rhythms. Journal of cell science. 2011;124(Pt 3):311-20. Epub 2011/01/19. doi: 10.1242/jcs.065771. PubMed PMID: 21242310; PubMed Central PMCID: PMCPMC3021995.
11. Hirano A, Fu YH, Ptacek LJ. The intricate dance of post-translational modifications in the rhythm of life. Nat Struct Mol Biol. 2016;23(12):1053-60. Epub 2016/12/07. doi: 10.1038/nsmb.3326. PubMed PMID: 27922612.
12. Gobet C, Naef F. Ribosome profiling and dynamic regulation of translation in mammals. Current opinion in genetics & development. 2017;43:120-7. Epub 2017/04/01. doi: 10.1016/j.gde.2017.03.005. PubMed PMID: 28363112.
13. Koike N, Yoo SH, Huang HC, Kumar V, Lee C, Kim TK, et al. Transcriptional

architecture and chromatin landscape of the core circadian clock in mammals. Science. 2012;338(6105):349-54. doi: 10.1126/science.1226339. PubMed PMID: 22936566; PubMed Central PMCID: PMCPMC3694775.

14. Menet JS, Rodriguez J, Abruzzi KC, Rosbash M. Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. eLife. 2012;1:e00011. doi: 10.7554/eLife.00011. PubMed PMID: 23150795; PubMed Central PMCID: PMCPMC3492862.

15. Trott AJ, Menet JS. Regulation of circadian clock transcriptional output by CLOCK:BMAL1. PLoS genetics. 2018;14(1):e1007156. Epub 2018/01/05. doi: 10.1371/journal.pgen.1007156. PubMed PMID: 29300726; PubMed Central PMCID: PMCPMC5771620.

16. Luck S, Thurley K, Thaben PF, Westermark PO. Rhythmic degradation explains and unifies circadian transcriptome and proteome data. Cell reports. 2014;9(2):741-51. doi: 10.1016/j.celrep.2014.09.021. PubMed PMID: 25373909.

17. Wang J, Symul L, Yeung J, Gobet C, Sobel J, Luck S, et al. Circadian clock-dependent and -independent posttranscriptional regulation underlies temporal mRNA accumulation in mouse liver. Proceedings of the National Academy of Sciences of the United States of America. 2018;115(8):E1916-E25. Epub 2018/02/13. doi: 10.1073/pnas.1715225115. PubMed PMID: 29432155; PubMed Central PMCID: PMCPMC5828596.

18. Keene JD. Minireview: global regulation and dynamics of ribonucleic Acid. Endocrinology. 2010;151(4):1391-7. Epub 2010/03/25. doi: 10.1210/en.2009-1250. PubMed PMID: 20332203; PubMed Central PMCID: PMCPMC2850242.

19. Kojima S, Gatfield D, Esau CC, Green CB. MicroRNA-122 modulates the rhythmic expression profile of the circadian deadenylase Nocturnin in mouse liver. PloS one. 2010;5(6):e11264. doi: 10.1371/journal.pone.0011264. PubMed PMID: 20582318; PubMed Central PMCID: PMCPMC2889834.

20. Du NH, Arpat AB, De Matos M, Gatfield D. MicroRNAs shape circadian hepatic gene expression on a transcriptome-wide scale. eLife. 2014;3:e02510. Epub 2014/05/29. doi: 10.7554/eLife.02510. PubMed PMID: 24867642; PubMed Central PMCID: PMCPMC4032493.

21. Kim TD, Kim JS, Kim JH, Myung J, Chae HD, Woo KC, et al. Rhythmic serotonin N-acetyltransferase mRNA degradation is essential for the maintenance of its circadian oscillation. Molecular and cellular biology. 2005;25(8):3232-46. Epub 2005/03/31. doi: 10.1128/MCB.25.8.3232-3246.2005. PubMed PMID: 15798208; PubMed Central PMCID: PMCPMC1069600.

22. Beilharz TH, Preiss T. Widespread use of poly(A) tail length control to accentuate expression of the yeast transcriptome. Rna-a Publication of the Rna Society. 2007;13(7):982-97. doi: 10.1261/rna.569407. PubMed PMID: WOS:000247465800005.

23. Jalkanen AL, Coleman SJ, Wilusz J. Determinants and implications of mRNA poly(A) tail size--does this protein make my tail look big? Semin Cell Dev Biol. 2014;34:24-32. Epub 2014/06/10. doi: 10.1016/j.semcdb.2014.05.018. PubMed PMID: 24910447; PubMed Central PMCID: PMCPMC4163081.

24. Eckmann CR, Rammelt C, Wahle E. Control of poly(A) tail length. Wiley interdisciplinary reviews RNA. 2011;2(3):348-61. doi: 10.1002/wrna.56. PubMed PMID: 21957022.

25. Kojima S, Sher-Chen EL, Green CB. Circadian control of mRNA polyadenylation dynamics regulates rhythmic protein expression. Genes & development. 2012;26(24):2724-36. doi: 10.1101/gad.208306.112. PubMed PMID: 23249735; PubMed Central PMCID:

PMCPMC3533077.
26. Robinson BG, Frim DM, Schwartz WJ, Majzoub JA. Vasopressin mRNA in the suprachiasmatic nuclei: daily regulation of polyadenylate tail length. Science. 1988;241(4863):342-4. Epub 1988/07/15. PubMed PMID: 3388044.
27. Gerstner JR, Vanderheyden WM, LaVaute T, Westmark CJ, Rouhana L, Pack AI, et al. Time of day regulates subcellular trafficking, tripartite synaptic localization, and polyadenylation of the astrocytic Fabp7 mRNA. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2012;32(4):1383-94. Epub 2012/01/27. doi: 10.1523/JNEUROSCI.3228-11.2012. PubMed PMID: 22279223; PubMed Central PMCID: PMCPMC3564590.
28. Stubblefield JJ, Gao P, Kilaru G, Mukadam B, Terrien J, Green CB. Temporal Control of Metabolic Amplitude by Nocturnin. Cell reports. 2018;22(5):1225-35. Epub 2018/02/02. doi: 10.1016/j.celrep.2018.01.011. PubMed PMID: 29386110; PubMed Central PMCID: PMCPMC5815321.
29. Wang Y, Osterbur DL, Megaw PL, Tosini G, Fukuhara C, Green CB, et al. Rhythmic expression of Nocturnin mRNA in multiple tissues of the mouse. BMC Dev Biol. 2001;1:9. Epub 2001/06/08. PubMed PMID: 11394964; PubMed Central PMCID: PMCPMC32249.
30. Weill L, Belloc E, Bava FA, Mendez R. Translational control by changes in poly(A) tail length: recycling mRNAs. Nature Structural & Molecular Biology. 2012;19(6):577-85. doi: 10.1038/nsmb.2311. PubMed PMID: WOS:000304958200004.
31. Wells SE, Hillner PE, Vale RD, Sachs AB. Circularization of mRNA by eukaryotic translation initiation factors. Molecular cell. 1998;2(1):135-40. Epub 1998/08/14. doi: 10.1016/S1097-2765(00)80122-7. PubMed PMID: 9702200.
32. Jacobson A, Favreau M. Possible involvement of poly(A) in protein synthesis. Nucleic acids research. 1983;11(18):6353-68. Epub 1983/09/24. doi: 10.1093/nar/11.18.6353. PubMed PMID: 6137807; PubMed Central PMCID: PMCPMC326378.
33. Gallie DR. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. Genes & development. 1991;5(11):2108-16. Epub 1991/11/01. doi: 10.1101/gad.5.11.2108. PubMed PMID: 1682219.
34. Gallie DR. A tale of two termini: a functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. Gene. 1998;216(1):1-11. Epub 1998/08/26. PubMed PMID: 9714706.
35. Zhao J, Hyman L, Moore C. Formation of mRNA 3 ' ends in eukaryotes: Mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiol Mol Biol R. 1999;63(2):405-+. PubMed PMID: WOS:000080690800008.
36. Garneau NL, Wilusz J, Wilusz CJ. The highways and byways of mRNA decay. Nature reviews Molecular cell biology. 2007;8(2):113-26. Epub 2007/01/25. doi: 10.1038/nrm2104. PubMed PMID: 17245413.
37. Mendez R, Richter JD. Translational control by CPEB: A means to the end. Nat Rev Mol Cell Bio. 2001;2(7):521-9. doi: 10.1038/35080081. PubMed PMID: WOS:000169789100015.
38. Charlesworth A, Meijer HA, de Moor CH. Specificity factors in cytoplasmic polyadenylation. Wiley interdisciplinary reviews RNA. 2013;4(4):437-61. Epub 2013/06/19. doi: 10.1002/wrna.1171. PubMed PMID: 23776146; PubMed Central PMCID: PMCPMC3736149.
39. Cui J, Sartain CV, Pleiss JA, Wolfner MF. Cytoplasmic polyadenylation is a major mRNA regulator during oogenesis and egg activation in Drosophila. Developmental biology. 2013;383(1):121-31. doi: 10.1016/j.ydbio.2013.08.013. PubMed PMID: WOS:000326057800013.

40. Kwak JE, Drier E, Barbee SA, Ramaswami M, Yin JCP, Wickens M. GLD2 poly(A) polymerase is required for long-term memory. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(38):14644-9. doi: 10.1073/pnas.0803185105. PubMed PMID: WOS:000259592400075.
41. Wu X, Brewer G. The regulation of mRNA stability in mammalian cells: 2.0. Gene. 2012;500(1):10-21. doi: 10.1016/j.gene.2012.03.021. PubMed PMID: 22452843; PubMed Central PMCID: PMCPMC3340483.
42. Schoenberg DR, Maquat LE. Regulation of cytoplasmic mRNA decay. Nature reviews Genetics. 2012;13(4):246-59. Epub 2012/03/07. doi: 10.1038/nrg3160. PubMed PMID: 22392217; PubMed Central PMCID: PMCPMC3351101.
43. Decker CJ, Parker R. A Turnover Pathway for Both Stable and Unstable Messenger-Rnas in Yeast - Evidence for a Requirement for Deadenylation. Genes & development. 1993;7(8):1632-43. doi: 10.1101/gad.7.8.1632. PubMed PMID: WOS:A1993LR77600014.
44. Muhlrad D, Parker R. Mutations affecting stability and deadenylation of the yeast MFA2 transcript. Genes & development. 1992;6(11):2100-11. Epub 1992/11/01. doi: 10.1101/gad.6.11.2100. PubMed PMID: 1427074.
45. Proudfoot NJ. Ending the message: poly(A) signals then and now. Genes & development. 2011;25(17):1770-82. doi: 10.1101/gad.17268411. PubMed PMID: 21896654; PubMed Central PMCID: PMCPMC3175714.
46. Sobol IM. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math Comput Simulat. 2001;55(1-3):271-80. doi: 10.1016/S0378-4754(00)00270-6. PubMed PMID: WOS:000167385700029.
47. Iooss B, Lemaitre P. A Review on Global Sensitivity Analysis Methods. In: Dellino G, Meloni C, editors. Uncertainty Management in Simulation-Optimization of Complex Systems. New York: Springer Science+Business Media; 2015.
48. Kornmann B, Schaad O, Bujard H, Takahashi JS, Schibler U. System-driven and oscillator-dependent circadian transcription in mice with a conditionally active liver clock. PLoS biology. 2007;5(2):e34. Epub 2007/02/15. doi: 10.1371/journal.pbio.0050034. PubMed PMID: 17298173; PubMed Central PMCID: PMCPMC1783671.
49. Panda S, Antoch MP, Miller BH, Su AI, Schook AB, Straume M, et al. Coordinated transcription of key pathways in the mouse by the circadian clock. Cell. 2002;109(3):307-20. doi: Doi 10.1016/S0092-8674(02)00722-5. PubMed PMID: WOS:000175412100007.
50. Storch KF, Lipan O, Leykin I, Viswanathan N, Davis FC, Wong WH, et al. Extensive and divergent circadian gene expression in liver and heart. Nature. 2002;417(6884):78-83. doi: DOI 10.1038/nature744. PubMed PMID: WOS:000175307200041.
51. Mermet J, Yeung J, Naef F. Systems Chronobiology: Global Analysis of Gene Regulation in a 24-Hour Periodic World. Csh Perspect Biol. 2017;9(3). doi: ARTN a028720 10.1101/cshperspect.a028720. PubMed PMID: WOS:000397038600006.
52. Yeung J, Mermet J, Jouffe C, Marquis J, Charpagne A, Gachon F, et al. Transcription factor activity rhythms and tissue-specific chromatin interactions explain circadian gene expression across organs. Genome research. 2018;28(2):182-91. Epub 2017/12/20. doi: 10.1101/gr.222430.117. PubMed PMID: 29254942; PubMed Central PMCID: PMCPMC5793782.
53. Pizarro A, Hayer K, Lahens NF, Hogenesch JB. CircaDB: a database of mammalian circadian gene expression profiles. Nucleic acids research. 2013;41(Database issue):D1009-13. Epub 2012/11/28. doi: 10.1093/nar/gks1161. PubMed PMID: 23180795; PubMed Central

PMCID: PMCPMC3531170.
54. Saltelli A. Global sensitivity analysis : the primer. Chichester, England ;: John Wiley; 2008. Available from: http://dx.doi.org/10.1002/9780470725184.
55. Mckay MD, Beckman RJ, Conover WJ. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. Technometrics. 1979;21(2):239-45. doi: Doi 10.2307/1268522. PubMed PMID: WOS:A1979GW16700012.
56. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. Nature. 2011;473(7347):337-42. doi: 10.1038/nature10098. PubMed PMID: 21593866.
57. Yi H, Park J, Ha M, Lim J, Chang H, Kim VN. PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay. Molecular cell. 2018;70(6):1081-8 e5. Epub 2018/06/23. doi: 10.1016/j.molcel.2018.05.009. PubMed PMID: 29932901.
58. Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput Phys Commun. 2010;181(2):259-70. doi: 10.1016/j.cpc.2009.09.018. PubMed PMID: WOS:000273192100002.

Supporting Information

S1 File: Supplementary Materials.

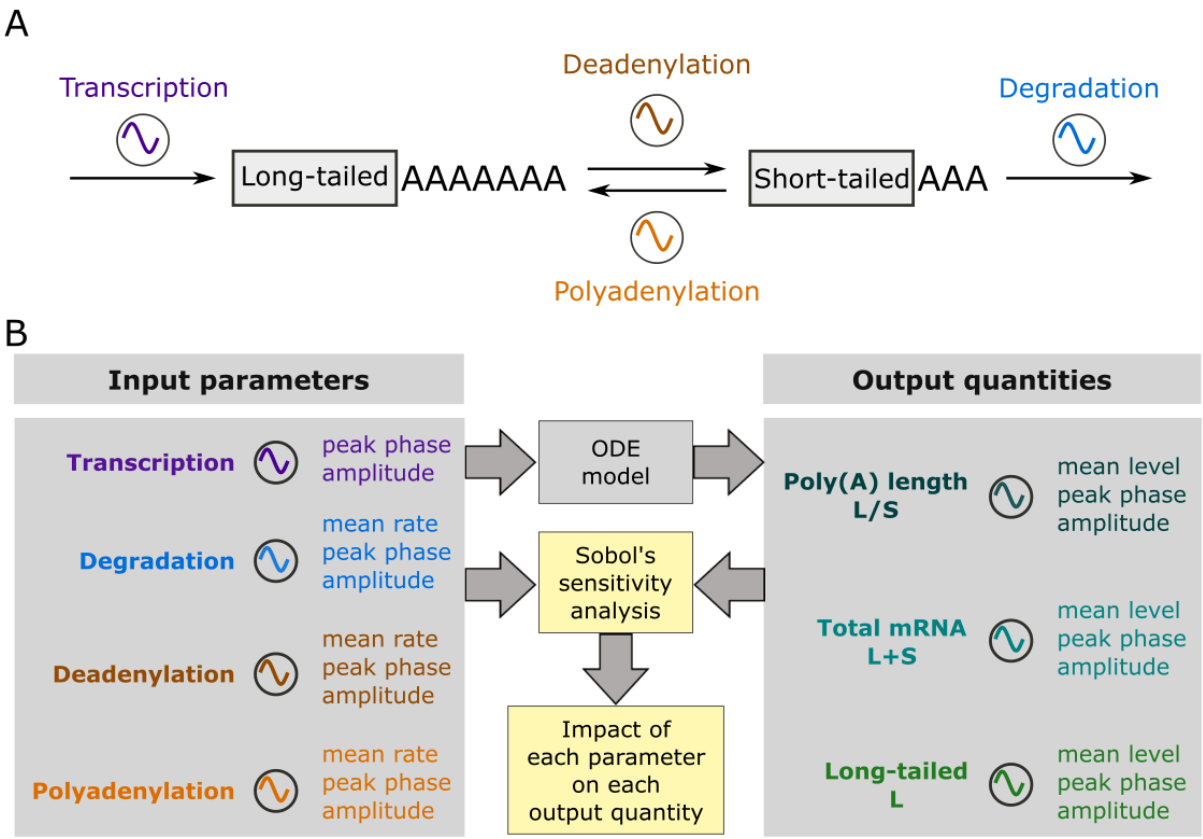

Figure 1. Schematic diagram of the model. (A) The model describes four processes which control the poly(A) tail length and mRNA abundance: transcription, degradation, cytoplasmic deadenylation and polyadenylation. All four processes are allowed variable rhythmicity. (B) Work flow of the study. Numeric simulations of the ODE model using different sets of input parameters (sampled according to **Table 1, Fig. S1**) generate the output quantities. The input parameters and output quantities are analyzed through the global sensitivity analysis to quantify the impact of each parameter on each output quantity over the global parameter space.

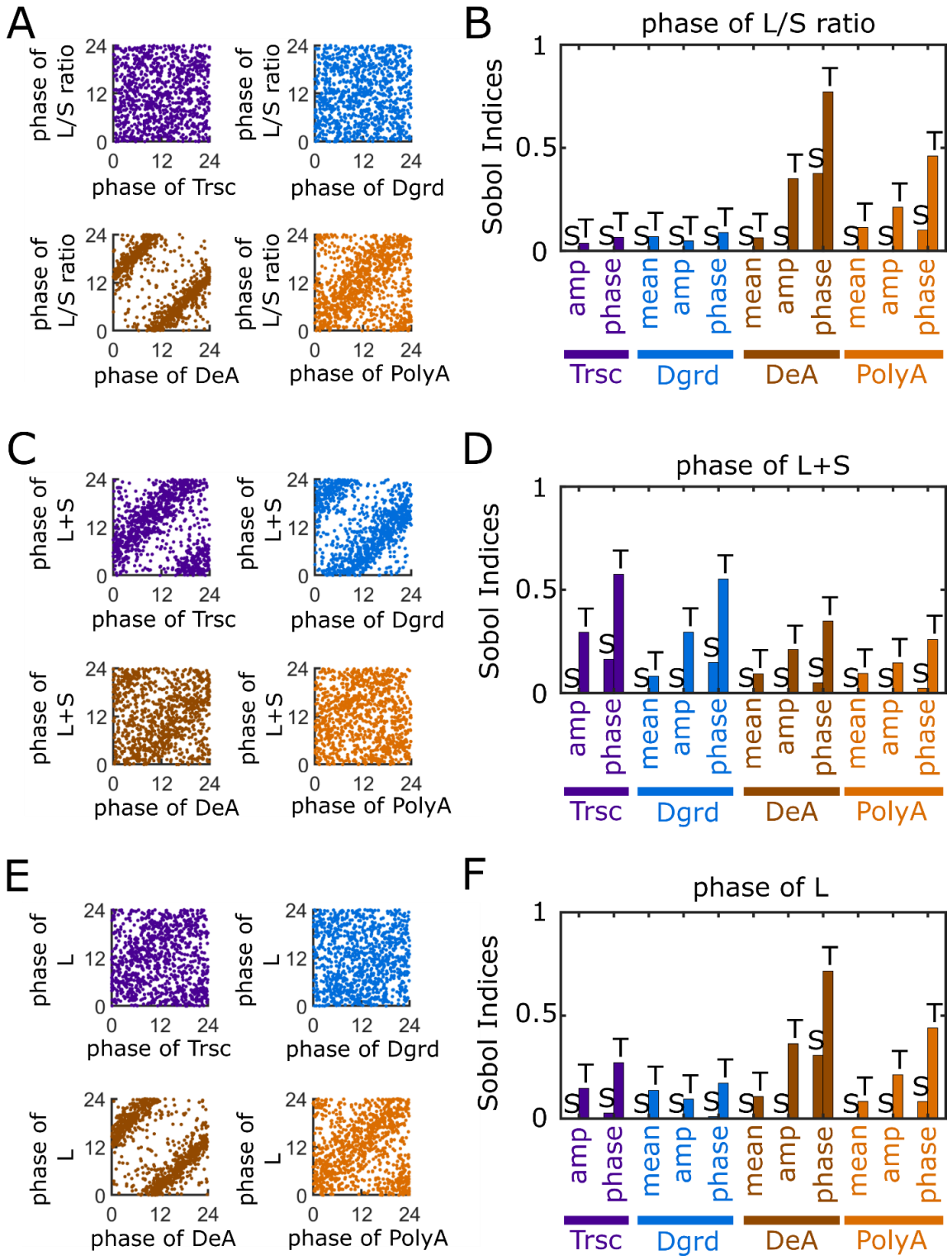

Figure 2. **Rhythmicities of poly(A) tail length and long-tailed mRNA abundance are strongly controlled by rhythmic deadenylation.** (A) Scatter plot of the peak phases of input

processes versus the peak phases of L/S ratio. (B) Sobol indices for the peak phase of L/S ratio (i.e., poly(A) length metric). (C) Scatter plot of the peak phases of input processes versus the peak phases of L+S (i.e., total mRNA abundance). (D) Sobol indices of the peak phase of L+S. (E) Scatter plot of the peak phases of input processes versus the peak phases of L (i.e., long-tailed mRNA abundance). (F) Sobol indices of the peak phase of L. (A, C, E) Each scatter plot shows 10,000 data points randomly chosen from the original simulations for the sake of visual clarity. (B, D, F) Bars with “S” on top: single Sobol indices. Bars with “T” on top: total Sobol indices. Mean values of the Sobol indices are shown, because the variances are too small for clear visualization (**Fig. S2**).

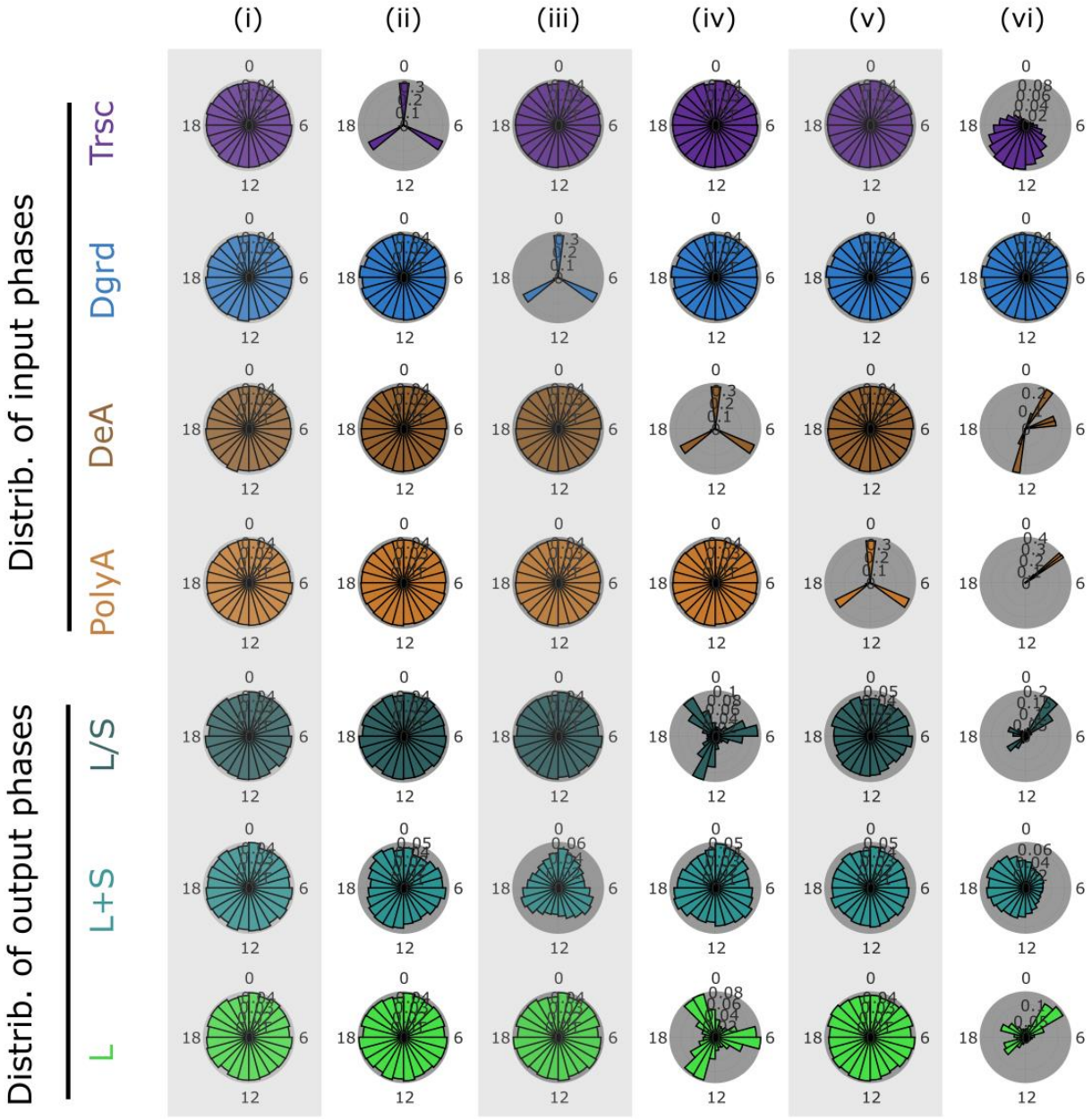

Figure 3. Distinct peak phases in deadenylases group transcripts by their peak phases of poly(A) tail length and long-tailed mRNA abundance (i) Transcription, degradation, deadenylation and polyadenylation phases evenly distributed around the clock. (ii) Transcription phases within three narrow windows at ZT 0, 8, and 16. Degradation, deadenylation and polyadenylation phases evenly distributed around the clock. (iii) Degradation phases within three narrow windows at ZT 0, 8, and 16. Transcription, deadenylation and polyadenylation phases

evenly distributed around the clock. (iv) Deadenylation phases within three narrow windows at ZT 0, 8, and 16. Transcription, degradation and polyadenylation phases evenly distributed around the clock. (v) Polyadenylation phases within three narrow windows at ZT 0, 8, and 16. Transcription, degradation and deadenylation phases evenly distributed around the clock. (vi) Peak phases of transcription follow transcriptome data reported by [13]. Deadenylation phases within three narrow windows at ZT 2, 5, and 13, and polyadenylation phases within one narrow windows at ZT 3.5, based on the data from [25], while degradation phases evenly distributed around the clock. Mean rates and relative amplitudes follow **Table 1** and **Fig. S1**.

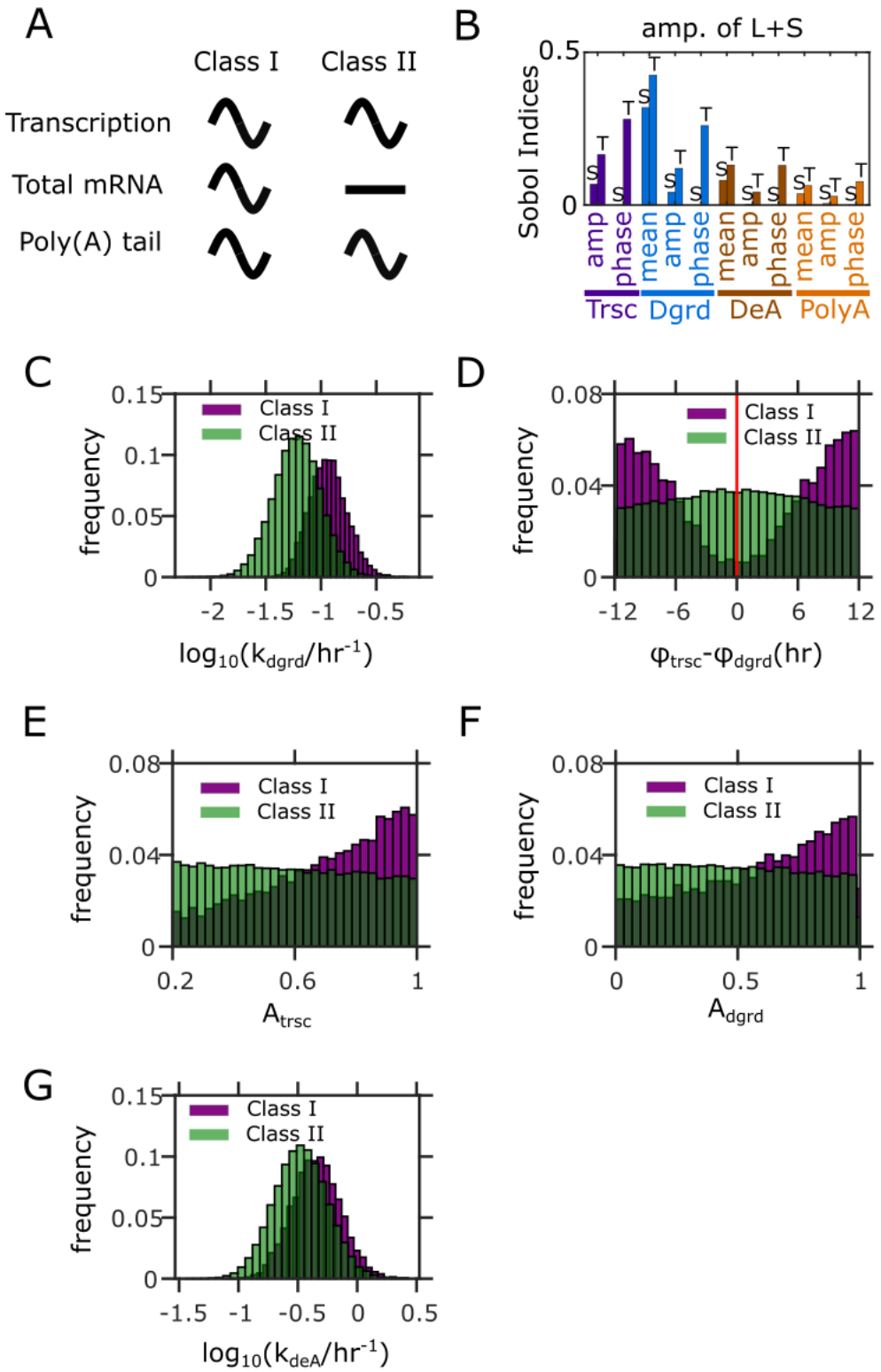

Figure 4. Factors distinguishing between Class I and Class II PAR mRNAs. (A) Characteristics of Class I and Class II PAR mRNAs. (B) Sobol indices for the amplitude of L+S

(i.e., total mRNA abundance). Bars with “S” on top: single Sobol indices. Bars with “T” on top: total Sobol indices. (C) Distributions of mean mRNA degradation rates for the two classes. (D) Distributions of peak phase differences between transcription and degradation for the two classes. The red line corresponds to the peak phase of rhythmic degradation. (E) Distributions of relative amplitudes of transcription for the two classes. (F) Distributions of relative amplitudes of degradation for the two classes. (G) Distribution of mean deadenylation rates for the two classes. Results in (C-G) from 100,000 simulations with parameters randomly sampled according to **Table 1**. Parameter sets with ≥ 0.2 relative amplitude in both L+S and L/S ratio are defined as Class I, while those with < 0.2 relative amplitude in L+S and ≥ 0.2 relative amplitude in L/S ratio are defined as Class II.

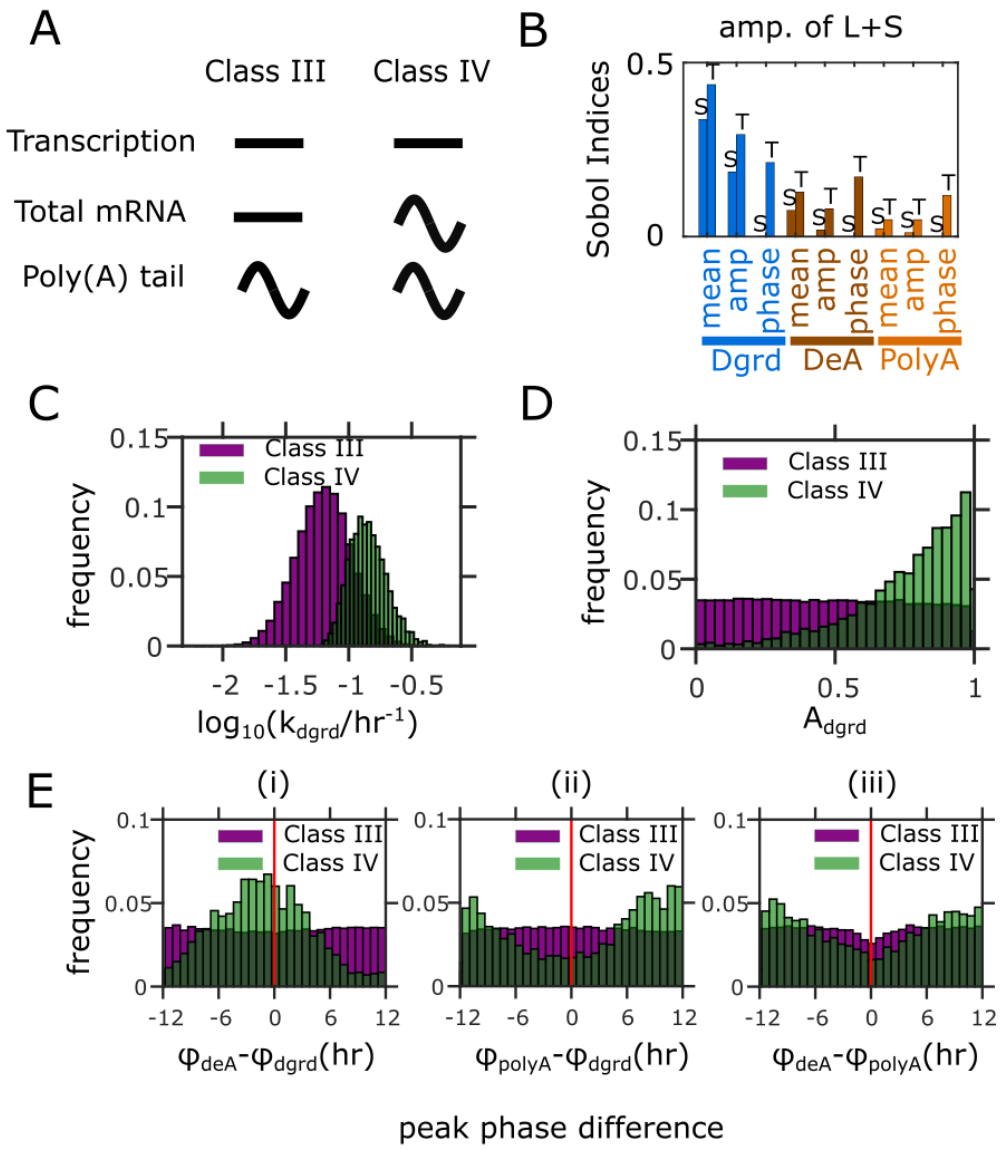

Figure 5. Factors distinguishing between Class III and Class IV PAR mRNAs. (A) Characteristics of Class III and the hypothetical Class IV mRNAs. (B) Sobol indices of the amplitude of L+S (i.e., total mRNA abundance) for the model without rhythmic transcription. Bars with "S" on top: single Sobol indices. Bars with "T" on top: total Sobol indices. (C) Distributions of mean mRNA degradation rates for the two classes. (D) Distributions of relative amplitudes of degradation for the two classes. (E) Distributions of peak phase differences (i) between deadenylation and degradation, (ii) between polyadenylation and degradation, and (iii)

between deadenylation and polyadenylation for the two classes. The red lines correspond to the peak phase of degradation (cases (i) and (ii)) or polyadenylation (case (iii)). Results in (C-E) from 100,000 simulations with parameters randomly sampled according to **Table 1**, but without rhythmic transcription ($A_{\mathrm{trsc}} = 0$). Parameter sets with ≥ 0.2 relative amplitude in L/S ratio and < 0.2 relative amplitude in L+S are defined as Class III, while those with and ≥ 0.2 relative amplitude in both L/S ratio and L+S are defined as Class IV.