

RESEARCH ARTICLE

Temporal precision of regulated gene expression

Shivam Gupta¹, Julien Varennes¹, Hendrik C. Korswagen², Andrew Mugler¹*

- 1 Department of Physics and Astronomy, Purdue University, West Lafayette, Indiana, United States of America, 2 Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences and University Medical Center Utrecht. Utrecht. the Netherlands
- * amugler@purdue.edu





Citation: Gupta S, Varennes J, Korswagen HC, Mugler A (2018) Temporal precision of regulated gene expression. PLoS Comput Biol 14(6): e1006201. https://doi.org/10.1371/journal.pcbi.1006201

Editor: Matthew R Bennett, Rice University, UNITED STATES

Received: December 1, 2017
Accepted: May 14, 2018
Published: June 7, 2018

Copyright: © 2018 Gupta et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data are available at https://github.com/gupta583/Temporal-precision-of-regulated-gene-expression. Data in Fig 3A are from the published study (Mentink et al, Cell 31:188, 2014), whose authors may be contacted at r.korswagen@hubrecht.eu.

Funding: This work was supported by Human Frontier Science Program (www.hfsp.org) grant RGP0030/2016. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Important cellular processes such as migration, differentiation, and development often rely on precise timing. Yet, the molecular machinery that regulates timing is inherently noisy. How do cells achieve precise timing with noisy components? We investigate this question using a first-passage-time approach, for an event triggered by a molecule that crosses an abundance threshold and that is regulated by either an accumulating activator or a diminishing repressor. We find that either activation or repression outperforms an unregulated strategy. The optimal regulation corresponds to a nonlinear increase in the amount of the target molecule over time, arises from a tradeoff between minimizing the timing noise of the regulator and that of the target molecule itself, and is robust to additional effects such as bursts and cell division. Our results are in quantitative agreement with the nonlinear increase and low noise of *mig-1* gene expression in migrating neuroblast cells during *Caenorhabditis elegans* development. These findings suggest that dynamic regulation may be a simple and powerful strategy for precise cellular timing.

Author summary

Cells control important processes with precise timing, even though their underlying molecular machinery is inherently imprecise. In the case of *Caenorhabditis elegans* development, migrating neuroblast cells produce a molecule until a certain abundance is reached, at which time the cells stop moving. Precise timing of this event is critical to *C. elegans* development, and here we investigate how it can be achieved. Specifically, we investigate regulation of the molecule production by either an accumulating activator or a diminishing repressor. Our results are consistent with the nonlinear increase and low noise of gene expression observed in the *C. elegans* cells.

Introduction

Proper timing is crucial for biological processes, including cell division [1–3], cell differentiation [4], cell migration [5], viral infection [6], embryonic development [7, 8], and cell death



Competing interests: The authors have declared that no competing interests exist.

[9]. These processes are governed by molecular events inside cells, i.e., production, degradation, and interaction of molecules. Molecular events are subject to unavoidable fluctuations, because molecule numbers are small and reactions occur at random times [10, 11]. Cells combat these fluctuations using networks of regulatory interactions among molecular species. This raises the fundamental question of whether there exist regulatory strategies that maximize the temporal precision of molecular events and, in turn, cellular behaviors.

A canonical mechanism by which a molecular event triggers a cellular behavior is accumulation to a threshold [3, 4, 12–14]: molecules are steadily produced by the cell, and once the molecule number crosses a particular threshold, the behavior is initiated. The temporal precision of the behavior is therefore bounded by the temporal precision of the threshold crossing. Threshold crossing has been shown to underlie cell cycle progression [3] and sporulation [4], although alternative strategies, such as derivative [9] or integral thresholding [15], oscillation [16], and dynamical transitions in the regulatory network [8], have also been investigated.

Recent work has investigated the impact of auto-regulation (i.e., feedback) on the temporal precision of threshold crossing [12, 13]. Interestingly, it was found that auto-regulation generically decreases the temporal precision of threshold crossing, meaning that the optimal strategy is a linear increase of the molecule number over time with no auto-regulation [12] (although auto-regulation can help if there is a large timescale separation and the threshold itself is also subject to optimization [13]). Indeed, even when the molecule also degrades, the optimal precision is achieved when positive auto-regulation counteracts the effect of degradation, preserving the linear increase over time [12]. However, in many biological processes, such as the temporal control of neuroblast migration in *Caenorhabditis elegans* [5], the molecular species governing the behavior increases nonlinearly over time. This suggests that other regulatory interactions beyond auto-regulation may play an important role in determining temporal precision. In particular, the impact of activation and repression on temporal precision, where the activator or repressor has its own stochastic dynamics, remains unclear.

Here we investigate the temporal precision of threshold crossing for a molecule that is regulated by either an accumulating activator or a degrading repressor. Using a first-passage-time approach [12, 17-19] and a combination of computational and analytic methods, we find that, unlike in the case of auto-regulation, the presence of either an activator or a repressor increases the temporal precision beyond that of the unregulated case. Furthermore, the optimal regulatory strategy for either an activator or a repressor corresponds to a nonlinear increase in the regulated molecule number over time. We elucidate the mechanism behind these optimal strategies, which stems from a tradeoff between reducing the noise of the regulator and that of the target molecule, and is similar to the fact that a sequence of time-ordered stochastic events becomes more precisely timed with more events. These findings are robust to more complex features of the regulation process, including bursts of molecule production, more complex regulator dynamics, and cell division. Our results are quantitatively consistent with both the temporal precision and nonlinearity of the mig-1 mRNA dynamics of the migrating neuroblast cells in C. elegans larvae [5]. The agreement of our simple model with these data suggests that many molecular timing processes may benefit from the generic regulatory strategies we identify here.

Results

We consider a molecular species X whose production is regulated by a second species, either an activator A or a repressor R (Fig 1A). The regulator undergoes its own dynamics: the activator undergoes pure production at a zeroth-order rate k whereas the repressor undergoes pure degradation at a first-order rate μ , such that in either case the production rate of X increases

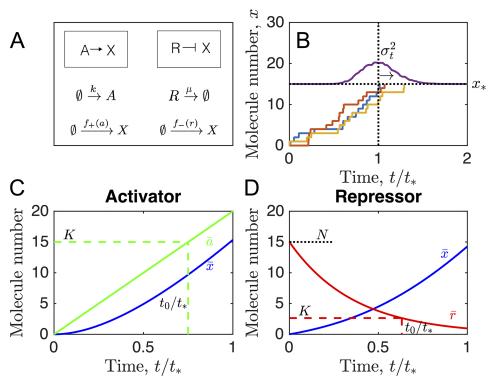


Fig 1. Threshold crossing of a regulated molecular species. (A) A target species X is regulated by either an accumulating activator A or a degrading repressor R. (B) Temporal precision is quantified by the variance σ_t^2 of the first-passage time, at which the stochastic molecule number x first crosses the threshold x_* . (C, D) Deterministic dynamics illustrate the effects of regulation. Parameters are $kt_* = 20$ and K = 15 in C; $\mu t_* = 2.75$, K = 2.6, and N = 15 in B and D; and $x_* = 15$ and H = 1 throughout. t_0 is defined by $\bar{a}(t_0) = K$ in C and $\bar{r}(t_0) = K$ in D.

https://doi.org/10.1371/journal.pcbi.1006201.g001

over time. The activator does not degrade and the repressor is not produced, although we later relax this assumption. For the regulation function we take a Hill function, which is a generic model of cooperative regulation [12, 13, 20],

$$f_{+}(a) = \frac{\alpha a^{H}}{a^{H} + K^{H}}$$
 (activator), (1)

$$f_{-}(r) = \frac{\alpha K^{H}}{r^{H} + K^{H}}$$
 (repressor). (2)

Here a and r are the molecule numbers of A and R, respectively, α is the maximal production rate of X, K is the half-maximal regulator number, and H is the cooperativity. First we neglect additional complexities such as bursts of production, more complex regulator dynamics, cell division, auto-regulation, longer regulatory cascades, or transcriptional delay. Later we check the robustness of our results to bursts, more complex regulator dynamics, and cell division, and we speculate upon the effects of auto-regulation, longer regulatory cascades, and delay in the Discussion.

We suppose that a behavior is initiated when the molecule number x crosses a threshold x_* (Fig 1B). Because the production of X and the dynamics of the regulator are stochastic, the time at which x first reaches x_* is a random variable. We characterize the precision of this event by the mean \bar{t} and variance σ_t^2 of this first-passage time, which we compute numerically from the master equation corresponding to the reactions in Fig 1A (see Materials and methods). The

maximal production rate α is set to ensure that \bar{t} is equal to a target time t_* , which we assume is set by functional constraints on the initiated behavior. This leaves k, K, and H as free parameters of the regulation (with α a function of these parameters). In principle, these parameters can be optimized to minimize the timing variance σ_t^2 .

The deterministic dynamics, illustrated in Fig 1C and 1D, neglect fluctuations but give an intuitive picture of the regulation. Whereas the amount of activator increases linearly over time, the amount of repressor decays exponentially from an initial molecule number *N*:

$$\bar{a}(t) = kt, \tag{3}$$

$$\bar{r}(t) = Ne^{-\mu t}. (4)$$

In either case, the production rate f_{\pm} of X increases over time, such that \bar{x} increases nonlinearly. N is an additional free parameter in the repressor case.

Regulation increases temporal precision

To investigate the effects of regulation on temporal precision, we consider the timing variance σ_t^2 as a function of the parameters k and K, or μ and K. The special case of no regulation corresponds to the limits $k\to\infty$ and $K\to0$ in the case of activation, or $\mu\to\infty$ and $K\to\infty$ in the case of repression. In this case, the production of X occurs at the constant rate α . Reaching the threshold requires x_* sequential events, each of which occurs in a time that is exponentially distributed with mean $1/\alpha$. The total completion time for such a process is given by a gamma distribution with mean $\overline{t}=x_*/\alpha$ and variance $\sigma_t^2=x_*/\alpha^2$ [19]. Ensuring that $\overline{t}=t_*$ requires $\alpha=x_*/t_*$, for which the variance satisfies $\sigma_t^2x_*/t_*^2=1$. This expression gives the timing variance for the unregulated process.

In Fig 2 we plot the scaled variance $\sigma_t^2 x_* / t_*^2$ as a function of the parameters k and K, or μ and K, for cooperativity H=3 (color maps). In the case of activation (Fig 2A), the variance decreases with increasing k and K. This means that the temporal precision is highest for an activator that accumulates quickly and requires a high abundance to produce K. In the case of repression (Fig 2B), the variance has a global minimum as a function of μ and K. This means that the temporal precision is highest for a repressor with a particular well-defined degradation rate and abundance threshold. Importantly, we see that for both activation and repression, the scaled variance can be less than one, meaning that regulation allows improvement of the

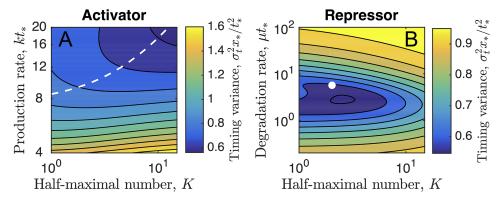


Fig 2. Optimal regulatory strategies. Timing variance as a function of the regulatory parameters reveals (A) a trajectory along which the variance decreases in the case of the activator and (B) a global minimum in the case of the repressor. White dashed line in A and white dot in B show the analytic approximations in Eqs 9 and 11, respectively. Parameters are N = 15 in B, and $x_* = 15$ and H = 3 in both.

https://doi.org/10.1371/journal.pcbi.1006201.g002



temporal precision beyond that of the unregulated process. We have checked that this result holds for $H \ge 1$.

Optimal regulation balances regulator and target noise

To understand the dependencies in Fig 2, we develop analytic approximations. First, we assume that $H \to \infty$, such that the regulation functions in Eqs 1 and 2 become threshold functions. In this limit, the production rate of X is zero if a < K or r > K, and α otherwise. The deterministic dynamics of X become piecewise-linear,

$$\bar{x}(t) = \begin{cases} 0 & t < t_0 \\ \alpha(t - t_0) & t \ge t_0, \end{cases}$$
 (5)

where t_0 is determined by either $\bar{a}(t_0) = K$ or $\bar{r}(t_0) = K$ according to Eqs 3 and 4. Then, to set α , we use the condition $\bar{x}(t_*) = x_*$, which results in $\alpha = x_*/(t_* - t_0)$.

Lastly, we approximate the variance in the first-passage time using the variance in the molecule number and the time derivative of the mean dynamics [13]. Specifically, the timing variance of X arises from two sources: (i) uncertainty in the time when the regulator crosses its threshold K, which determines when the production of the target X begins, and (ii) uncertainty in the time when X crosses its threshold X_* , given that production begins at a particular time. The first source is regulator noise, and the second source is target noise. We estimate these timing variances from the associated molecule number variances, propagated via the time derivatives,

$$\sigma_t^2 \approx \sigma_y^2 \left(\frac{d\bar{y}}{dt} \right)^{-2} \Big|_{t_0} + \sigma_x^2 \left(\frac{d\bar{x}}{dt} \right)^{-2} \Big|_{t_*}, \tag{6}$$

where $y \in \{a, r\}$ denotes the regulator molecule number.

For the activator, which undergoes a pure production process with rate k, the molecule number obeys a Poisson distribution with mean kt. Therefore, the molecule number variance at time t_0 is $\sigma_a^2 = kt_0$. For the repressor, which undergoes a pure degradation process with rate μ starting from N molecules, the molecule number obeys a binomial distribution with number of trials N and success probability $e^{-\mu t}$. Therefore, the molecule number variance at time t_0 is $\sigma_r^2 = Ne^{-\mu t_0}(1-e^{-\mu t_0})$. For the target molecule, which undergoes a pure production process with rate α starting at time t_0 , the molecule number obeys a Poisson distribution with mean α ($t-t_0$). Therefore, the molecule number variance at time t_* is $\sigma_x^2 = \alpha(t_* - t_0)$. Inserting these expressions into Eq.6, along with the derivatives calculated from Eqs.3–5 and the appropriate expressions for α and t_0 , we obtain

$$\frac{\sigma_{l}^{2} x_{*}}{t_{*}^{2}} \approx \frac{K x_{*}}{\left(k t_{*}\right)^{2}} + \left(1 - \frac{K}{k t_{*}}\right)^{2} \quad \text{(activator)},\tag{7}$$

$$\frac{\sigma_t^2 x_*}{t_*^2} \approx \frac{(N-K)x_*}{NK(\mu t_*)^2} + \left[1 - \frac{\log(N/K)}{\mu t_*}\right]^2 \quad \text{(repressor)}. \tag{8}$$

As a function of kt_* and K, the global minimum of Eq 7 occurs as $kt_* \to \infty$ and $K \to \infty$. The path of descent toward this minimum is given by differentiating with respect to K at fixed kt_*

and setting the result to zero, which yields the curve

$$K = \begin{cases} 0 & kt_* < \frac{x_*}{2} \\ kt_* - \frac{x_*}{2} & kt_* \ge \frac{x_*}{2}, \end{cases}$$
 (9)

along which the variance satisfies

$$\frac{\sigma_t^2 x_*}{t_*^2} = \begin{cases} 1 & kt_* < \frac{x_*}{2} \\ \frac{x_*}{kt_*} \left(1 - \frac{x_*}{4kt_*} \right) & kt_* \ge \frac{x_*}{2}, \end{cases}$$
(10)

where the first case comes from the fact that K must be nonnegative. In contrast, the global minimum of Eq.8 occurs at finite μt_* and K: differentiating with respect to each and setting the results to zero gives the values

$$K = e^{-2}N, (11a)$$

$$\mu t_* = \frac{e^2 x_*}{2N} + 2,\tag{11b}$$

$$\frac{\sigma_t^2 x_*}{t_*^2} = \frac{x_*}{x_* + 4e^{-2}N},\tag{12}$$

where we have assumed that $K/N \ll 1$ (see Materials and methods), which is justified post-hoc by Eq 11a.

These analytic approximations are compared with the numerical results for the activator in Fig 2A (white dashed line, Eq 9) and for the repressor in Fig 2B (white circle, Eq 11). In Fig 2A we see that the global minimum indeed occurs as $kt_* \to \infty$ and $K \to \infty$, and the predicted curve agrees well with the observed descent. In Fig 2B we see that the predicted global minimum lies very close to the observed global minimum. We have also checked along specific slices in the (K, kt_*) or $(K, \mu t_*)$ plane and found that the analytic approximations generally differ from the numerical results by about 10% or less, despite the fact that the approximations take $H \to \infty$ whereas the numerics in Fig 2 use H = 3.

The success of the approximations means that Eq 6 describes the key mechanism leading to the optimal temporal precision. Eq 6 demonstrates that the optimal regulatory strategy arises from a tradeoff between minimizing regulator and target noise. On the one hand, minimizing only the regulator noise would require that the regulator cross its threshold K with a steep slope $d\bar{y}/dt$ and therefore at an early time, meaning that the target molecule would be effectively unregulated and would increase linearly over time. On the other hand, minimizing only the target noise would require that the regulator cross its threshold only shortly before the target time t_* , such that the target molecule would cross its threshold x_* with a steep slope $d\bar{x}/dt$, leading to a highly nonlinear increase of the target molecule over time. In actuality, the optimal strategy is somewhere in between, with the regulator crossing its threshold at some intermediate time t_0 , and the target molecule exhibiting moderately nonlinear dynamics as in Fig 1C and 1D.

Eqs 10 and 12 demonstrate that the timing variance is small for large kt_*/x_* in the case of activation, and small for large N/x_* in the case of repression. This makes intuitive sense because each of these quantities scales with the number of regulator molecules: k is the production rate of activator molecules, while N is the initial number of repressor molecules. To make this



intuition quantitative, we define a cost as the time-averaged number of regulator molecules,

$$\langle a \rangle = \frac{1}{t_*} \int_0^{t_*} dt \ \bar{a}(t) = \frac{1}{2} k t_*,$$
 (13)

$$\langle r \rangle = \frac{1}{t_{\circ}} \int_{0}^{t_{*}} dt \ \bar{r}(t) = \frac{N}{\mu t_{\circ}} (1 - e^{-\mu t_{*}}),$$
 (14)

where the second steps follow from Eqs 3 and 4. We see that, indeed, $\langle a \rangle$ scales with k, and $\langle r \rangle$ scales with N. Thus, Eqs 10 and 12 demonstrate that increased temporal precision comes at a cost, in terms of the number of regulator molecules that must be produced.

Model predictions are consistent with neuroblast migration data

We test our model predictions using data from neuroblast cells in *C. elegans* larvae [5]. During *C. elegans* development, particular neuroblast cells migrate from the posterior to the anterior of the larva. It has been shown that the migration terminates not at a particular position, but rather after a particular amount of time, and that the termination time is controlled by a temporal increase in the expression of the *mig-1* gene [5]. Since *mig-1* is known to be subject to regulation [21], we investigate the extent to which the dynamics of *mig-1* can be explained by the predictions of our model.

Fig 3A shows the number x of mig-1 mRNA molecules per cell as a function of time t, obtained by single-molecule fluorescent in situ hybridization (from [5]). We analyze these data in the following way (see Materials and methods for details). First, noting that the dynamics are nonlinear, we quantify the linearity using the area under the curve, normalized by that for a perfectly linear trajectory $x_*t_*/2$,

$$\rho = \frac{2}{x_* t_*} \int_0^{t_*} dt \ x(t). \tag{15}$$

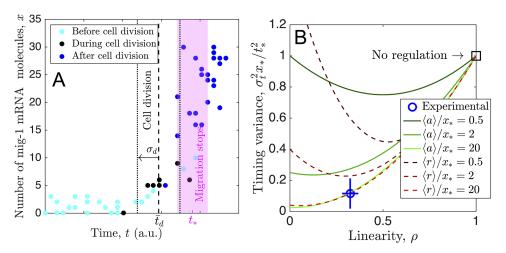


Fig 3. Model predictions agree with neuroblast migration data. (A) Number of mig-1 mRNA molecules per cell as a function of time t, obtained by single-molecule fluorescent in situ hybridization, from [5]. Magenta shows approximate range of times when cell migration terminates. Black lines show mean \bar{t}_d (dashed) and standard deviation σ_d of cell division times (black points). (B) Timing variance vs. linearity of x(t), both for experimental data in A (blue circle) and our model (curves, Eqs 16 and 17). Data analyzed using ranges of threshold $10 \le x_* \le 25$ and bin size $3 \le \Delta x \le 12$; error bars show standard deviations of these results. We see that for sufficiently large cost $\langle a \rangle / x_*$ or $\langle r \rangle / x_*$, model predictions agree with experimental data point.

https://doi.org/10.1371/journal.pcbi.1006201.g003

By this definition, $\rho = 1$ for perfectly linear dynamics, and $\rho \to 0$ for maximally nonlinear dynamics (a sharp rise at t_*). Then, we estimate x_* , t_* , and the timing variance σ_t^2 from the data. Specifically, migration is known to terminate between particular reference cells in the larva [5], which gives an estimated range for the termination time t_* . This range is shown in magenta in Fig 3A and corresponds to a threshold within the approximate range $10 \le x_* \le 25$. Therefore, we divide the x axis into bins of size Δx , choose bin midpoints x_* within this range, and for each choice compute the mean t_* and the variance σ_t^2 of the data in that bin. Fig 3B shows the average and standard deviation of results for different values of x_* and Δx (blue circle).

The experimental data point in Fig 3B exhibits two clear features: (i) the dynamics are non-linear (ρ is significantly below 1), and (ii) the timing variance is low ($\sigma_t^2 x_*/t_*^2$ is significantly below 1). Neither feature can be explained by a model in which the production of x is unregulated, since that would correspond to a linear increase of molecule number over time ($\rho = 1$) and a timing variance that satisfies $\sigma_t^2 x_*/t_*^2 = 1$ (square in Fig 3B). Furthermore, since autoregulation has been shown to generically increase timing variance beyond the unregulated case [12], it is unlikely that feature (ii) can be accounted for by a model with auto-regulation alone. Can these data be accounted for by our model with regulation?

To address this question we calculate ρ and $\sigma_t^2 x_*/t_*^2$ from our model. For simplicity, we focus on the analytic approximations in Eqs 7 and 8, since they have been validated in Fig 2. In these approximations, since $\bar{x}(t)$ is piecewise-linear (Eq 5), calculating ρ via Eq 15 is straightforward: $\rho = 1 - t_0/t_*$, where t_0 is once again determined by either $\bar{a}(t_0) = K$ or $\bar{r}(t_0) = K$ according to Eqs 3 and 4. For a given ρ and cost $\langle a \rangle/x_*$ or $\langle r \rangle/x_*$, we calculate the minimum timing variance $\sigma_t^2 x_*/t_*^2$. For the activator, we use the expression for ρ along with Eq 13 to write Eq 7 in terms of ρ and $\langle a \rangle/x_*$,

$$\frac{\sigma_t^2 x_*}{t_*^2} = \frac{x_*}{2\langle a \rangle} (1 - \rho) + \rho^2.$$
 (16)

For the repressor, we use the expression for ρ along with Eq 14 to write Eq 8 in terms of ρ and $\langle r \rangle / x_*$, and then minimize over N (see Materials and methods) to obtain

$$\frac{\sigma_t^2 x_*}{t_*^2} = \frac{e^3}{27} \frac{x_*}{\langle r \rangle} (1 - \rho)^3 + \rho^2.$$
 (17)

Eqs 16 and 17 are shown in Fig 3B (green solid and red dashed curves, respectively), and we see the same qualitative features for both cases: all curves satisfy $\sigma_t^2 x_*/t_*^2 = 1$ at $\rho = 1$, as expected; and as ρ decreases, each curve exhibits a minimum whose depth and location depend on cost. Specifically, as cost increases (lighter shades of green or red), the variance decreases, as expected. Importantly, we see that at a cost on the order of $\langle a \rangle/x_* = \langle r \rangle/x_* \sim 10$, the model becomes consistent with the experimental data: both the low timing variance and the low degree of linearity predicted by either the activator or repressor case agree quantitatively with the experiment. This suggests that either an accumulating activator or a degrading repressor is sufficient to account for the temporal precision observed in *mig-1*-controlled neuroblast migration.

Results are robust to additional complexities including cell division

Our minimal model neglects common features of gene expression such as bursts in molecule production [22] and additional sources of noise. Therefore we test the robustness of our findings to these effects here. First, we test the robustness of the results to the presence of bursts by replacing the Poisson process governing the activator production with a bursty production process. Specifically, we assume that each production event increases the activator molecule

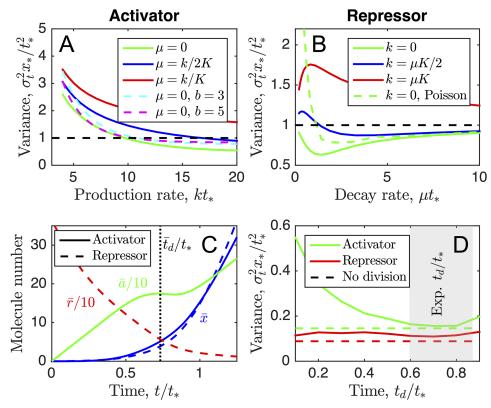


Fig 4. Results are robust to additional complexities including cell division. (A, B) Green solid curves show slices from Fig 2 with K=10 while black dashed line shows unregulated limit $\sigma_r^2 x_*/t_*^2=1$. We see that regulation can reduce timing variance even with bursts in activator production of mean size b (A, cyan and magenta dashed), initial Poisson noise in repressor number (B, green dashed), or steady state k/μ in regulator dynamics (blue) unless it approaches regulation threshold K (red). (C) Mean dynamics of activator model (solid) and repressor model (dashed) in which cell division occurs at time \bar{t}_d on average. Abrupt reductions in molecule numbers are smoothed by noise in t_d and by binomial partitioning of molecules. (D) Timing variance approaches that with no division (dashed) within experimental division region (gray). In A and B, parameters are as in Fig 2. In C and D, parameters are $x^*=15$, $\langle a \rangle/x_*=\langle r \rangle/x_*=10$, and H=3, with kt_* , μt_* , and K set to optimal values (Fig 2) and \bar{t}_d and σ_d set to experimental values. In all cases, α is set to ensure that mean threshold crossing time \bar{t} equals t_* .

https://doi.org/10.1371/journal.pcbi.1006201.g004

count by an integer in $[1, \infty)$ drawn from a geometric distribution with mean b [23, 24]. The limiting case of b=1 recovers the original Poisson process. The results are shown in Fig 4A for b=1, 3, and 5 (green solid, cyan dashed, and magenta dashed curves). We see that bursts in the activator increase the timing variance of the target molecule, as expected, but that there remain parameters for which the variance is less than that for the unregulated case, $\sigma_t^2 x_*/t_*^2 = 1$ (dashed black line). This result shows that even with bursts, regulation by an accumulating activator is beneficial for timing precision.

We also recognize that whereas the activator can be assumed to start with exactly zero molecules, it is more realistic for the repressor to start with an initial number of molecules that has its own variability. We incorporate this additional variability into the model by performing stochastic simulations [25] of the reactions in Fig 1A and drawing the initial repressor molecule number from a Poisson distribution across simulations. The result is shown by the green dashed curve in Fig 4B. We see that the additional variability gives rise to an increase in the timing variance of the target molecule, as expected (compare with the green solid curve). However, for most of the range of degradation rates, including the optimal degradation rate, the



variance remains less than that of the unregulated case, $\sigma_t^2 x_* / t_*^2 = 1$ (dashed black line). This result indicates that the benefit of repression is robust to this additional source of noise.

Then, we test the robustness of the results to our assumptions that the activator undergoes pure production and the repressor undergoes pure degradation. Specifically, we introduce a degradation rate μ for the activator, and a production rate k for the repressor, such that either regulator reaches a steady state of k/μ . The blue curves in Fig 4A and 4B show the case where the increasing activator's steady state k/μ is twice its regulation threshold K, or the decreasing repressor's steady state k/μ is half its regulation threshold K, respectively. In both cases, we see that the timing variance of the target molecule increases because the regulator slows down on the approach to its regulation threshold. Nonetheless, we see that it is still possible for the variance to be lower than that of the unregulated case. The red curves show the case where the regulator's steady state is equal to its regulation threshold. Here we are approaching the regime in which threshold crossing is an exponentially rare event. As a result, the variance further increases, to the point where it is above that of the unregulated case for the full range of parameters shown. These results demonstrate that the benefit of regulation is robust to more complex regulator dynamics, but only if the regulator still crosses its regulation threshold at a reasonable mean velocity.

Finally, we test the robustness of the results to a feature exhibited by the experimental mig-1 data: near the end of migration, cell division occurs (Fig 3A, black data). One daughter cell continues migrating (Fig 3A, dark blue data), while the other undergoes programmed cell death [5]. To investigate the effects of cell division, we perform stochastic simulations, and at a given time t_d we assume that the cell volume V is reduced by a factor of two. For each simulation, t_d is drawn from a Gaussian distribution with mean \overline{t}_d and variance σ_d^2 determined by the data (Fig 3A, black). At t_d , we reduce the molecule numbers of both the regulator and the target molecule assuming symmetric partitioning, such that the molecule number after division is drawn from a binomial distribution with total number of trials equal to the molecule number before division and success probability equal to one half. We also reduce the molecule number threshold K by a factor of two because it is proportional to the cell volume via $K = K_d V$, where K_d is the dissociation constant.

Fig 4C shows the dynamics of the mean molecule numbers of the activator (green solid) and its target (blue solid), or the repressor (red dashed) and its target (blue dashed). We see that the activator, repressor, and target drop in molecule number at division but that the abruptness of the drop is smoothed by the variability in the division time. The smoothing is more pronounced in the cases of the repressor and the target because the molecule numbers of these species are smaller at division. Thus, for either the activator or repressor mechanism, we see that the experimentally observed variability in division time is sufficient to smooth out the dynamics of the target molecule number, consistent with the experimental data in Fig 3A.

Additionally, we see in Fig 4D that the timing variance of the target molecule in both the activator and repressor cases is similar to that without division in the region of the experimental division time. This further indicates that either model remains sufficient to account for the low experimental timing variance, even with the experimentally observed cell division. Taken together, the results in Fig 4C and 4D show that the key results of the model are robust to the effects of cell division.

Discussion

We have demonstrated that regulation by an accumulating activator or a diminishing repressor increases the precision of threshold crossing by a target molecule, beyond the precision achievable with constitutive expression alone. The increase in precision results from a tradeoff



between reducing the noise of the regulator and reducing the noise of the target molecule itself. Our minimal model is sufficient to account for both the high degree of nonlinearity and the low degree of noise in the dynamics of *mig-1* in *C. elegans* neuroblasts, providing evidence for the hypothesis that these cells use regulated expression to terminate their migration with increased temporal precision. These results suggest that regulation by a dynamic upstream species is a simple and generic method of increasing the temporal precision of cellular behaviors governed by threshold-crossing events.

Why does regulation increase temporal precision, whereas it has been shown that auto-regulation (feedback) does not [12]? After all, either regulation or positive feedback can produce an acceleration in molecule number over time, leading to a steeper threshold crossing. The reason is likely that positive feedback relies on self-amplification. In addition to amplifying the mean, positive feedback also amplifies the noise. In contrast, regulation by an external species does not involve self-amplification. Therefore, the noise increase is not as strong. The target molecule certainly inherits noise from the regulator (Eq 6), but the increase in noise does not outweigh the benefit of the acceleration, as it does for feedback. Future work could investigate the interplay of regulation and feedback, as well as active degradation of the target molecule, especially as *mig-1* is thought to be subject to feedback and degradation in addition to external regulation [5, 21].

Our finding that regulation increases temporal precision is related to the more basic phenomenon that a sequence of ordered events has a lower relative timing error than a single event [19, 26]. Specifically, if a single event occurs in a time that is exponentially distributed with a mean τ , then the relative timing error is $\sigma/\tau=1$. For n such events that must occur in sequence, the total completion time follows a gamma distribution with relative timing error $\sigma/\tau=1/\sqrt{n}$, which decreases with increasing n. Thus, at a coarse-grained level, the addition of a regulator can be viewed as increasing the length of the sequence of threshold-crossing events from one to two, and thus decreasing the timing error. This perspective suggests that the timing error could be decreased even further via a cascade of regulators.

Although we have demonstrated that our findings are robust to complexities such as bursts and cell division (Fig 4), our model neglects additional features of regulated gene expression such as transcriptional delay. Transcriptional delay has been shown to play an important role in regulation [27, 28] and to have consequences for the mean and variance of threshold-crossing times [29]. If a delay were to arise due to a sequence of stochastic but irreversible steps, then we conjecture that the relative timing error would decrease with the number of these steps, due to the same cascading mechanism mentioned in the previous paragraph. However, it has been shown that if there are reversible steps or cycles within a multistep process, then the first passage time distribution can approach an exponential as the number of steps becomes large [26]. In this case the timing statistics would be captured by our simple model, which assumes single exponentially distributed waiting times. Future work could explore the effects of transcriptional delay in more detail.

Finally, we have shown that the *mig-1* data from migrating neuroblasts in *C. elegans* are quantitatively consistent with either the accumulating activator or diminishing repressor model, but the data do not distinguish between the two models. A direct approach to search for a distinction would be to use genetic knockout techniques to screen directly for regulators of *mig-1* and their effects on its abundance. A less direct approach would be to more closely investigate the effects of the cell division that occurs during migration. For example, we assumed in this study that the volume fraction is equal to the average molecule number fraction in the surviving cell after division. However, if they were found to be unequal for either *mig-1* or its regulator(s), then the concentrations of these species could undergo an abrupt



change after division, which may have opposing consequences for the activator vs. the repressor mechanism. Future studies could use these or related approaches to more concretely identify the role of gene regulation in achieving precise timing during cellular processes.

Materials and methods

Computation of the first-passage time statistics

We compute the first-passage time statistics \bar{t} and σ_t^2 numerically from the master equation following [12], generalized to a two-species system. Specifically, the probability F(t) that the molecule number crosses the threshold x_* at time t is equal to the probability P_y , $x_*-1(t)$ that there are y regulator molecules (where $y \in \{a, r\}$) and $x_* - 1$ target molecules, and that a production reaction occurs with rate $f_\pm(y)$ to bring the target molecule number up to x_* . Since this event can occur for any regulator molecule number y, we sum over all y,

$$F(t) = \sum_{y=0}^{Y} f_{\pm}(y) P_{y,x_*-1}(t), \tag{18}$$

where $Y = \{a_{\text{max}}, N\}$. The repressor has a maximum number of molecules N, whereas the activator number is unbounded, and therefore we introduce the numerical cutoff $a_{\text{max}} = kt_* + \sqrt{10kt_*}$. The probability P_{yx} evolves in time according to the master equation corresponding to the reactions in Fig 1A,

$$\dot{P}_{ax} = kP_{a-1,x} + f_{+}(a)P_{a,x-1} - [k + f_{+}(a)]P_{ax}, \tag{19a}$$

$$\dot{P}_{rx} = \mu(r+1)P_{r+1,x} + f_{-}(r)P_{r,x-1} - [\mu r + f_{-}(r)]P_{rx}.$$
(19b)

The moments of Eq 18 are

$$\langle t^m \rangle = \sum_{y=0}^{Y} f_{\pm}(y) \int_0^{\infty} dt \ t^m P_{y,x_*-1}(t),$$
 (20)

where $\bar{t} = \langle t \rangle$ and $\sigma_t^2 = \langle t^2 \rangle - \langle t \rangle^2$. Therefore computing \bar{t} and σ_t^2 requires solving for the dynamics of P_{vx} .

Because Eq 19 is linear in P_{yx} , it is straightforward to solve by matrix inversion. We rewrite P_{yx} as a vector by concatenating its columns, $\vec{P}^{\top} = [[P_{00}, \dots, P_{Y0}], \dots, [P_{0,x_*-1}, \dots, P_{Y,x_*-1}]]$, such that Eq 19 becomes $\dot{\vec{P}} = \mathbf{M}\vec{P}$, where

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}^{(1)} & & & & & & \\ \mathbf{M}^{(2)} & \mathbf{M}^{(1)} & & & & & \\ & \mathbf{M}^{(2)} & \mathbf{M}^{(1)} & & & & \\ & & \ddots & \ddots & & & \\ & & & \mathbf{M}^{(2)} & \mathbf{M}^{(1)} \end{bmatrix}.$$
(21)

Here, for $i,j\in\{0,\ldots,Y\}$, the x_*-1 subdiagonal blocks are the diagonal matrix $\mathbf{M}_{ij}^{(2)}=f_\pm(i)\delta_{ij}$, and the x_* diagonal blocks are the subdiagonal matrix $\mathbf{M}_{ij}^{(1)}=-[k(1-\delta_{ia_{\max}})+f_+(i)]\delta_{ij}+k\delta_{i-1,j}$ or the superdiagonal matrix $\mathbf{M}_{ij}^{(1)}=-[\mu i+f_-(i)]\delta_{ij}+\mu(i+1)\delta_{i+1,j}$ for the activator or repressor case, respectively. The $\delta_{ia_{\max}}$ term prevents activator production beyond a_{\max} molecules. The final $\mathbf{M}^{(1)}$ matrix in Eq 21 contains f_\pm production terms that are not balanced by equal and

opposite terms anywhere in their columns. These terms correspond to the transition from $x_* - 1$ to x_* target molecules, for which there is no reverse transition. Therefore, the state with x_* target molecules (and any number of regulator molecules) is an absorbing state that is outside the state space of \vec{P} [12]. Consequently, probability leaks over time, and $\vec{P}(t \to \infty) = \vec{\emptyset}$. Equivalently, the eigenvalues of \mathbf{M} are negative.

The solution of the dynamics above $\underline{\text{Eq }21}$ is $\vec{P}(t) = \exp(\mathbf{M}t)\vec{P}(0)$. Therefore, $\underline{\text{Eq }20}$ becomes $\langle t^m \rangle = \vec{V}^\top [\int_0^\infty dt \ t^m \exp(\mathbf{M}t)] \vec{P}(0)$, where \vec{V}^\top is a row vector of length $x_*(Y+1)$ consisting of $[f_\pm(0), \ldots, f_\pm(Y)]$ preceded by zeros. We solve this equation via integration by parts [12], noting that the boundary terms vanish because the eigenvalues of \mathbf{M} are negative, to obtain

$$\langle t^m \rangle = (-1)^{m+1} m! \vec{V}^{\top} (\mathbf{M}^{-1})^{m+1} \vec{P}(0).$$
 (22)

We see that computing $\bar{t} = \langle t \rangle$ and $\sigma_t^2 = \langle t^2 \rangle - \langle t \rangle^2$ requires inverting **M**, which we do numerically in Matlab. We initialize \vec{P} as $P_{ax}(0) = \delta_{a0} \, \delta_{x0}$ or $P_{rx}(0) = \delta_{rN} \, \delta_{x0}$ for the activator or repressor case, respectively.

When including cell division, we compute \bar{t} and σ_t^2 from 50,000 stochastic simulations [25].

Deterministic dynamics

The dynamics of the mean regulator and target molecule numbers are obtained by calculating the first moments of Eq 19, $\partial_t \bar{y} = \sum_{yx} y \dot{P}_{yx}$ and $\partial_t \bar{x} = \sum_{yx} x \dot{P}_{yx}$, where $y \in \{a, r\}$. For the regulator we obtain $\partial_t \bar{a} = k$ or $\partial_t \bar{r} = -\mu \bar{r}$ in the activator or repressor case, respectively, which are solved by Eqs 3 and 4. For the target molecule we obtain $\partial_t \bar{x} = \langle f_\pm(y) \rangle$, which is not solvable because f_\pm is nonlinear (i.e., the moments do not close). A deterministic analysis conventionally assumes $\langle f_\pm(y) \rangle \approx f_\pm(\bar{y})$, for which the equation becomes solvable by separation of variables. For example, in the case of H=1, using Eqs 1–4, we obtain

$$\bar{x}(t) = \begin{cases} \alpha t - (\alpha K/k) \log[(kt+K)/K] & (\text{activator}) \\ (\alpha/\mu) \log[(N+Ke^{\mu t})/(N+K)] & (\text{repressor}). \end{cases}$$
 (23)

Eq 23 is plotted in Fig 1C and 1D.

When including cell division, we compute the mean dynamics from the simulation trajectories (Fig 4C).

Details of the analytic approximations

To find the global minimum of Eq.8, we differentiate with respect to kt_* and K and set the results to zero, giving two equations. kt_* can be eliminated, leaving one equation for K,

$$\frac{1}{2}\log\frac{N}{K} = 1 - \frac{K}{N} \tag{24}$$

This equation is transcendental. However, in the limit $K \ll N$, we neglect the last term, which gives Eq 11.

To derive Eq 17, we use

$$\rho = 1 - \frac{t_0}{t_0} = 1 - \frac{\log N/K}{\mu t_0} \tag{25}$$



where the second step follows from $\bar{r}(t_0) = K$ according to Eq. 4; and, from Eq. 14,

$$\langle r \rangle = \frac{N}{\mu t_*} (1 - e^{-\mu t_*}) \approx \frac{N}{\mu t_*} \tag{26}$$

where the second step assumes that the repressor is fast-decaying, $\mu t_* \gg 1$. We use Eqs 26 and 25 to eliminate μt_* and K from Eq 8 in favor of ρ and $\langle r \rangle$,

$$\frac{\sigma_t^2 x_*}{t_*^2} \approx x_* \left(e^{N(1-\rho)/\langle r \rangle} - 1 \right) \frac{\langle r \rangle^2}{N^3} + \rho^2. \tag{27}$$

For nonlinear dynamics (ρ < 1) we may safely neglect the -1 in Eq 27. Then, differentiating Eq 27 with respect to N and setting the result to zero, we obtain $N = 3\langle r \rangle/(1-\rho)$. Inserting this result into Eq 27 produces Eq 17.

Analysis of the experimental data

To estimate the time at which migration terminates in Fig 3A, we refer to [5]. There, the position at which neuroblast migration terminates is measured with respect to seam cells V1 to V6 in the larva (see Fig. 4D in [5]). In particular, in wild type larvae, migration terminates between V2 and the midpoint of V2 and V1. This range corresponds to the magenta region in Fig 3A (see Fig. 4B, upper left panel, in [5]). Under the assumptions of constant migration speed and equal distance between seam cells, the horizontal axis in Fig 3A represents time.

To compute ρ for the experimental data in Fig 3A according to Eq 15 we use a trapezoidal sum. For the choices of x_* and t_* described in the text, this produces the ρ values in Fig 3B.

Acknowledgments

We thank Jeroen van Zon and Pieter Rein ten Wolde for helpful discussions, and van Zon for preliminary simulations that inspired the work.

Author Contributions

Conceptualization: Hendrik C. Korswagen, Andrew Mugler.

Data curation: Hendrik C. Korswagen.

Formal analysis: Shivam Gupta, Andrew Mugler.

Funding acquisition: Hendrik C. Korswagen, Andrew Mugler.

Investigation: Shivam Gupta, Julien Varennes, Andrew Mugler.

Methodology: Shivam Gupta, Julien Varennes, Hendrik C. Korswagen, Andrew Mugler.

Project administration: Hendrik C. Korswagen, Andrew Mugler.

Resources: Andrew Mugler. **Supervision:** Andrew Mugler.

Writing - original draft: Shivam Gupta, Julien Varennes, Andrew Mugler.

Writing – review & editing: Shivam Gupta, Julien Varennes, Hendrik C. Korswagen, Andrew Mugler.



References

- Bean JM, Siggia ED, Cross FR. Coherence and timing of cell cycle start examined at single-cell resolution. Molecular cell. 2006; 21(1):3–14. https://doi.org/10.1016/j.molcel.2005.10.035 PMID: 16387649
- Nachman I, Regev A, Ramanathan S. Dissecting timing variability in yeast meiosis. Cell. 2007; 131 (3):544–556. https://doi.org/10.1016/j.cell.2007.09.044 PMID: 17981121
- Schneider BL, Zhang J, Markwardt J, Tokiwa G, Volpe T, Honey S, et al. Growth rate and cell size modulate the synthesis of, and requirement for, G1-phase cyclins at start. Molecular and cellular biology. 2004; 24(24):10802–10813. https://doi.org/10.1128/MCB.24.24.10802-10813.2004 PMID: 15572683
- Carniol K, Eichenberger P, Losick R. A threshold mechanism governing activation of the developmental regulatory protein σF in Bacillus subtilis. Journal of Biological Chemistry. 2004; 279(15):14860–14870. https://doi.org/10.1074/jbc.M314274200 PMID: 14744853
- Mentink RA, Middelkoop TC, Rella L, Ji N, Tang CY, Betist MC, et al. Cell intrinsic modulation of Wnt signaling controls neuroblast migration in C. elegans. Developmental cell. 2014; 31(2):188–201. https://doi.org/10.1016/j.devcel.2014.08.008 PMID: 25373777
- Amir A, Kobiler O, Rokney A, Oppenheim AB, Stavans J. Noise in timing and precision of gene activities in a genetic cascade. Molecular Systems Biology. 2007; 3(1):71. https://doi.org/10.1038/msb4100113 PMID: 17299413
- 7. Meinhardt H. Models of biological pattern formation. Academic Press Inc; 1982.
- Tufcea DE, François P. Critical timing without a timer for embryonic development. Biophysical journal. 2015; 109(8):1724–1734. https://doi.org/10.1016/j.bpj.2015.08.024 PMID: 26488664
- Roux J, Hafner M, Bandara S, Sims JJ, Hudson H, Chai D, et al. Fractional killing arises from cell-to-cell variability in overcoming a caspase activity threshold. Molecular systems biology. 2015; 11(5):803. https://doi.org/10.15252/msb.20145584 PMID: 25953765
- 10. Van Kampen N. Stochastic Processes in Physics and Chemistry. vol. 1. Elsevier; 1992.
- McAdams HH, Arkin A. Stochastic mechanisms in gene expression. Proceedings of the National Academy of Sciences. 1997; 94(3):814–819. https://doi.org/10.1073/pnas.94.3.814
- Ghusinga KR, Dennehy JJ, Singh A. First-passage time approach to controlling noise in the timing of intracellular events. Proceedings of the National Academy of Sciences. 2017; 114(4):693–698. https:// doi.org/10.1073/pnas.1609012114
- Co AD, Lagomarsino MC, Caselle M, Osella M. Stochastic timing in gene expression for simple regulatory strategies. Nucleic acids research. 2017; 45(3):1069–1078. https://doi.org/10.1093/nar/gkw1235 PMID: 28180313
- **14.** Yurkovsky E, Nachman I. Event timing at the single-cell level. Briefings in functional genomics. 2012; p. els057.
- Jafarpour F, Vennettilli M, Iyer-Biswas S. Biological timekeeping in the presence of stochasticity. arXiv preprint arXiv:170310058. 2017;.
- Monti M, ten Wolde PR. The accuracy of telling time via oscillatory signals. Physical biology. 2016; 13(3):035005. https://doi.org/10.1088/1478-3975/13/3/035005 PMID: 27203353
- 17. Redner S. A guide to first-passage processes. Cambridge University Press; 2001.
- Chou T, D'orsogna M. First passage problems in biology. First-Passage Phenomena and Their Applications. 2014; 35:306. https://doi.org/10.1142/9789814590297_0013
- **19.** Iyer-Biswas S, Zilman A. First-Passage Processes in Cellular Biology. Advances in Chemical Physics, Volume 160. 2016; p. 261–306.
- Walczak AM, Mugler A, Wiggins CH. Analytic methods for modeling stochastic regulatory networks. Computational Modeling of Signaling Networks. 2012; p. 273–322. https://doi.org/10.1007/978-1-61779-833-7_13
- 21. Ji N, Middelkoop TC, Mentink RA, Betist MC, Tonegawa S, Mooijman D, et al. Feedback control of gene expression variability in the Caenorhabditis elegans Wnt pathway. Cell. 2013; 155(4):869–880. https://doi.org/10.1016/j.cell.2013.09.060 PMID: 24209624
- Raj A, van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. Cell. 2008; 135(2):216–226. https://doi.org/10.1016/j.cell.2008.09.050 PMID: 18957198
- 23. Friedman N, Cai L, Xie XS. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. Physical review letters. 2006; 97(16):168302. https://doi.org/10.1103/ PhysRevLett.97.168302 PMID: 17155441
- Shahrezaei V, Swain PS. Analytical distributions for stochastic gene expression. Proceedings of the National Academy of Sciences. 2008; 105(45):17256–17261. https://doi.org/10.1073/pnas. 0803850105



- Gillespie DT. Exact stochastic simulation of coupled chemical reactions. The journal of physical chemistry. 1977; 81(25):2340–2361. https://doi.org/10.1021/j100540a008
- Bel G, Munsky B, Nemenman I. The simplicity of completion time distributions for common complex biochemical processes. Physical biology. 2009; 7(1):016003. https://doi.org/10.1088/1478-3975/7/1/016003 PMID: 20026876
- McAdams HH, Shapiro L. Circuit simulation of genetic networks. Science. 1995; 269(5224):650–656. https://doi.org/10.1126/science.7624793 PMID: 7624793
- Mahaffy J, Pao C. Models of genetic control by repression with time delays and spatial effects. Journal of mathematical biology. 1984; 20(1):39–57. https://doi.org/10.1007/BF00275860 PMID: 6491544
- Josić K, López JM, Ott W, Shiau L, Bennett MR. Stochastic delay accelerates signaling in gene networks. PLoS computational biology. 2011; 7(11):e1002264. https://doi.org/10.1371/journal.pcbi. 1002264 PMID: 22102802