Is the Cell Nucleus a Necessary Component in Precise Temporal Patterning?

Jaroslav Albert[1,2]*, Marianne Rooman[1]

1 BioModeling, BioInformatics & BioProcesses, Université Libre de Bruxelles, Brussels, Belgium, 2 Applied Physics Research Group, Vrije Universiteit Brussel, Brussels, Belgium

* jalbert@ulb.ac.be

Abstract

One of the functions of the cell nucleus is to help regulate gene expression by controlling molecular traffic across the nuclear envelope. Here we investigate, via stochastic simulation, what effects, if any, does segregation of a system into the nuclear and cytoplasmic compartments have on the stochastic properties of a motif with a negative feedback. One of the effects of the nuclear barrier is to delay the nuclear protein concentration, allowing it to behave in a switch-like manner. We found that this delay, defined as the time for the nuclear protein concentration to reach a certain threshold, has an extremely narrow distribution. To show this, we considered two models. In the first one, the proteins could diffuse freely from cytoplasm to nucleus (simple model); and in the second one, the proteins required assistance from a special class of proteins called importins. For each model, we generated fifty parameter sets, chosen such that the temporal profiles they effectuated were very similar, and whose average threshold time was approximately 150 minutes. The standard deviation of the threshold times computed over one hundred realizations were found to be between 1.8 and 7.16 minutes across both models. To see whether a genetic motif in a prokaryotic cell can achieve this degree of precision, we also simulated five variations on the coherent feed-forward motif (CFFM), three of which contained a negative feedback. We found that the performance of these motifs was nowhere near as impressive as the one found in the eukaryotic cell; the best standard deviation was 6.6 minutes. We argue that the significance of these results, the fact and necessity of spatio-temporal precision in the developmental stages of eukaryotes, and the absence of such a precision in prokaryotes, all suggest that the nucleus has evolved, in part, under the selective pressure to achieve highly predictable phenotypes.

Introduction

Biological cells house an enormous number of diverse molecules that interact with the cells' environment and among each other. Such a complex network of interactions makes the evolution of any molecule stochastic, be it the change in its position, conformation, or copy number. The effects stochasticity has on the phenotype of a cell are many, ranging from significant to negligible; in some cases they are detrimental [1], in others they play an important role [2]. For

example, a population of *Escherichia coli* may survive the attack from an antibiotic drug thanks to random switching between two phenotypes—one sensitive, the other resistant to the drug [3]. Sporulation in yeast is another example; in this case the concentration level of a single protein determines the fate (death or survival) of a cell [4]. In eukaryotes, too, one can observe noise-induced subpopulations at work, such as was done for the immune T-cells [5]. Even in the early development of eukaryotes, where spatio-temporal precision is essential, stochastic variations can be of service, as was exemplified by the segmentation clock [6]: the oscillating expression of the gene *her1* is sustained by virtue of copy number fluctuations.

As essential as these mechanisms may appear, observation, experimentation and modeling of gene expression suggest that important genes must be (and are) tightly regulated [7–9]. And in fact, numerous mechanisms of noise reduction and noise filtering have been discovered. Negative transcriptional [10, 11] and translational [8, 12] autoregulation, non-linear protein degradation [13], multimer formation [14–16] are just a few examples worth mentioning here. Especially in the developmental stages of eukaryotes is noise reduction desirable: the spatio-temporal execution of the genetic program must be as noise free as possible [17, 18].

Previously, studies on the regulation of noise in development have focused mainly on stabilization of newly differentiated cell states [17], noise filtering of morphogen gradients [18], how cellular boundaries are formed and maintained [19], and how the segmentation clock, e. g. in Zebrafish, can lead to synchronized oscillations due to cell signaling [20, 21]. In this paper, we investigate how the segregation of the nuclear space from that of the cytoplasm affects the timing accuracy of gene expression. The effects of such division have been investigated in previous studies, e. g. in the context of Notch signaling and synchronized oscillations [20], the Hes1 and p53-Mdm2 pathways [22], and how the intrinsic noise, coupled with the effects of the nuclear membrane, can explain heterogeneity in embryonic stem cell differentiation [23]. Motifs with autoregulation have also been studied with emphasis on the spatial degree of freedom [24]. While all these studies do incorporate the presence of the nuclear membrane, they do not focus specifically on the question of whether such separation can lead to enhanced timing accuracy in gene transcription. It is this niche that we try to fill in this paper.

What we consider are two models of protein transport from cytoplasm to nucleus. In the first model, we allow the protein to diffuse freely across the nuclear envelope. The second model involves a three-gene motif comprising of the gene of interest and two other genes that code for importin-α and importin-β—two out of a class of proteins responsible for facilitating nuclear import of other proteins. Both models include a negative feedback. The most simplistic version of this motif, one which does not involve the nuclear envelope and hence importins, has been extensively studied and is well understood: the time for a gene to reach a steady state is fast and the noise diminished compared to the situation with no negative feedback. In order to prolong the time to reach a steady state, a positive feedback has been the traditional device; however, this leads to an increase of noise. We demonstrate that the inclusion of the cell nucleus can also lead to a prolonged transition into a steady state, however, at the same time, to a reduction of noise—the very opposite of the nucleus-free situation.

Furthermore, we explore the possibility that without the nucleus it is not possible to generate a delayed transition while also reducing the noise. We do this by comparing the coherent feed-forward motif (CFFM), and variations thereof, to the motif involving a nucleus. The CFFM is highly abundant in many prokaryotes [25], which makes it a good first step. Our results show that although it is possible to match the temporal profile of the protein concentration inside the nucleus with that of the protein coded for by the third and last gene downstream of the CFFM, the noise in the latter will tend to be much greater during the transition into a steady state. In light of these results, we dare ask whether the cell nucleus is an indispensable component of the near deterministic workings of the developmental gene regulatory networks in eukaryotes.

Results

Modeling mRNA and protein concentration in eukaryotes

For an mRNA molecule to be translated, it must find its way from the nucleus into the cytoplasm. The mRNA recruits special proteins with which it forms a complex [26] that freely diffuses into the cytoplasm [27]. Once in the cytoplasm, the mRNA rarely returns back to the nucleus, and so, from now on we will consider its transport a one way trip. Also, it is only in the cytoplasm that mRNA is actively degraded; in the nucleus, only damaged mRNA is degraded. The mechanism of intra-compartmental locomotion of proteins depends (in part) on their size: small proteins ($< \sim$ 40kDa) tend to diffuse freely across the membrane, while large ones need assistance from a family of proteins known as importins and exportins [28]. Those proteins that are needed back in the nucleus after translation, e.g. transcription factors, contain an amino acid sequence that the importins recognize and bind. The simplest scenario involves only one type of importin, importin-β. In another, more complex, situation two importins, importin-α and importin-β, work together to facilitate transport of cargo proteins across the nuclear membrane. In the most likely scenario, importin-α binds to importin-β before it can bind to the cargo protein. Only in this three-protein complex is import into the nucleus most efficient—though the importin-β–cargo complex, too, can sneak in but with a significantly smaller rate. Fig 1 shows in detail the essential reactions involved in synthesizing and transporting mRNA and protein. Models of this type of system have been studied previously, both theoretically and experimentally [28–30]

In the next two sections, we set up the deterministic equations for both the free diffusion (simple) and the importin-assisted (extended) models, and in the section after that we analyze the stochastic dynamics of these two models.

Deterministic model—simple. We model the eukaryotic transcriptional and translational machinery in a two-compartment cell: a negatively auto-regulated gene is transcribed in the nucleus; the generated mRNA diffuses through the membrane from the nucleus to the cytoplasm, where it is translated into protein; the protein then diffuses freely into the nucleus. Let us write down the set of equations governing the dynamics of mRNA and protein concentrations. We denote the four essential processes, the rate of transcription, translation, and mRNA and protein degradation, as r, K, k and q respectively. Then, the concentrations of mRNA, X_n, and protein, Y_n, inside the nucleus, and of the mRNA, X_c, and protein, Y_c, in the cytoplasm obey:

$$
\begin{aligned}
\dot{X}_n &= r(1 - P_1 - P_2) - \lambda X_n, \\
\dot{X}_c &= \lambda X_n - k X_c, \\
\dot{Y}_n &= \kappa Y_c - q Y_n, \\
\dot{Y}_c &= K X_c - (q + \kappa) Y_c,
\end{aligned}
\tag{1}
$$

where λ is the rate of mRNA diffusion into the cytoplasm, while κ is the rate of protein diffusion from the cytoplasm into the nucleus. The promoter states P_1 and P_2 are governed by:

$$
\begin{aligned}
\dot{P}_1 &= w_1 Y_n (1 - P_1 - P_2) + \tilde{w}_{-1} P_2 - w_{-1} P_1 - w_1 Y_n P_1 \\
\dot{P}_2 &= w_1 Y_n P_1 - \tilde{w}_{-1} P_2,
\end{aligned}
\tag{2}
$$

where w_1 and w_{-1} are the rates at which Y_n binds and dissociates a free promoter; $\tilde{w}_{-1}$ is the rate at which Y_n decouples from a fully occupied promoter P_2. In Fig 2, we show the dynamics of the variables Y_n and Y_c for typical parameters.
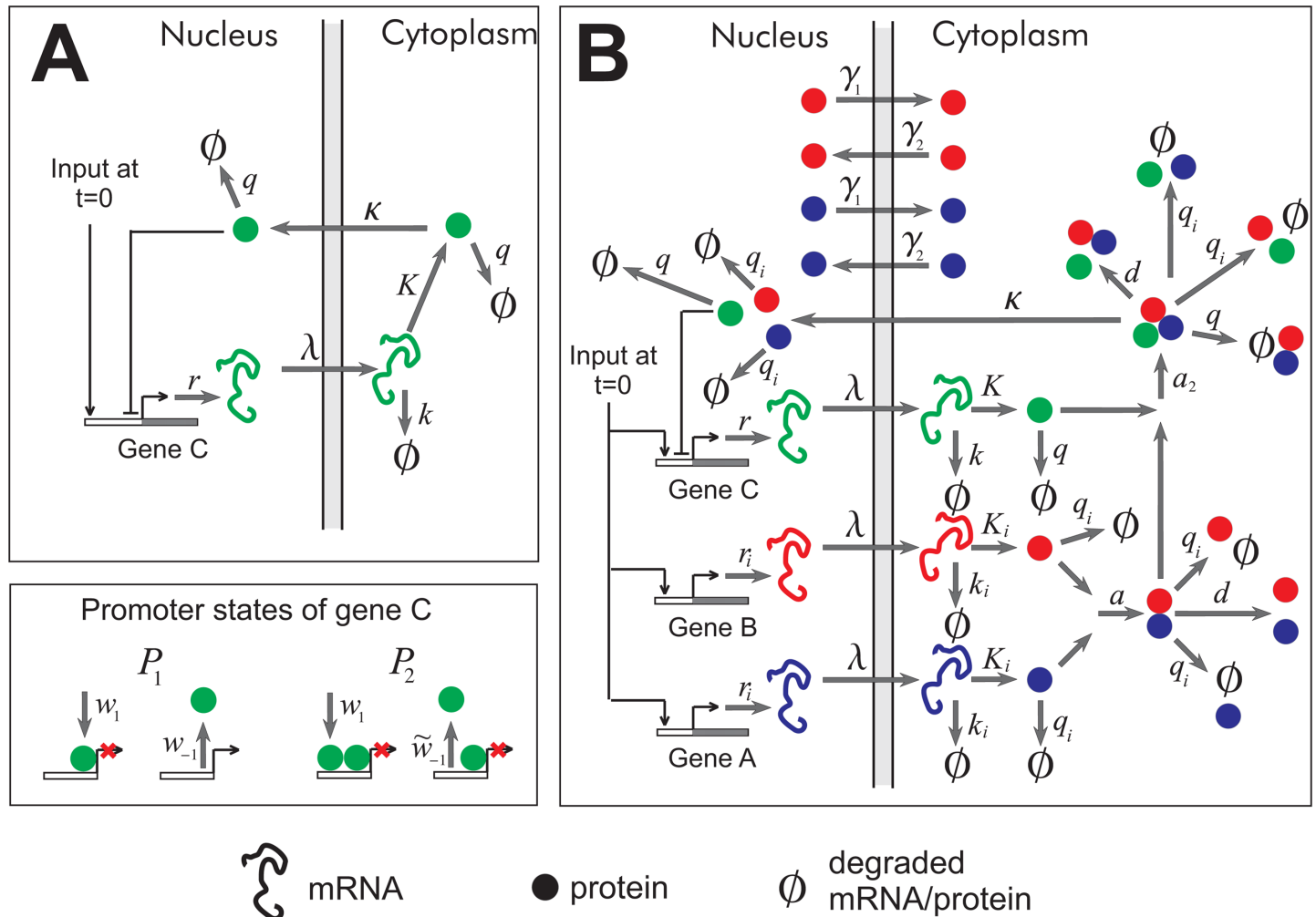
Fig 1. Two models of nuclear transport. A) Simple model: the protein of gene C diffuses freely from cytoplasm into the nucleus. B) Extended model: a gene network involving importin-α (blue), importin-β (red) and cargo protein (green). The symbol $\emptyset$ indicates degraded protein and mRNA. The reactions are labeled as in Eqs (3, 2, 5, 6): r—transcription rate; λ—export rate of mRNA into the cytoplasm; K—translation rate; k—rate of mRNA degradation; q—rate of protein degradation; a (d)—rate of association (dissociation) between importins α and β; a_2 (d)—rate of association (dissociation) between the importin-α–β complex and cargo protein; κ—import rate of the importin-α–β-cargo complex; γ_1—rate export of importins α and β into cytoplasm; w_1 (w_{-1})—rate of association (dissociation) between a cargo protein and a free promoter of gene C; $\tilde{w}_{-1}$—rate of dissociation of a cargo protein from a fully occupied promoter.

doi:10.1371/journal.pone.0134239.g001

Deterministic model—extended. Let us now consider the role of importins in the transport of cargo proteins. Before proceeding, we make the following assumptions:

- mRNA does not degrade inside the nucleus;

- mRNA does not diffuse back into the nucleus;

- The rate of diffusion into the cytoplasm is the same for all species of mRNA λ;

- There is no export of cargo proteins from the nucleus to the cytoplasm;

- Importins α and β diffuse freely in and out of the nucleus;

- The rates of transcription (r_i), translation (K_i), mRNA degradation (k_i), and protein degradation (q_i) associated with importin-α, are identical to those associated with importin-β;
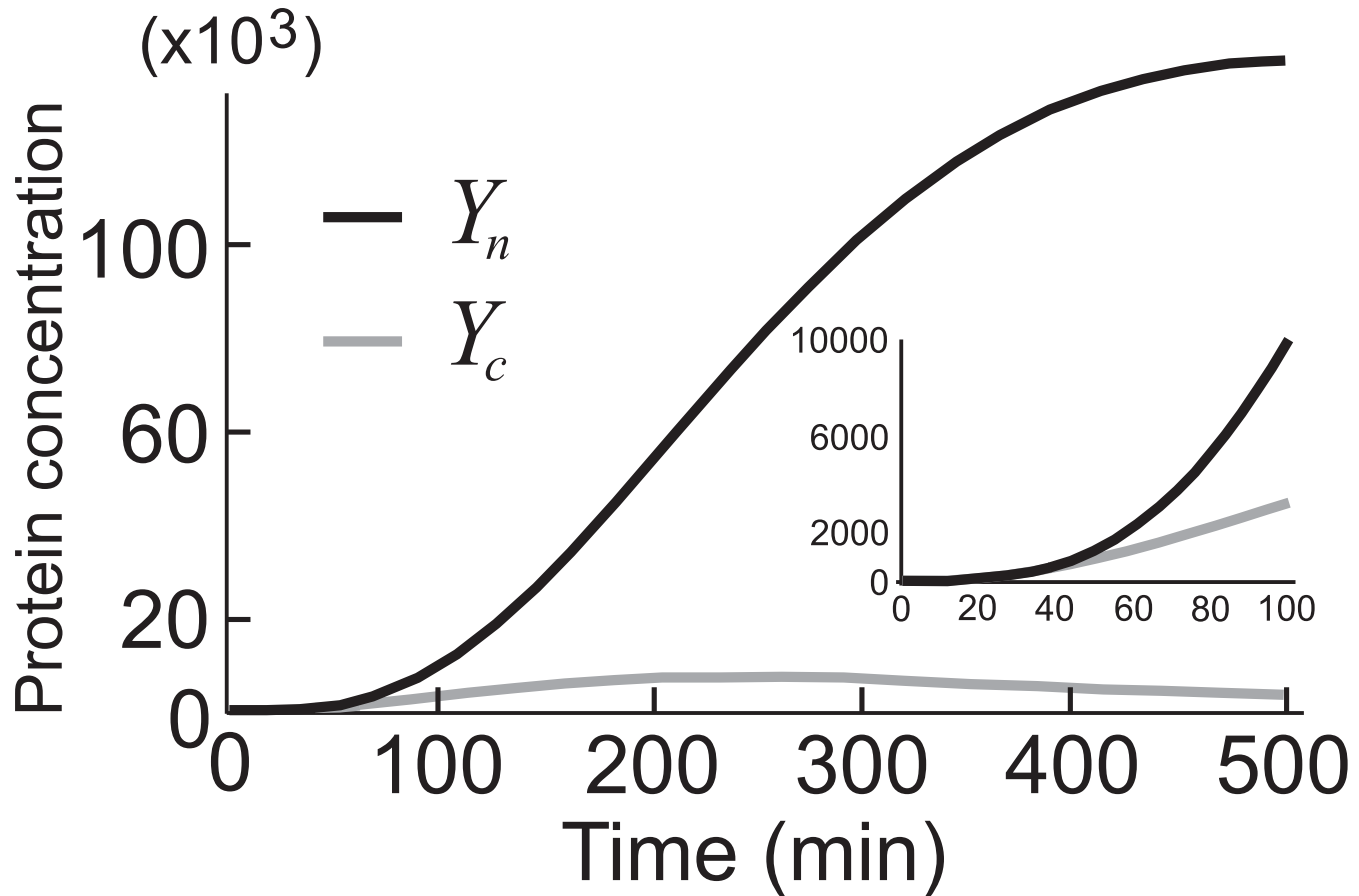
Fig 2. Evolution of average protein concentrations. A close-up view of the figure is given on the right-hand side. The nuclear and the cytoplasmic concentrations rise at approximately the same rate until $t \approx 40$, after which time the nuclear concentration begins to increase much more rapidly.

- The rate of dissociation b of importin-α from importin-β is equivalent to that at which the importin-α-importin-β complex dissociates from the cargo protein;

- Proteins in the nucleus degrade at the same rate as those in the cytoplasm;

- Degradation rates for all species of protein remain constant under multimerization;

- The promoter occupancy of gene C, i.e. P_1 and P_2 in Fig 1B, corresponding to one and two proteins bound to it respectively, occurs with the same rate (w_1).

- The concentrations of all gene products are zero at t = 0.

All these assumptions are true to various degrees for various gene products in the real world. Our main reason for making them is to reduce the number of parameters, which, we argue, does not diminish the significance of our results: whatever this system can achieve with fewer parameters, it will necessarily achieve with more parameters. The last assumption may be an oversimplification, as the regulation of both importin-α and importin-β is more complicated by virtue of being part of a larger gene network [31–33]; same applies to the cargo protein. However, for our purposes here, which are to investigate the role of the cell nucleus in concentration dynamics and in the regulation of intrinsic noise, it will suffice as the first step. Discussion of the situation where the importins start out with a non-zero concentration, e. g. in a steady state, will follow in the "Discussion" section.

Let us write down the set of equations governing the dynamics of mRNA and protein concentrations under the above mentioned assumptions. Using the same notation for the cytoplasmic and nuclear mRNA and cargo protein as in the previous section, the equations of motion read:

$$
\begin{aligned}
\dot{X}_n &= r(1 - P_1 - P_2) - \lambda X_n, \\
\dot{X}_c &= \lambda X_n - k X_c, \\
\dot{Y}_n &= \kappa Y_c - q Y_n, \\
\dot{Y}_c &= K X_c + (d + 2q_i)Z_3 - a_2 Z_2 Y_c - q Y_c,
\end{aligned}
\tag{3}
$$

where Z_2 is the importin-α–importin-β heterodimer, and Z_3 is the importin-α–importin-β–cargo complex. The coefficient a_2 is the binding rate of Y_c to Z_2, and d the rate at which they dissociate. As before, κ is the transport rate into the nucleus, but this time it is Z_3 that is being transported. Lastly, q_i stands for the degradation rate of both types of importins.

$$
\begin{aligned}
\dot{P}_1 &= w_1 Y_n (1 - P_1 - P_2) + \tilde{w}_{-1} P_2 - w_{-1} P_1 - w_1 Y_n P_1 \\
\dot{P}_2 &= w_1 Y_n P_1 - \tilde{w}_{-1} P_2,
\end{aligned}
\tag{4}
$$

The equations for the α- and β-importins in the cytoplasm and nucleus $Y_{\alpha n}$, $Y_{\alpha c}$, $Y_{\beta n}$, $Y_{\beta c}$ and the corresponding mRNAs, $X_{\alpha n}$, $X_{\alpha c}$, $X_{\beta n}$, $X_{\beta c}$ from which they are translated, read:

$$
\begin{aligned}
\dot{X}_{\alpha n} &= r_i - \lambda X_{\alpha n}, \\
\dot{X}_{\alpha c} &= \lambda X_{\alpha n} - k_i X_{\alpha c}, \\
\dot{X}_{\beta n} &= r_i - \lambda X_{\beta n}, \\
\dot{X}_{\beta c} &= \lambda X_{\beta n} - k_i X_{\beta c}, \\
\dot{Y}_{\alpha n} &= \gamma_2 Y_{\alpha c} - \gamma_1 Y_{\alpha n} + \kappa Z_3 - q_i Y_{\alpha n} \\
\dot{Y}_{\alpha c} &= -\gamma_2 Y_{\alpha c} + \gamma_1 Y_{\alpha n} + K_i X_{\alpha c} + (d + q_i)Z_2 + q_i Z_2 - a Y_{\alpha c} Y_{\beta c} - q_i Y_{\alpha c} \\
\dot{Y}_{\beta n} &= \gamma_2 Y_{\beta c} - \gamma_1 Y_{\beta n} + \kappa Z_3 - q_i Y_{\beta n} \\
\dot{Y}_{\beta c} &= -\gamma_2 Y_{\beta c} + \gamma_1 Y_{\beta n} + K_i X_{\beta c} + (d + q_i)Z_2 + q_i Z_2 - a Y_{\alpha c} Y_{\beta c} - q_i Y_{\beta c}.
\end{aligned}
\tag{5}
$$

The coefficients $\gamma_1 = \gamma$ and $\gamma_2 = \nu\gamma$ are the free diffusion rates for importins, with $\nu = V_n/V_c$ being the ratio of the volume of the nucleus to the volume of the cytoplasm. For the protein complexes, we have

$$
\begin{aligned}
\dot{Z}_2 &= a Y_{\beta c} Y_{\alpha c} - (d + 2q_i)Z_2, \\
\dot{Z}_3 &= a_2 Z_2 Y_c - (d + 2q_i + q + \kappa)Z_3.
\end{aligned}
\tag{6}
$$

The importin-α- importin-β heterodimer binds to (dissociates from) Y_c at the rate of a_2 (d).

Let us now select values for all the parameters and solve numerically Eqs (3), (5) and (6). Fig 2 shows the temporal profiles of all proteins and their complexes in a) cytoplasm and b) the nucleus. All three genes are turned on at $t = 0$; this may be the result of a signaling molecule or the unraveling of the chromatin(s) harboring these genes. The most relevant feature for our purposes is the delayed production of Y_n. For the first forty minutes or so its concentration remains low, and then begins rapidly increasing. This switch-like behavior is common in many biological systems. In what follows, we look at the stochastic behavior of this system and how the delays in the production of Y_n change from cell to cell; stated otherwise, we analyze the noise.

Stochastic model. Our goal in this section is not only to look at noise in the system defined by the parameters in Figs 2 and 3, but also to study how much the noise changes when we
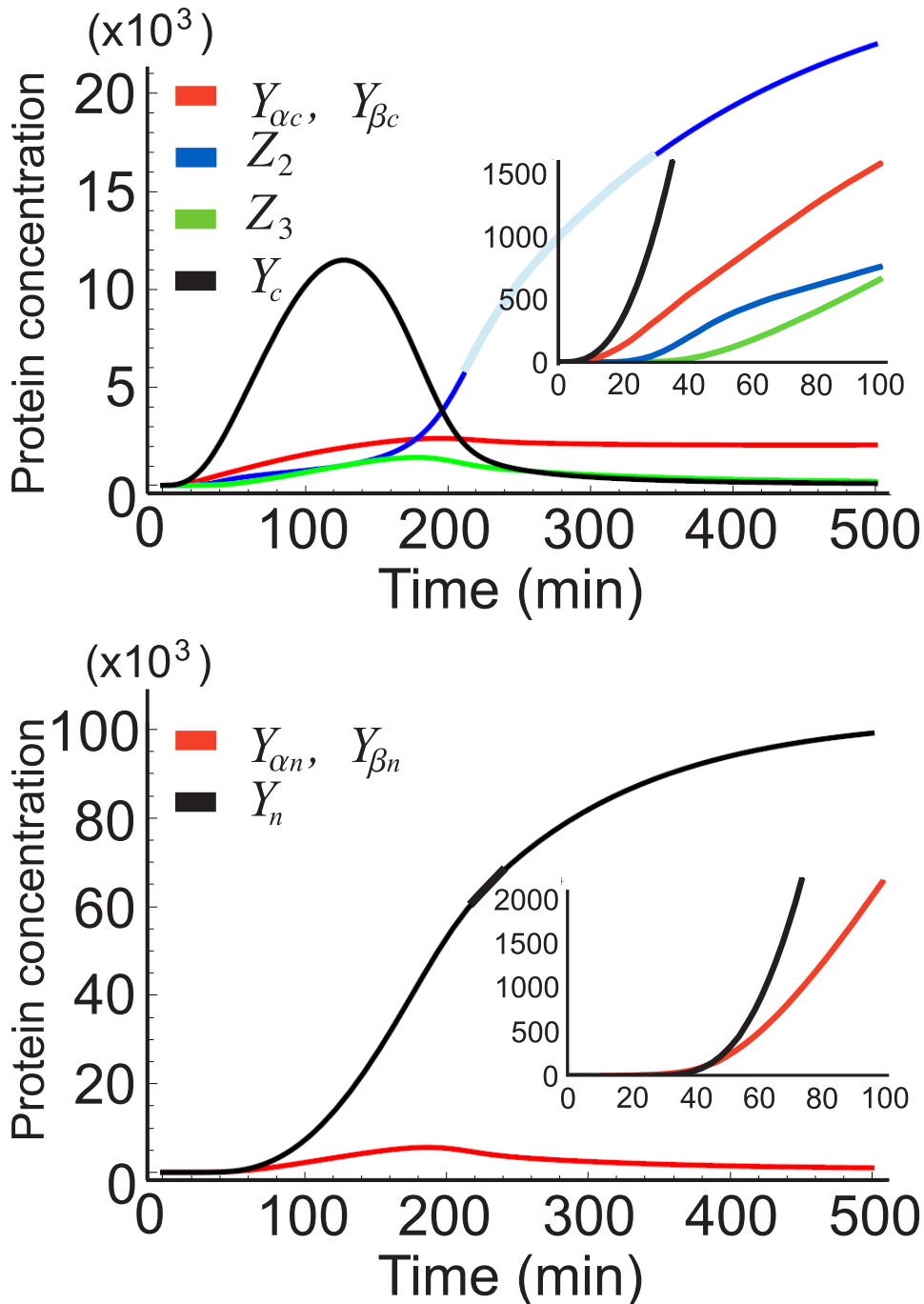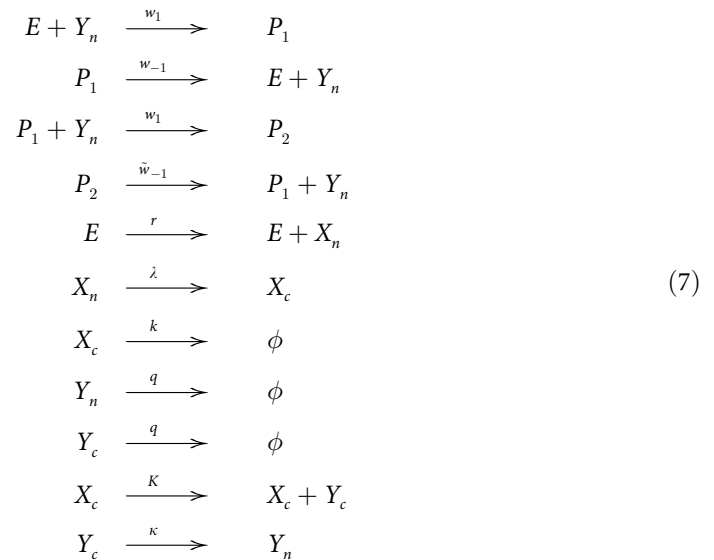
Fig 3. Evolution of average protein concentrations. A close-up view of the figure is given on the right-hand side A) Protein concentrations in the cytoplasm. At t = 0 all three genes are turned on. The cargo protein in the cytoplasm (black) begins to accumulate first, followed by the importin α and β (red). At about t = 20 min, importins α and β start forming heterodimers (blue), and 20 min after that, the concentration of importin-α–β-cargo complex (green) begins to rise. B) Protein concentrations in the nucleus. Both the cargo protein (black) and the importins α and β (red) start to accumulate at about t = 40 min.
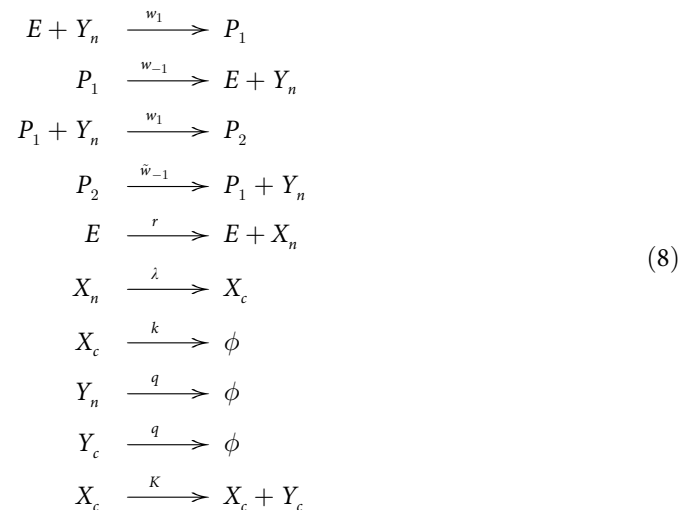
change these parameters. We will only consider those sets of parameters that give rise to a phenotype similar to the profile of Y_n in Fig 3. We will call this profile the reference curve and denote it by Y_r. There is nothing particularly special about this reference cure other than being more or less a typical concentration profile.

Since we are interested mostly in the transitional part of Y_r, we impose that at $t = 150$ min the concentrations of all generated profiles be within 5% of $Y_r(t = 150$ min$)$, and that their slopes fall within 5% of $\dot{Y}_r$ $(t = 150$ min$)$. Since we consider the steady state concentrations less important, we give them a latitude of 30%. Although we are less concerned with the profiles of all the other species, we do not want them behaving in a manner that is improbable, such as reaching unusually large steady state concentrations. To that end, we ensure that the sum of steady state concentrations of all species (proteins and their complexes) corresponding to a particular parameter set is no greater than 30% of the sum of steady state concentrations arising from the reference parameter set responsible for Y_r. This last constraint is related to the amount of energy the cell needs to expend in order to maintain the steady state concentrations. We will discuss this point further in the "Discussion" section. The fifty profiles we obtain for the two models are shown in Fig 4A and 4B. The range of values for each parameter, along the references from which they were taken, are listed in Table 1.
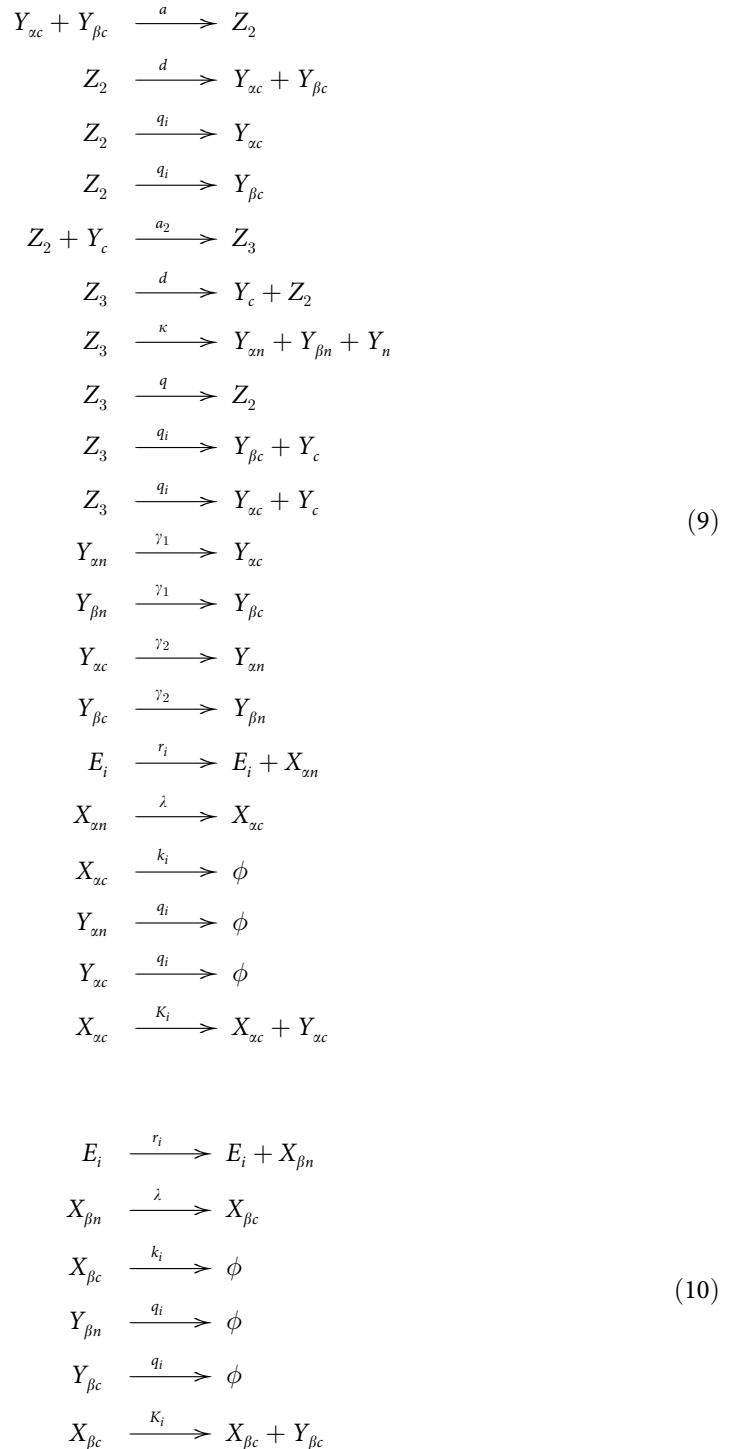
The next step is to simulate the reactions of Fig 1 using the Gillespie algorithm for each parameter set. In the standard notation, for the simple models, these reactions are:

$$
\begin{aligned}
E + Y_n &\xrightarrow{w_1} P_1 \\
P_1 &\xrightarrow{w_{-1}} E + Y_n \\
P_1 + Y_n &\xrightarrow{w_1} P_2 \\
P_2 &\xrightarrow{\tilde{w}_{-1}} P_1 + Y_n \\
E &\xrightarrow{r} E + X_n \\
X_n &\xrightarrow{\lambda} X_c \\
X_c &\xrightarrow{k} \phi \\
Y_n &\xrightarrow{q} \phi \\
Y_c &\xrightarrow{q} \phi \\
X_c &\xrightarrow{K} X_c + Y_c \\
Y_c &\xrightarrow{\kappa} Y_n
\end{aligned}
\tag{7}
$$

and for the extended model they read:

$$
\begin{aligned}
E + Y_n &\xrightarrow{w_1} P_1 \\
P_1 &\xrightarrow{w_{-1}} E + Y_n \\
P_1 + Y_n &\xrightarrow{w_1} P_2 \\
P_2 &\xrightarrow{\tilde{w}_{-1}} P_1 + Y_n \\
E &\xrightarrow{r} E + X_n \\
X_n &\xrightarrow{\lambda} X_c \\
X_c &\xrightarrow{k} \phi \\
Y_n &\xrightarrow{q} \phi \\
Y_c &\xrightarrow{q} \phi \\
X_c &\xrightarrow{K} X_c + Y_c
\end{aligned}
\tag{8}
$$

$$Y_{\alpha c} + Y_{\beta c} \xrightarrow{a} Z_2$$

$$Z_2 \xrightarrow{d} Y_{\alpha c} + Y_{\beta c}$$

$$Z_2 \xrightarrow{q_i} Y_{\alpha c}$$

$$Z_2 \xrightarrow{q_i} Y_{\beta c}$$

$$Z_2 + Y_c \xrightarrow{a_2} Z_3$$

$$Z_3 \xrightarrow{d} Y_c + Z_2$$

$$Z_3 \xrightarrow{\kappa} Y_{\alpha n} + Y_{\beta n} + Y_n$$

$$Z_3 \xrightarrow{q} Z_2$$

$$Z_3 \xrightarrow{q_i} Y_{\beta c} + Y_c$$

$$Z_3 \xrightarrow{q_i} Y_{\alpha c} + Y_c$$

$$Y_{\alpha n} \xrightarrow{\gamma_1} Y_{\alpha c}$$

$$Y_{\beta n} \xrightarrow{\gamma_1} Y_{\beta c}$$

$$Y_{\alpha c} \xrightarrow{\gamma_2} Y_{\alpha n}$$

$$Y_{\beta c} \xrightarrow{\gamma_2} Y_{\beta n}$$

$$E_i \xrightarrow{r_i} E_i + X_{\alpha n}$$

$$X_{\alpha n} \xrightarrow{\lambda} X_{\alpha c}$$

$$X_{\alpha c} \xrightarrow{k_i} \phi$$

$$Y_{\alpha n} \xrightarrow{q_i} \phi$$

$$Y_{\alpha c} \xrightarrow{q_i} \phi$$

$$X_{\alpha c} \xrightarrow{K_i} X_{\alpha c} + Y_{\alpha c}$$

$$(9)$$

$$E_i \xrightarrow{r_i} E_i + X_{\beta n}$$

$$X_{\beta n} \xrightarrow{\lambda} X_{\beta c}$$

$$X_{\beta c} \xrightarrow{k_i} \phi$$

$$Y_{\beta n} \xrightarrow{q_i} \phi$$

$$Y_{\beta c} \xrightarrow{q_i} \phi$$

$$X_{\beta c} \xrightarrow{K_i} X_{\beta c} + Y_{\beta c}$$

$$(10)$$

where E and E_i stand for the unoccupied promoter of the genes that code for X_n, $X_{\alpha n}$ and $X_{\beta n}$. From these reactions, one can obtain Eqs (1), (2), (3), (5) and (6) by using the standard assumption that the average of any product can be approximated by the product of the individual averages, e. g. $\langle Y_{\alpha c} Y_{\beta c} \rangle \approx \langle Y_{\alpha c} \rangle \langle Y_{\beta c} \rangle$.

What is most interesting to us is the distribution of times at which the protein concentrations of Y_n cross a certain threshold. The concentration threshold of 26.5×10^3 is reached after
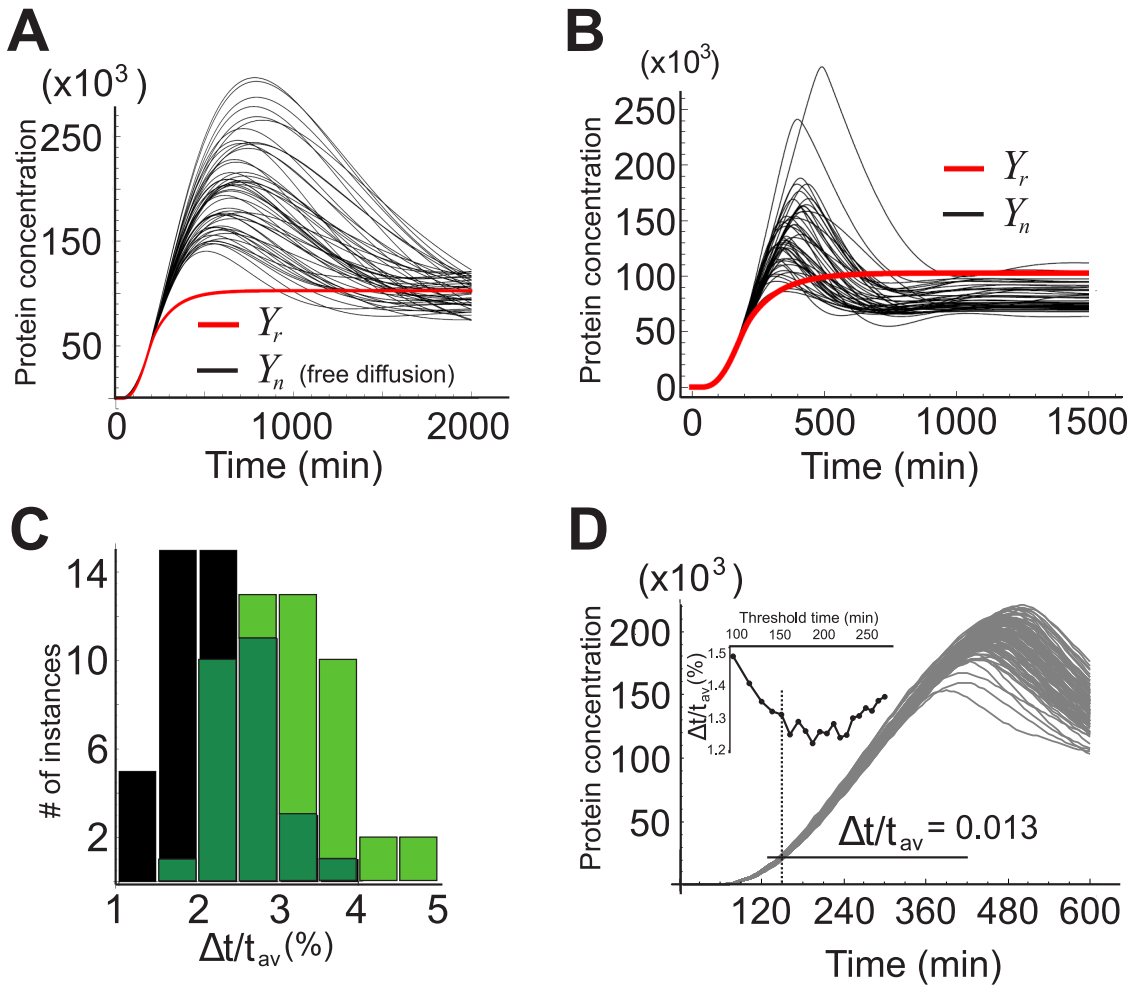
Fig 4. Results for the simple and the extended models. Profiles for the nuclear concentration corresponding to the 50 parameter sets for A) the simple model, and B) the extended model. In red is the reference curve Y_r. C) Distribution of threshold times for the simple model (green) and the extended model (black). D) Superposition of 100 realizations for the best case offered by the extended model: $\Delta t/t_{av} = 0.013$. The smaller plot within shows values of $\Delta t/t_{av}$ for different threshold times t_{av}.

doi:10.1371/journal.pone.0134239.g004

the average time $t_{av} = 150$ min, as can be worked out from Eqs (3), (5) and (6). As with the reference curve Y_r t being 150 min is not particularly special, but such a delay does conform to what is observed in the developmental stages of eukaryotes [34, 35]. The results for all 50 sets, for both the simple and the extended models, is plotted in Fig 4C as a histogram. Appearing on the horizontal axis is the standard deviation of the threshold times Δt divided by the average time $t_{av} \approx 150$ min. The vertical axis shows the number of cases, out of 50, corresponding to a specific range of $\Delta t/t_{av}$. The superposition of histograms corresponding to the two models reveals that the extended model does better than the simple model. This may seem counterintuitive, given that the presence of importins should introduce an additional source of intrinsic noise which is not present in the simple model. Nevertheless, the degree of noise filtering achieved by both models is equally impressive. Fig 4D shows the best case scenario, belonging to the extended model, for various threshold times. There is another interesting feature to be observed here: 1) the noise during the transition, more specifically between 80 and about 300 min, is very small (see Fig 4C). 2) the noise at equilibrium is quite large, giving a Fano factor

Table 1. Range of values for all parameters.

Parameter	Range	Reference
r, r_i	1–10/min	[34, 35]
K, K_i	maximum: 11/min/mRNA	[34, 35]
k	0.5–0.005/min	[34, 35]
q	0.5–0.005/min	[34, 35]
λ	maximum 0.0534/min	[27, 36]
q_i	0.5775–6.79×10^{-3}/min	[37]
γ	0.01–0.2/min	[30]
a	0.01975–1.225×10^{-4}/min	[38]
a_2	$0.306 a_2$	[38]
d	0.015–0.102/min	[38]
κ	0.0248–0.426/min	[29]
V_n(Nuclear volume)	$4\mu m^3$	[29]
V_C(Cytoplasmic volume)	$40\mu m^3$	[29]
Number of nuclear pore complexes	80–170	[29]

Parameters (in symbols), their range, and the reference from which they were taken.

doi:10.1371/journal.pone.0134239.t001

for the concentration Y_n of 1540. This may seem surprising: negative autoregulation is known to decrease noise. However, a limit on this effect has been shown to exist when delays are included in stochastic models [39, 40]. Although we did not add delays in our model explicitly, they are implicit in the fact that between transcription and promoter association there are several reactions that must take place. Looking at the histogram in Fig 4C, it is tempting to conclude that accuracy in timing is built into this type of system. We defer a more extensive discussion of this point for the "discussion" section.

Delay generating circuits in prokaryotes

Coherent feed-forward motif. We would like to turn attention now to gene networks which do not involve a physical barrier between different reactions, such as those found in prokaryotes, and ask the question: can these types of network mimic the temporal profile of Y_r of Fig 3, and if so, can they do it while maintaining such low level of noise? This question is too hypothetical to be graced with a definite answer; the number of possible motifs is too large to explore. However, we can start by looking at some delay generating motifs common to prokaryotes and see if they could be modified, e.g. by tuning their parameters, to minimize noise while keeping a temporal profile similar to Y_r. Ideally, this (typically three-gene) motif should resemble the one of Fig 1B as closely as possible, i.e. the concentration of the protein for which gene C codes can grow only when genes A and B are on and their gene products abundant. The most common motif that fits the bill is the coherent feed-forward motif (CFFM), occurring at a frequency of 70% in *E. coli*[25]. It can come in two flavors, as depicted in Fig 5. In both cases, initiation of gene C can begin only when the transcription factors of both gene A and gene B are bound to the promoter of this gene; however, they may bind separately, one after the other (Fig 5B), or as a heterodimer (Fig 5C). In what follows, we explore both cases. For now we do not include a negative feedback; this feature will be added in the next section.

We are interested in making a comparison between the system shown in Fig 1B and the two versions of the CFFM of Fig 5. In particular, the temporal profiles of Y_r and the protein belonging to gene C should be close. To make this comparison as fair as possible, we apply the same
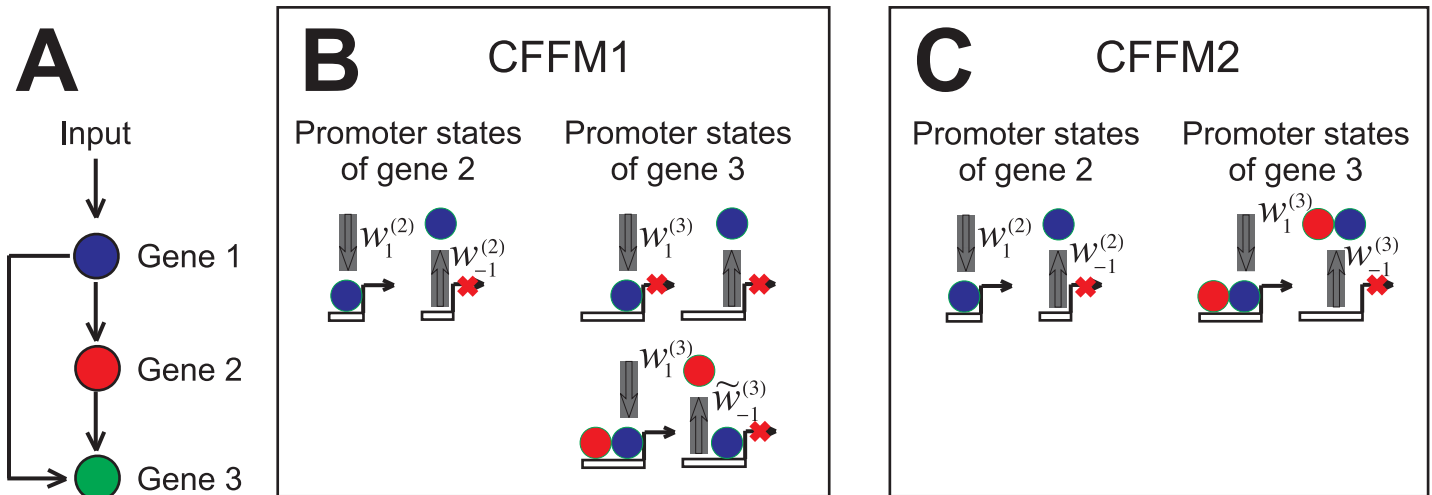
Fig 5. Two types of the coherent feed-forward motif. A) A schematic of the CFFM. The transcription factor of gene A (TFA) enables transcription of gene B, and TFA and TFB together enable transcription of gene C. B) First, TFA binds to the promoter of C; then, either it dissociates or TFB binds to the promoter next to TFA. TFB can bind to the promoter only when TFA is already there. Only when TFA and TFB are both in place can transcription of gene C be initiated. C) TFA and TFB form a heterodimer which then binds the promoter of gene 3, initiating its transcription.

conditions and constraints as in the previous section. The reader should refer to this section and merely replace Eqs (3), (5) and (6) with the following equations:

$$
\begin{aligned}
\dot{X}_1 &= r_1 - k_1 X_1 \\
\dot{X}_2 &= r_2 S - k_2 X_2 \\
\dot{X}_3 &= r_3 P_2 - k_3 X_3 \\
\dot{Y}_1 &= K_1 X_1 - q_1 Y_1 \\
\dot{Y}_2 &= K_2 X_2 - q_2 Y_2 \\
\dot{Y}_3 &= K_3 X_3 - q_3 Y_3 \\
\dot{S} &= w_1^{(2)} Y_1 (1 - S) - w_{-1}^{(2)} S \\
\dot{P}_1 &= w_1^{(3)} Y_1 (1 - P_1 - P_2) + \tilde{w}_{-1}^{(3)} P_2 - w_1^{(3)} Y_2 P_1 - w_{-1}^{(3)} P_1 \\
\dot{P}_2 &= w_1^{(3)} Y_2 P_1 - \tilde{w}_{-1}^{(3)} P_2
\end{aligned}
\tag{11}
$$

for CFFM1, and

$$
\begin{aligned}
\dot{X}_1 &= r - k_1 X_1 \\
\dot{X}_2 &= r S - k_1 X_2 \\
\dot{X}_3 &= r_3 P - k_3 X_3 \\
\dot{Y}_1 &= K X_1 - q Y_1 \\
\dot{Y}_2 &= K X_2 - q Y_2 \\
\dot{Y}_3 &= K_3 X_3 - q_3 Y_3 \\
\dot{Z} &= a Y_1 Y_2 + w_{-1}^{(3)} P - w_1^{(3)} Z (1 - P) - (b + 2q) Z \\
\dot{S} &= w_1^{(2)} Y_1 (1 - S) - w_{-1}^{(2)} S \\
\dot{P} &= w_1^{(3)} Z (1 - P) - w_{-1}^{(3)} P
\end{aligned}
\tag{12}
$$

for CFFM2. As before, the variables X and Y denote mRNA and protein concentrations respectively, with the subscript indicating the gene, and Z refers to the heterodimer comprising of Y_1 and Y_2. For CFFM1, the association and dissociation rates between the protein (either monomer or heterodimer) and a gene's promoter are given by $w_1^{(i)}$ and $w_{-1}^{(i)}$ respectively, with the upper index indicating the gene: $i = A,B,C$. The rate at which a protein encoded by gene A binds with that of gene B is represented by a, while the rate of their dissociation is given by b. In deriving Eqs (11) and (12), we made the same assumptions (those relevant to the present system) listed at the beginning of section "Deterministic model". When all the steps of the previous section are carried out, the results are as presented in Fig 6.

The difference between the system of Fig 1B and the two types of the CFFM is immediately apparent: the distribution of $\Delta t / t_{av}$ for CFFM1 and CFFM2 is shifted to the right; the first is also more spread out. Interestingly, the best case for CFFM2 is nearly as good as the worst case for the system with a nucleus and importins.

Variations on the coherent feed-forward motif. Since in the nucleus bearing system the gene of interest auto-repressed itself, it seems only fair that we add such a feature to the CFFM. Instead of considering only the case where gene C represses itself, we look at two other variations on negative feedback within the CFFM: gene C represses gene A; and gene C represses gene B. The equations governing the promoter occupancy of the gene being repressed are the following:

Gene C represses gene A (CFFM1(-)):

$$
\begin{aligned}
\dot{P}_1 &= w_1^{(1)} Y_3 (1 - P_1 - P_2) + \tilde{w}_{-1}^{(1)} P_2 - w_1^{(1)} Y_3 P_1 - w_{-1}^{(1)} P_1 \\
\dot{w}_2 &= w_1^{(1)} Y_3 P_1 - \tilde{w}_{-1}^{(1)} P_2
\end{aligned}
\tag{13}
$$

Gene C represses gene B (CFFM2(-)):

$$
\begin{aligned}
\dot{S} &= w_1^{(2)} Y_1 (1 - S - P_1 - P_2) - w_{-1}^{(2)} p_1 \\
\dot{P}_1 &= w_1^{(2)} Y_3 (1 - S - P_1 - P_2) + \tilde{w}_{-1}^{(2)} P_2 - w_1^{(2)} Y_3 P_1 - w_1^{(2)} P_1 \\
\dot{P}_2 &= w_1^{(2)} Y_3 P_1 - \tilde{w}_{-1}^{(2)} P_2
\end{aligned}
\tag{14}
$$

Gene C represses itself (CFFM3(-)):

$$
\begin{aligned}
\dot{P}_1 &= w_1^{(3)} Y_1 (1 - P_1 - P_2 - P_3 - P_4) + w_{-1}^{(3)} P_2 - w_1^{(3)} Y_2 P_1 - w_{-1}^{(3)} P_1 \\
\dot{P}_2 &= w_1^{(3)} Y_2 P_1 - \tilde{w}_{-1} P_2 \\
\dot{P}_3 &= w_1^{(3)} Y_3 (1 - P_1 - P_2 - P_3 - P_4) + \tilde{w}_1^{(3)} P_4 - w_1^{(3)} Y_3 P_3 - w_{-1}^{(3)} P_3 \\
\dot{P}_4 &= w_1^{(3)} Y_3 w_1 - \tilde{w}_{-1}^{(3)} P_4
\end{aligned}
\tag{15}
$$

The new variables P_1 and P_2 refer to the promoter occupancy of gene C by one and two copies of Y_3 respectively. Repeating the now familiar procedures of sections "Stochastic model" and "Coherent feed-forward motif", renders the results shown in Fig 7. The profile for the best case, corresponding to the motif CFFM2(-), is shown in Fig 7D. The quantity $\Delta t / t_{av}$ in this case is 2 times greater than what it is for the worst case for the system with a nucleus, which is $\Delta t / t_{av} = 0.05$, and 6 times greater than its best case (Fig 4B). The eukaryotic motif wins out again.
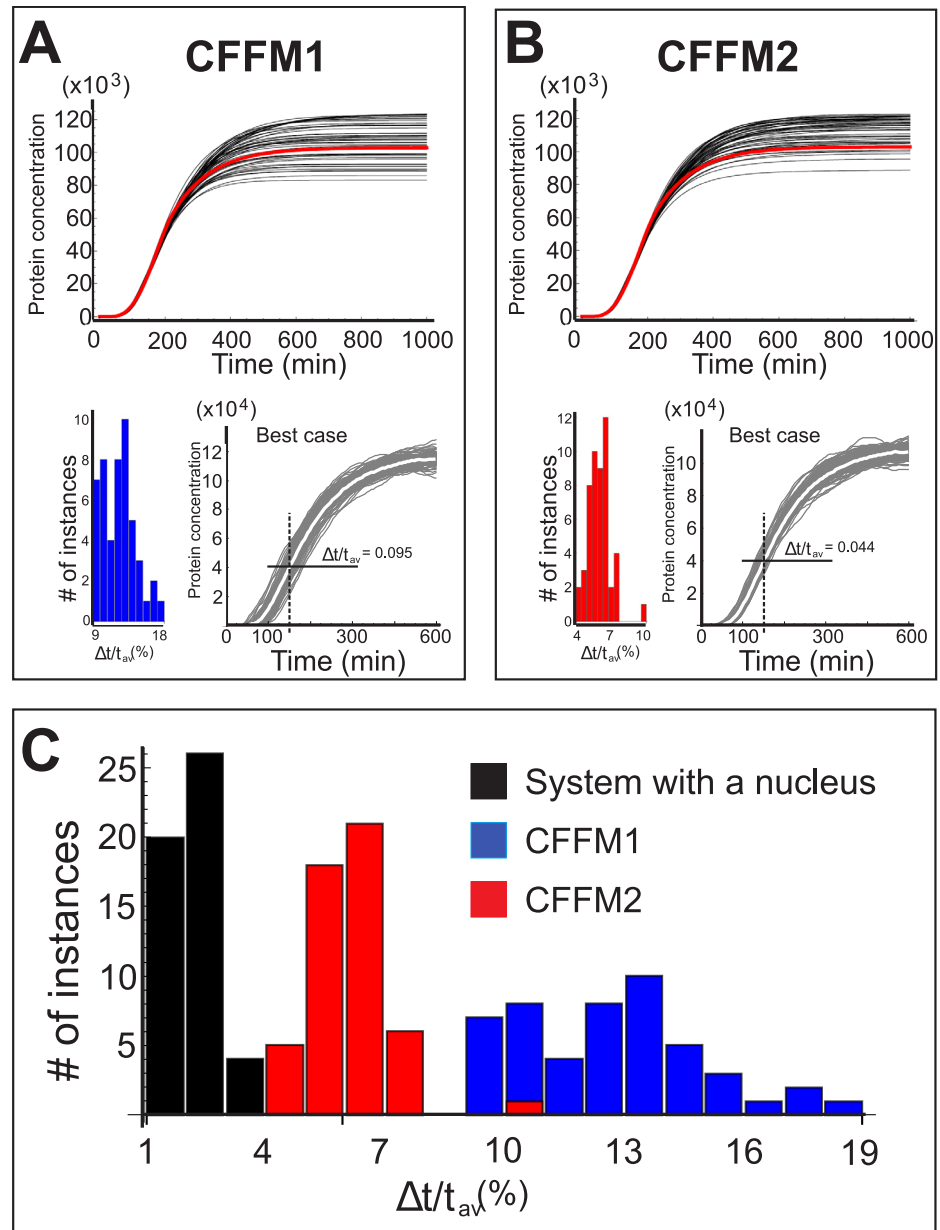
Fig 6. Results for CFFM1 and CFFM2. Shown here are the profiles of all fifty cases (top) and the resulting threshold times distributions (bottom), including the best case for A) CFFM1 and B) CFFM2. C) Threshold times distribution for the system with a nucleus (black) superimposed on CFFM1 (blue) and CFFM2 (red).

doi:10.1371/journal.pone.0134239.g006

Discussion

We have shown that in the eukaryotic cell, a negative feedback combined with an import of proteins from cytoplasm to nucleus, either active or diffusive, can not only lead to a delayed switch-like transition, but also to a highly precise timing of this transition. Conventionally, a negative feedback has been used to speed up the transition into a steady state, not slow it down. However, the presence of a nuclear barrier ensures that proteins inside the nucleus do not amass until the proteins in the cytoplasm gain abundance. Once they do, the nucleic proteins
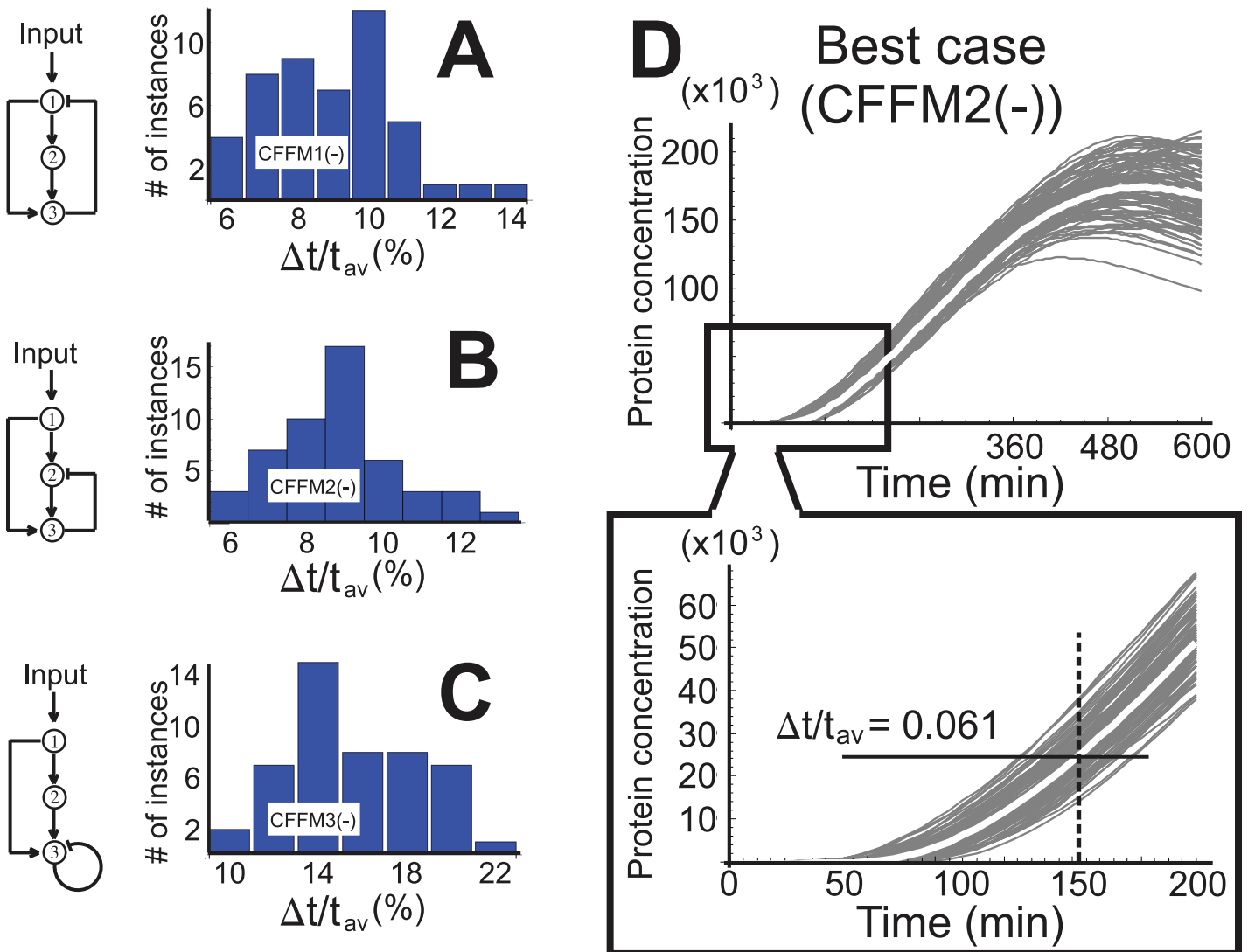
Fig 7. Three variations on the coherent feed-forward motif. The histograms in A, B and C show the distributions for the three variations on the CFFM the schematics of which are to the left. In all three cases, the transcription factors bind to the different promoters one at a time: no (hetero) dimers. D) Best case, corresponding to CFFM2(-) with $\Delta t/t_{av} = 0.061$.

doi:10.1371/journal.pone.0134239.g007

start growing at a large rate, giving rise to a switch-like behavior. As for the timing accuracy, this is likely a result of the fact that when a process, e.g. protein synthesis, consists of many steps, the uncertainty in "when" does the last step occur is inversely proportional to the number of steps [41]. In the present case, going from mRNA to nucleic protein, takes four steps, for the simple model, and five steps for the extended model (refer to Fig 1).

We note that the results of the extended model rely on the requirement that all three genes are turned on at the same time. In real systems this may not always be the case. In many situations the importins may have reached steady state well before the gene of interest is turned on. In fact, experimental studies reveal that neither case is strictly true; the importins themselves are regulated by other molecules, making their concentration vary over time [31–33]. For the sake of completion, however, we did consider a situation where, at time $t = 0$, all the gene products pertaining to importins are in a steady state. As an example, we took the parameters

corresponding to the best case scenario, shown in Fig 4D, and modified them such that the temporal profile of Y_n under the new conditions matches the reference profile Y_r according to the constraints of section "Stochastic model". To have a good match, we only had to let $q \rightarrow q/$ 2.4 and $a_2 \rightarrow a_2/50$. Fig 8 shows the result. The temporal precision, while still very good, is diminished by about 30%. This result does not necessarily mean that the precision is always diminished for the latter, as the parameters for the best case scenario may be completely different. An analysis similar to the one carried out in section "Stochastic model" will be necessary in order to determine whether starting out with importins, and the mRNAs from which they are translated, already in a steady state follow the distribution similar to the ones in Fig 4C. We
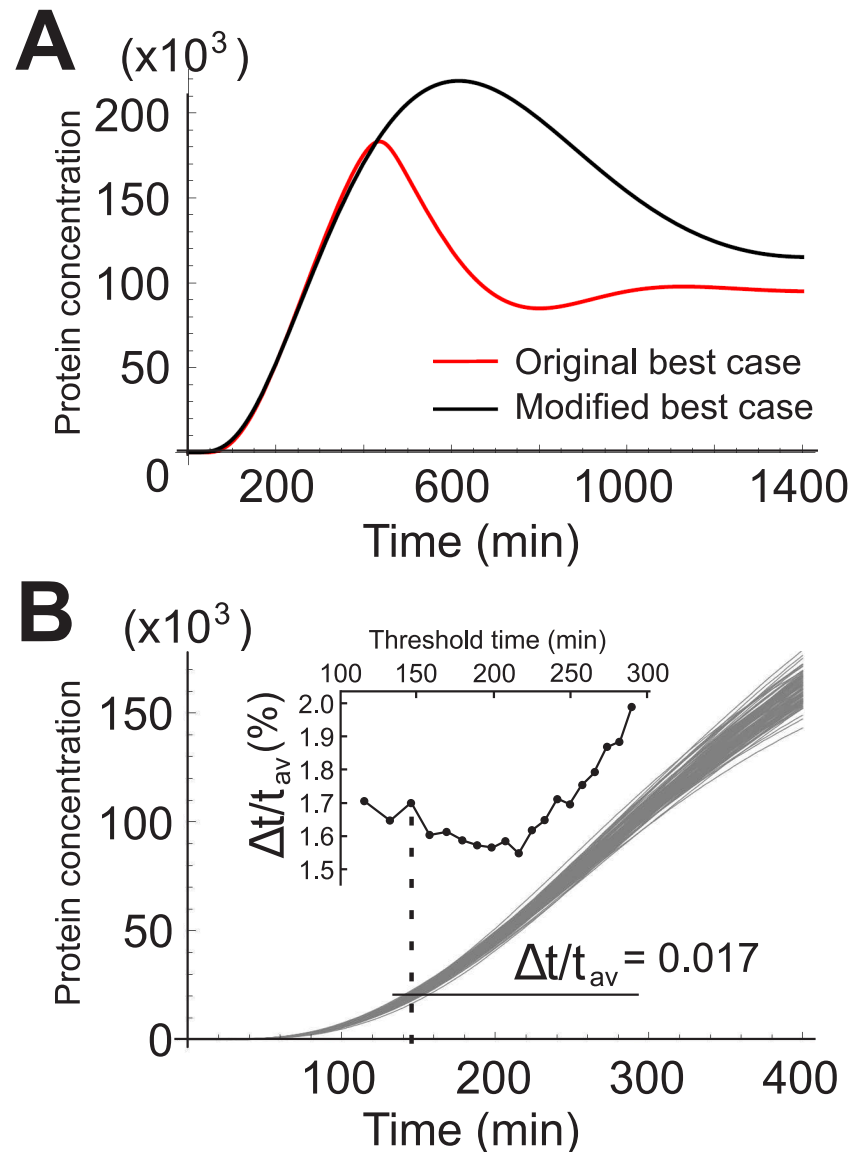
Fig 8. **Importins in a steady state at *t* = 0.** A) Average protein concentrations for Y_n. The red curve corresponds to the best case discussed in Fig 4. The black curve is the profile obtained when the concentrations $X_{\alpha n}$, $X_{\alpha c}$, $X_{\beta n}$, $X_{\beta c}$, $Y_{\alpha n}$, $Y_{\alpha c}$, $Y_{\beta n}$, $Y_{\beta c}$ and Z_2 start out in a steady state, and $q \rightarrow q/2.4$ and $a_2 \rightarrow a_2/50$. B) 100 profiles generated by the Gillespie algorithm. At *t* = 150, $\Delta t/t_{av} = 0.017$.

doi:10.1371/journal.pone.0134239.g008

leave this analysis for the future. The proven capability to control importins experimentally [42] makes our results testable.

With regards to the prokaryotic cell, we asked whether a genetic motif could be constructed such that its protein output would mimic the output of the nucleic protein. We considered two types of the coherent feed-forward motif, and three variations on it, and showed that while all of these motifs can be tuned to reproduce the temporal profile of the nucleic protein concentration, they cannot compete with the eukaryotic motif when it comes to temporal precision. The difference is not merely noticeable; it is remarkable. In the best case scenario, provided by the motif CFFM2, the relative deviation from the average threshold time amounted to 4.4%, a value 3.5 times greater than that found in the eukaryotic cell.

These findings suggest that one of the functions of the cell nucleus is to regulate fluctuations in protein abundance during transitions into steady states. From our limited number of cases, it appears that cells without a nucleus (and importins) do not have the means to reduce these fluctuations enough in order to effectuate reliable changes in any genetic program downstream. There may be, of course, larger motifs consisting of many genes that may cater to one specific gene in order to reduce noise associated with its corresponding protein. However, such gene regulatory networks, while conceivable, would demand a lot of energy to function. Research suggests [43] that early in the history of eukaryotes, energy was an important selection factor; networks that required less energy were more likely to persist. This begs the question: could organisms with complex spatio-temporal patterns have evolved without a nucleus? Or is the fact that in billions of years prokaryotes have not evolved into such organisms just a coincidence?

Supporting Information

S1 File. Lists of parameters for all the models and cases. The parameters are ordered as indicated in the "Table". Where applicable, other information is provided under each "Table". (DOC)

Acknowledgments

Author Contributions

Conceived and designed the experiments: JA. Performed the experiments: JA. Analyzed the data: JA MR. Contributed reagents/materials/analysis tools: JA MR. Wrote the paper: JA MR.

References

1. Pilpel Y. Noise in biological systems: pros, cons, and mechanisms of control. Methods Mol. Biol. 2011; 759: 409–25.

2. Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. Nature Reviews. 2010; 467: 09326

3. Balaban NQ, Merrin J, Chait R, Kowalik L, Leibler S. Bacterial persistence as a phenotypic switch. Science. 2004; 305: 1622–1625. doi: 10.1126/science.1099390 PMID: 15308767

4. Nachman I, Regev A, Ramanathan S. Dissecting timing variability in yeast meiosis. Cell. 2007; 131(3): 544–56. doi: 10.1016/j.cell.2007.09.044 PMID: 17981121

5. Fienerman O, Veiga J, Dorfman JR, Germain RN, Altan-Bonnet G. Variability and robustness in T cell activation from regulated heterogeneity in protein levels. Science. 2008; 321: 1081–1084. doi: 10.1126/science.1158013

6. Lewis J. Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator. Curr. Biol. 2003; 13 1398–408. doi: 10.1016/S0960-9822(03)00534-7 PMID: 12932323

7. Fraser HB, Hirsh AE, Giaever G, Kumm J, Eisen MB. Noise minimization in eukaryotic gene expression. PLOS Biol. 2004; 2(6): 0834. doi: 10.1371/journal.pbio.0020137

8. Joshi A, Beck Y, Michoel T. Post-transcriptional regulatory networks play a key role in noise reduction that is conserved from micro-organisms to mammals. FEBC Journal. 2012; 279: 3501–3512. doi: 10.1111/j.1742-4658.2012.08571.x

9. Little SC, Tikhonov M, Gregor T. Precise Developmental Gene Expression Arises from Globally Stochastic Transcriptional Activity. Cell. 2013; 154: 789–800. doi: 10.1016/j.cell.2013.07.025 PMID: 23953111

10. Dublanche Y, Michalodimitrakis K, Kümmere N, Foglierini N, Serrano L. Noise in transcription negative feedback loops: simulation and experimental analysis. Mol. Syst. Biol. 2006; 2: 41 doi: 10.1038/msb4100081 PMID: 16883354

11. Lestas I, Vinnicombe G, Paulsson J. Fundamental limits on the suppression of molecular fluctuations. Nature. 2010; 467: 09333 doi: 10.1038/nature09333

12. Swain PS. Efficient attenuation of stochasticity in gene expression through post-transcriptional control. J. Mol. Biol. 2010; 344: 965–976. doi: 10.1016/j.jmb.2004.09.073

13. Komorowski M, Miekisz J, Stumpf MPH. Decomposing Noise in Biochemical Signaling Systems Highlights the Role of Protein Degradation. Biophys. J. 2013; 104: 1783–1793. doi: 10.1016/j.bpj.2013.02.027 PMID: 23601325

14. Rooman M, Albert J, Duerinckx M. Stochastic noise reduction upon complexification: Positively correlated birth-death type systems. J. Theor. Biol. 2014; 354: 113–123. doi: 10.1016/j.jtbi.2014.03.007 PMID: 24632443

15. Bundchuh R, Hayot R, Jayaprakash F. The role of dimerization in noise reduction of simple genetic networks. J. Theor. Biol. 2003; 220:161–9.

16. Ghim CM, Almaas E, Jayaprakash F. Genetic noise control via protein oligomerization. BMC Syst. Biol. 2008; 2: 94. doi: 10.1186/1752-0509-2-94 PMID: 18980697

17. Arias AM, Hayward P. Filtering transcriptional noise during development: concepts and mechanisms. Nature Reviews. 2006; 7: 34–44. doi: 10.1038/nrg1750 PMID: 16369570

18. Houchmandzadeh B, Wieschaus E, Leibler S. Establishment of developmental precision and proportions in the early Drosophila embryo. Nature. 2002 415: 798–802. doi: 10.1038/415798a PMID: 11845210

19. Dahman C, Oates AC, Brand M. Boundary formation and maintenance in tissue development Nature Reviews. 2011; 12: 43–55. doi: 10.1038/nrg2902

20. Terry J, Sturrock M, Dale JK, Maroto M, Chaplain MAJ. A spatio-temporal model of Notch signaling in the zebrafish segmentation clock: conditions for synchronized oscillatory dynamics PLoS ONE. 2011; 6(2):e16980 doi: 10.1371/journal.pone.0016980 PMID: 21386903

21. Mather W, Hasty J, Tsimring LS. Synchronization of Degrade-and-Fire Oscillations via a Common Activator PRL 2014; 113: 128102

22. Sturrock M, Terry JA, Xirodimas DP, Thompson AM, Chaplain MAJ. Influence of nuclear membrane, active transport and cell shape on the Hes1 and p53-Mdm2 pathways: Insights from spatio-temporal modelling Bulletin of Math. Biol. 2012; 74: 1531–1579

23. Sturrock M, Hellander A, Matzavinos A, Chaplain MAJ. Spatial stochastic modeling of the Hes1 gene regulatory network: intrinsic noise can explain heterogeneity in embryonic stem cell differentiation Journal of the Royal Society Interface. 2013; 10:20120988 doi: 10.1098/rsif.2012.0988

24. Sturrock M, Murray PJ, Matzavinos A, Chaplain MAJ. Mean field analysis of a spatial stochastic model of a gene regulatory network J Math. Biol. PMID: 25323318

25. Mangan S, Zaslaver A, Alon U. The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. J. Mol. Biol. 2003; 334: 197–204 doi: 10.1016/j.jmb.2003.09.049 PMID: 14607112

26. Köhler A, Hurt E. Exporting RNA from the nucleus to the cytoplasm. Nature Reviews. 2007; 8: 761–773. doi: 10.1038/nrm2255 PMID: 17786152

27. Vargas DY, Ray A, Marras SAE, Kramer FR, Tyagi S. Mechanism of mRNA transport in the nucleus. PNAS. 2005; 102: 47: 17008–17013. doi: 10.1073/pnas.0505580102 PMID: 16284251

28. Macara IG. Transport into and out of the nucleus. Microbiol. Mol. Biol. Rev. 2005; 65(4): 570–594. doi: 10.1128/MMBR.65.4.570-594.2001

29. Timney BL, Tetenbaum-Novatt J, Agate DS, Williams R, Zhang W, Chait BT, et al. Simple kinetic relationship and nonspecific competition govern nuclear import rates in vivo. J. Cell. Biol. 2006; 175: 579–593. doi: 10.1083/jcb.200608141 PMID: 17116750

30. Riddick G, Macara IG. A system analysis of importin-α–β mediated nuclear protein import. J. Cell. Biol. 2005; 168: 1027–1038. doi: 10.1083/jcb.200409024 PMID: 15795315

31. Yasuhara N, Oka M, Yoneda Y. The role of the nuclear transport system in cell differentiation. Semin Cell Dev Biol. 2009; 20(5): 590–9. doi: 10.1016/j.semcdb.2009.05.003 PMID: 19465141

32. You P, Peng Z, Wang Y, Tao T. Expression and subcellular distribution of imp13 are regulated in brain development. In Vitro Cell Dev Biol Anim., 2013; 49(5): 346–353. doi: 10.1007/s11626-013-9599-z PMID: 23605716

33. Okada N, Ishigami Y, Suzuki T, Kaneko A, Yasui K, Fukutomi R, et al. Importins and exportins in cellular differentiation. J. Cell. Mol. Med. 2008; 12, No 5B: 1863–1871. doi: 10.1111/j.1582-4934.2008.00437.x PMID: 18657223

34. Bolouri H, Davidson EH. Transcriptional regulatory cascades in development: Initial rates, not steady states, determine network kinetics. PNAS. 2003; 100(16): 9371–9376. doi: 10.1073/pnas.1533293100 PMID: 12883007

35. de-Loen SB-T, Davidson EH. Modeling the dynamics of transcriptional gene regulatory networks for animal development. Dev. Biol. 2009; 325: 317–328. doi: 10.1016/j.ydbio.2008.10.043

36. Audibert A, Weil D, Dautry F. In vivo kinetics of mRNA splicing and transport in mammalian cells, Mol. Cell. Biol. 2002; 22(19): 6706–6718. doi: 10.1128/MCB.22.19.6706-6718.2002 PMID: 12215528

37. Zienkiewicz J, Armitage A, Hawiger J. Targeting nuclear import shuttles, importins/karyopherins alpha by peptide mimicking the NF κB1/p50 nuclear localization sequence. J. Am. Heart. Assoc. 2013 2 e000386. doi: 10.1161/JAHA.113.000386 PMID: 24042087

38. Catimel B, Teh T, Fontes MR, Jennings JG, Jans DA, Howlett GJ, et al. Biophysical characterization of interactions involving importin-α during nuclear import. J. Biol. Chem. 2001; 276: 34189–34198. doi: 10.1074/jbc.M103531200 PMID: 11448961

39. Zavala E, Marquez-Lago TT. Delays Induce Novel Stochastic Effects in Negative Feedback Gene Circuits Biophys. J. 2014; 106(2):467–478.

40. Bratsun D, Volfson D, Tsimring LS, Hasty J. Delay-induced stochastic oscillations in gene regulation PNAS. 2005; 102(41): 14593–14598. doi: 10.1073/pnas.0503858102 PMID: 16199522

41. Bel G, Munsky B, Nemenman I. The simplicity of completion time distributions for common complex biochemical processes. Phys Biol. 2010; 7, 016003. doi: 10.1088/1478-3975/7/1/016003

42. Quensel C, Friedrich B, Sommer T, Hartmann E, Kohler M. In vivo analysis of importin alpha proteins reveals cellular proliferation inhibition and substrate specificity. Mol Cell Biol. 2004; 24(23): 10246–55. doi: 10.1128/MCB.24.23.10246-10255.2004 PMID: 15542834

43. Wagner A. Energy Constraints on the Evolution of Gene Expression. Mol. Biol. and Evol. 2005; 22(6): 1365–1374. doi: 10.1093/molbev/msi126