Division time-based amplifiers for stochastic gene expression

Journal:	*Molecular BioSystems*
Manuscript ID:	MB-ART-06-2015-000391
Article Type:	Paper
Date Submitted by the Author:	11-Jun-2015
Complete List of Authors:	Zhou, Tianshou; SunYat-Sen University, School of Math & Comput Sci Wang, Haohua; Sun Yat-Sen University, Yuan, Zhanjiang; Sun Yat-Sen University, Liu, Peijiang; Sun Yat-Sen University,

Division time-based amplifiers for stochastic gene expression

Haohua Wang[1,2], Zhanjiang Yuan[1], Peijiang Liu[1] and Tianshou Zhou[1]

1. *School of Mathematics and Computational Science, Sun Yat-Sen University, Guangzhou 510275, People's Republic of China*

2. Department of Mathematics College of Information Science and Technology Hainan University, Haikou 570228, People's Republic of China

ABSTRACT While cell-to-cell variability is a phenotypic consequence of gene expression noise, sources of this noise may be complex: apart from intrinsic sources such as the random birth/death of mRNA and stochastic switching between promoter states, there are also extrinsic sources of noise such as cell division where division times are either constant or random. However, how this time-based division affects gene expression as well as how it contributes to cell-to-cell variability remains unexplored. Using a computational model combined with experimental data, we show that the cell-cycle lengths defined as the differences between two sequential division times can significantly impact expression dynamics. Specifically, we find that both divisions (constant or random) always raise the mean level of mRNA and lengthen the mean first passage time. In contrast to constant division, random division always amplifies expression noise but tends to stabilize its temporal level, and unimodalizes the mRNA distribution but makes its tail become longer. These qualitative results reveal that cell division based on time is an effective mechanism of both raising expression levels and enhancing cell-to-cell variability.

1. Introduction

Gene expression involves transcription, translation, transitions between promoter states, cell division, etc.[1-4] These biochemical processes are essentially all stochastic, resulting in stochastic fluctuations in mRNA and protein levels. These fluctuations can become important when the mRNA or protein copy number is low[5,6]. While quantitative time-lapse fluorescence microscopy allows for tracing the real-time

expressions of genes in individual cells,[7-9] there is considerable interest in theoretically understanding how different molecular mechanisms of gene expression affect variations in mRNA and protein abundances across a population of genetically identical cells.[10-16] Quantifying the contributions of different sources of noise using stochastic gene models is an important step towards understanding fundamental intracellular processes and cell-to-cell variability.[17-26]

Cell growth and division are ubiquitous in natural systems. When cellular energy and growth factors are enough, cell growth and division are all periodic events[27-30] (traditionally, cell division means that a cell begins to divide when its volume reaches twice the original volume, so this kind of division is actually volume-based division. In this paper, we consider time-based division by which we mean that a cell divides at time points that are either deterministic or random). The time difference between two adjacent divisions (called cell cycle)[29,31,32] can be regulated by the extracellular environment.[33-35] In particular, when living cellular physiological states are changed, the time points of cell division can also alter,[29,6-41] resulting in variation in cell cycle. In addition, the way of cell division may be diverse and division times may exhibit big differences between distinct cells. For example, B cells exposed to Toll-like receptor 9 ligands are nonself-adherent, allowing individual cells and families to be followed in vitro for up to 5 days.[42-44] They undergo phases typical of an adaptive response, dividing up to 6 times before losing the impetus for further growth and division and eventually dying by apoptosis.[42-44] Another example is fast growing *E. coli* cells dividing as frequently as every 20 minutes, and experimental observations show that cellular activity including continuous reproduction of genome is at an extreme level[37] but DNA replication is slowed down at a low nutrient level.[37-39] Variations in division times between otherwise identical cells are bound to generating intercellular differences in the copy numbers of stable mRNAs or proteins.[31-37]

Effects of cell division, in particular those of random division (by which we mean that division times are random throughout this paper), are seldom considered in previous models of gene expression.[27,28,35,45] In those models, fundamental expression events such as transcription/translation and stochastic promoter switching are often

modeled as time-continuous and space-discrete Markovian processes and hence the effect of expression noise can be captured by chemical master equations (CMEs).[11,12,16,17] However, cell division creates a discontinuity in the time evolution of the probability density function of the mRNA or protein number. In particular, cell division time may be random[31,42-44] and the partition of molecules into the daughter cells may be asymmetrical, each creating noise[9,27,28,46-49] that additionally contributes to expression noise. In addition, cell growth and division may affect several kinetic parameters including global transcription and translation rates,[37,49] thus controlling and/or determining cellular activity. While capturing effects of random cell division (an extrinsic source of expression noise) is a computational challenge, it is of particular interesting yet important to explore how cell division affects expression levels as well as how different sources of expression noise altogether contribute to cell-to-cell variability.

Many cellular processes including cell-cycle progression, signal transduction and cell apoptosis[29,35,5,51] all involve mRNA or protein degradation. Traditionally, dynamics of the gene-product number was described simply by a balance of synthesis and degradation.[2,4,15-19] However, recent genome-wide experiments quantifying intracellular mRNA and protein expression levels and turnovers in parallel in a population of unperturbed mammalian cells have revealed that many gene-encoded stable proteins are not actively degraded but are diluted through cell division[31,50,52] with random division times.[29,42] Also traditionally, the diluting effect of cell growth and division is determined by the growth rate $\lambda = \ln(V_1/V_0)/\tau$, where V_0 and V_1 stand respectively for the volumes at the begin and end of cell division in a cell cycle and τ for the mean cell cycle. In common cases, the growth rate is calculated by $\lambda = \ln(2)/\tau$.[9,52] Differently, we introduce the so-called equivalent degradation rate of gene product, which is similar to the diluting role of (deterministic or random) cell division, to capture the diluting effect of time-based division.

In a word, cell growth and division are inevitable events taking place in the process of gene expression, and biological experiments have also provided evidences

for random division times, but mechanisms of how time-based division affects gene expression are unclear. In this paper, we use a toy model to investigate how cell-cycle variability contributes to cell-to-cell variability in a population of genetically identical cells. Although simple, our model still captures major events taking place in gene expression such as the random synthesis and degradation of mRNA and stochastic switching between promoter activity states as well as cell growth and division. By model analysis, we highlight the roles of time-based division in regulating gene expression, which include: (1) this kind of division always raises the mean level of gene expression and lengthens the men first passage time (MFPT); (2) in contrast to division with a constant cell cycle, division with cell-cycle variability amplifies expression noise but tends to stabilize their transient level, and unimodalizes the mRNA distributions but makes their tails become longer. These results, which are independent of the choice of model parameter values and even gene models and are therefore qualitative, not only reveal that time-based division is an effective mechanism of both increasing expression levels and promoting cell-to-cell variability but also indicate that random division acts as an amplifier for gene expression.[53]

2. Results

2.1 Time-based division always raises the mean expression level

Figure 1 shows the dependence relationship between the stable mean mRNA number and the mean cell-cycle length in three cases of gene switching: slow; asymmetric and fast (see Methods wherein some analytical results are also given). From this figure, we observe that the mean mRNA level is a monotonically increasing function of the mean cell-cycle length, implying that time-based division always raises the mean expression level (this is an interesting yet seemingly-counterintuitive result since the longer the cell cycle is, the larger is the cell volume). In particular, for slow or fast switching (Fig. 1A or C), the raised amplitude is larger in the case of random division than in the case of constant division (comparing the red line with circles or the green line with squares with the blue line with stars). This indicates that

the noise from the random division (named as division noise) positively contributes to the expression level. In the case of asymmetric switching (Fig. 1B), however, this effect of division noise is not apparent possibly since the effect of the promoter noise from asymmetric switching partially counteracts the effect of division noise in this case.

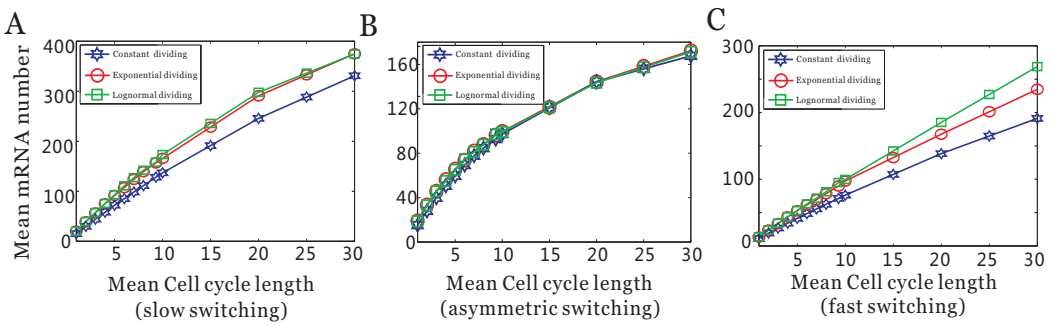


Fig. 1 The dependence of stable mean mRNA number on mean cell-cycle length: (A) slow switching; (B) asymmetric switching; (C) fast switching. The blue line with stars represents variations in the stable mean mRNA number in the constant division model; the red line with circles represents variations in the stable mean mRNA number in the exponential division model; and the green line with squares represents variations in the stable mean mRNA number in the lognormal division model.

We also observe from Fig. 1 that the increase of the stable mean mRNA number with the increase of the mean cell-cycle length is approximately linear in the case of slow or fast switching, implying that the equivalent degradation rate of mRNA due to cell growth and division is smaller in the former case than in the latter case. In contrast, this dependence relationship seems nonlinear in the case of asymmetric switching but the difference between the stable mean mRNA numbers in the random and deterministic cases of cell division is very small. This property to keep the mRNA number stable would imply that cells can use the internal noise (e.g., the promoter noise) to resist the external noise (e.g., the division noise generated from random division due to environmental fluctuations or changes in cellular physiological states) and to keep the intracellular mRNA level stable.[2,15-17] In addition, we observe that the stable mRNA number is smaller in the constant division case than in the stochastic

division case, and the difference becomes particularly apparent in the cases of symmetric switching. This would partially explain why stochasticity can make gene expression more efficient as well as why cells can use this noise for better survival. In addition, it would be used to interpret the molecular mechanism of why cancer cells are unstable (in this case, noise is large) but very vital.[41,51,54] The next two sections will give more explanations for the relevant molecular mechanism. Finally, this raised expression level due to cell division would be important for activation of downstream signals in gene regulatory networks.

3.2 Cell-cycle variability tends to stabilize the temporal level of mean expression

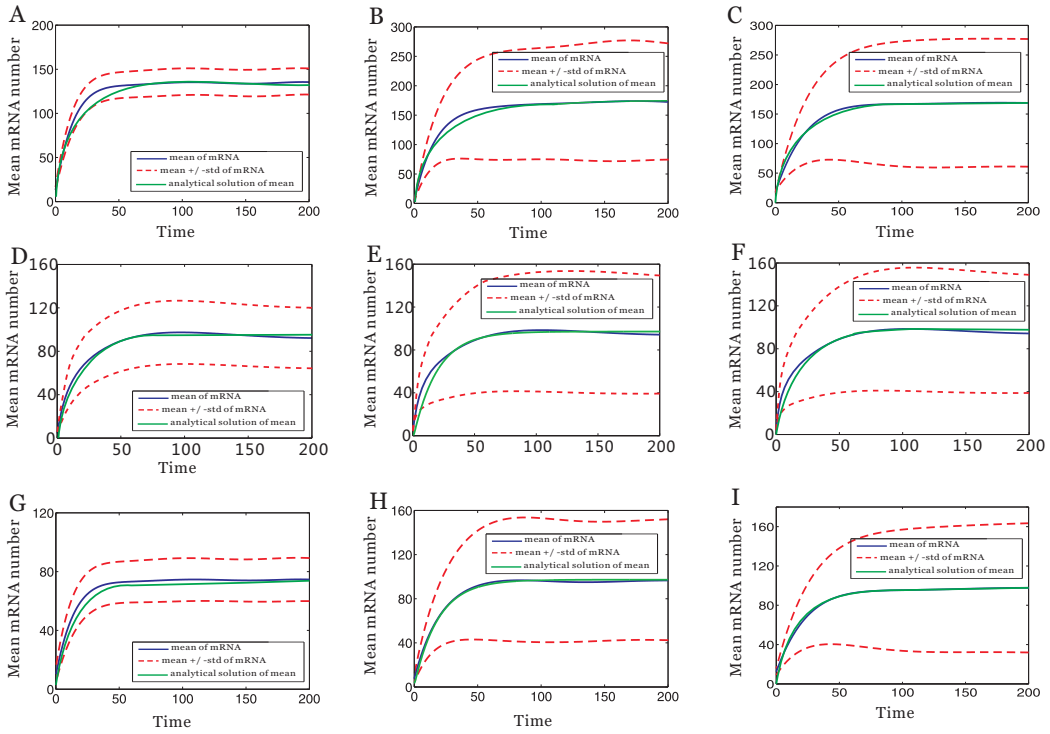


Fig. 2 Temporal changes of the mean mRNA number, where the solid line (blue) represents the mean whereas the dashed line (red) describes derivation from the mean. (A)(B)(C) correspond to slow switching, (D)(E)(F) to asymmetric switching, and (G)(H)(I) to fast switching. In (A)(D)(G), cell cycle is a constant; In (B)(E)(H), cell cycles are random, following an exponential distribution; In (C)(F)(I), cell cycles are also random but follow a lognormal distribution. The initial mRNA number is set as four. The green and blue solid lines represent the analytic and numerical solutions of

the time-dependent mean mRNA respectively, and the red dashed line to the mean plus/minus the standard deviation. In all the diagrams, every transient trajectory is obtained by averaging 10000 realizations.

The dependence curves shown in Fig. 1 well demonstrate the effects of cell division on the stable mean mRNA number, but more elaborate influence of the former on the latter should be reflected in the time evolution of the mean mRNA number. To demonstrate this temporal change, we plot Fig. 2 using the above-mentioned method. It is clearly seen from this figure that for each of two cell-division models, the mean mRNA number is time-dependent. Moreover, it first increases and then tends to a stable value after the time is sufficiently large. However, the mRNA variance in two cases of division time distribution is all larger than that in the case of constant division (comparing the first column with the second or third column).

Specifically, when promoter switching changes from fast to slow, the mean mRNA number increases but this increasing tendency is different from that in the case of constant division. Moreover, slow promoter switching makes the mean mRNA number drift towards the increasing direction, which is counter-intuitive in contrast to the case of stochastic gene expression with symmetric switching but without cell division.[53,55] Also, the mean mRNA number is higher in the case of random division than in the constant division, implying that gene expression in Model B (see Methods) is more efficient than that in Model A (also see Methods), although the amplitude of fluctuations becomes larger in the former than in the latter. This is not strange since the fact that cells can use stochasticity for better survival has been already reported in previous studies.[1,4,15,18,56] Compared subplots (B)(E)(H) with subplots (C)(F)(I), it is obvious that the difference in the mean mRNA number between two cases of random division is not large, but the fluctuation range in the case of lognormal division is larger than that in the case of exponential division, implying that even if the upstream noise level is fixed, the downstream noise may exhibit different levels possibly due to different molecular mechanisms.[3,15,17,24,46,54]

3.3 Cell-cycle variability amplifies expression noise

Cells that progress equally within each mean cell cycle can differ in the molecular content, because the number of species molecules produced and degraded varies between cells in time due to both intrinsic noise arising from biochemical reactions and cell-cycle variability from random division. The net variance in a molecule copy number is the result of the accumulation of various stochastic effects.[1,15,16,46] In principle, intrinsic noise and extrinsic noise altogether contribute to expression noise (or the total noise).[15,16,47,64] Some authors studied influences of the mean gene-product number on expression noise under different sources of noise,[1,15-17] and showed that the expression noise often becomes smaller when the mean number increases.[16,17,47,54] Here, we are interested mainly in effects of cell-cycle variability on expression noise. Note that if the mean cell cycle is lengthened, then the external noise will increase even though the stationary expression noise is approximately kept at a constant value even for different distributions of division times.

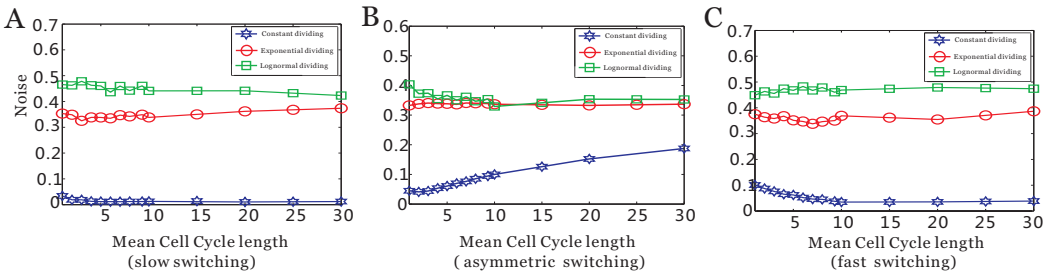


Fig. 3 The influence of the mean cell-cycle length on the stationary expression noise: (A) slow switching; (B) asymmetric switching; (C) fast switching. The blue line with stars represents the stable noise intensity in the case of constant division; the red line with circles sign represents the stable noise intensity in the case of exponential division; the green line with squares represents the stable noise intensity in the case of lognormal division.

First, we consider the stationary effect of cell-cycle variability on expression noise. For this, we plot Fig. 3, showing the dependence of the stable (or stationary) noise intensity on the mean cell-cycle length. From this figure, we first observe that the intensity of the total noise in the case of random division is always greater than

that in the case of constant division for the same mean cell-cycle length (comparing the red line with circles or the green line with squares with the blue line with stars), independent of promoter switching rates. This implies that random division always contributes positively to expression noise. Then, we observe that in the case of symmetric switching including fast or slow switching, the noise intensity in the case of lognormal distribution is larger than that in the case of exponential distribution, but the difference between them becomes smaller with the increase of the mean cell-cycle length. Third, we observe that the noise intensity is apparently a monotonic function of the mean cell-cycle length in the case of both constant division and asymmetric switching, but there is no apparent monotonic relationship between the noise intensity and the mean cell-cycle length in the other cases, in particular in cases of random division. Fourth, we observe that the noise intensity changes in a very narrow range in the cases of fast switching and slow switching when the mean cell-cycle length gradually increases, but the change range becomes relatively larger in the case of asymmetric switching, implying that the positive contribution of random division to expression noise depends on the symmetry of promoter switching.

Comparing Fig. 3 with Fig. 1, we see that the stationary intensity of expression noise is nearly independent of the stable mean mRNA number in the cases of random division. However, this intensity decreases to a stable value in a small amplitude manner under constant division plus symmetric switching. This change trend is opposite to that taking place in the case of asymmetric switching, possibly implying that a cell regulates the cell-cycle length to resist the interference of external stimulation or fluctuations in the cellular environment. Also, the change tendency would imply that there is a tradeoff between the cell-cycle length and the mean mRNA number in this case. The molecular mechanism behind it would be utilized by the cell to maintain the cytoactive or regulate its physiological states.[33-35]

3.4 Cell-cycle variability tends to stabilize the temporal level of expression noise

Here, we examine the temporal effect of cell-cycle variability on expression noise. This examination is more elaborate in contrast to the stationary case, and can

more precisely describe how cell-cycle variability affects temporal expression noise. The numerical results are shown in Fig. 4.

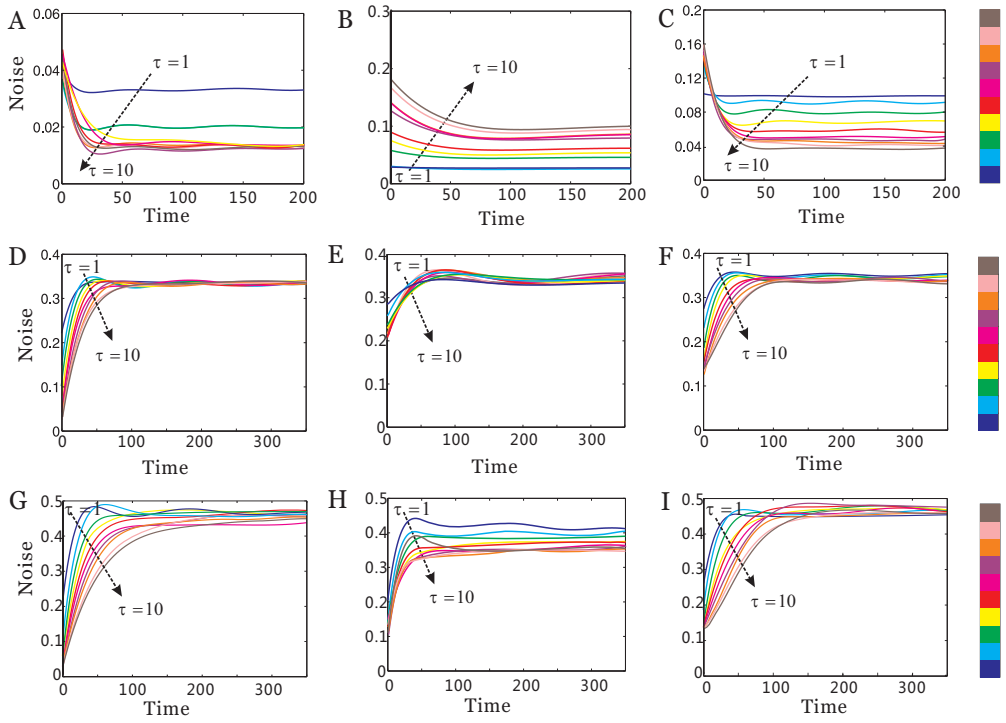


Fig. 4 The time evolution of expression noise for different mean cell-cycle lengths that changes from $\tau = 1$ to $\tau = 10$. (A)(B)(C) constant division; (D)(E)(F) exponential division; (G)(H)(I) lognormal division. (A)(D)(G) slow switching; (B)(E)(H) asymmetric switching; (C)(F)(I) fast switching. The color bar describes that different colors correspond to different cell cycles. In all the diagrams, every transient curve is obtained by averaging 10000 realizations.

From Fig. 4, we first observe that the intensity of expression noise tends to a constant value after the time is sufficiently large. This change trend is independent of the mode of cell division (constant or random), the speed of gene switching (slow, fast, asymmetric), and the mean length of cell cycle. Second, in the case of slow or fast switching, the noise intensity is fundamentally decreasing with the increase of the mean cell-cycle length at the initial stage but is almost independent of the size of the mean cell-cycle length after the time is beyond a certain critical value (see the first and third columns in Fig. 4). In contrast, the noise intensity in the case of asymmetric

switching may increase or decrease with the increase of the mean cell-cycle length at the initial stage, depending on the mode of division (constant or random). Third, in the case of constant division, the noise intensity is a monotonically decreasing function of the mean cell-cycle length in the case of slow or fast switching but this dependence relationship becomes opposite in the case of asymmetric switching. Fourth, the difference between noise intensities corresponding to the smallest and largest mean cell-cycle lengths in the case of constant division is apparently larger than that in the case of random division, implying that random division plays a role of stabilizing the level of temporal expression noise. Moreover, in the former case, different mean cell-cycle lengths lead to different levels of temporal expression noise that however are stable after the time is beyond a threshold (about 40 time units). Fifth, the noise intensity looks like having extreme values with regard to time in the case of random division but not in the case of constant division (see the diagrams in the second and third columns). Sixth, the stable noise levels in the case of constant division are much smaller than those in the case of stochastic division since the size of the former is higher one order than that of the latter, implying that random division positively contributes to expression noise. In addition, for constant division, the noise level in the case of slow switching is always lower than that in the case of fast switching for the same mean cell cycle. This is counter-intuitive since promoter noise in the former case is higher than that in the latter case,[17,23,54] but implies that the contribution of division noise to expression noise is inversely proportional to the speed of promoter switching.

The change trend of expression noise under constant division is distinct from that under random division, and in particular, the noise intensity in the latter case becomes more stable after a long time. With the increase of the mean cell-cycle length, the noise intensity becomes smaller and smaller in the case of both constant division and fast switching, but approaches to a stable value after the mean cell-cycle length is beyond a threshold ($\tau \approx 4$ time units). In addition, the noise intensity increases in the case of constant division plus asymmetric switching, but a different case takes place in the case of random division, where the noise intensity first increases and then

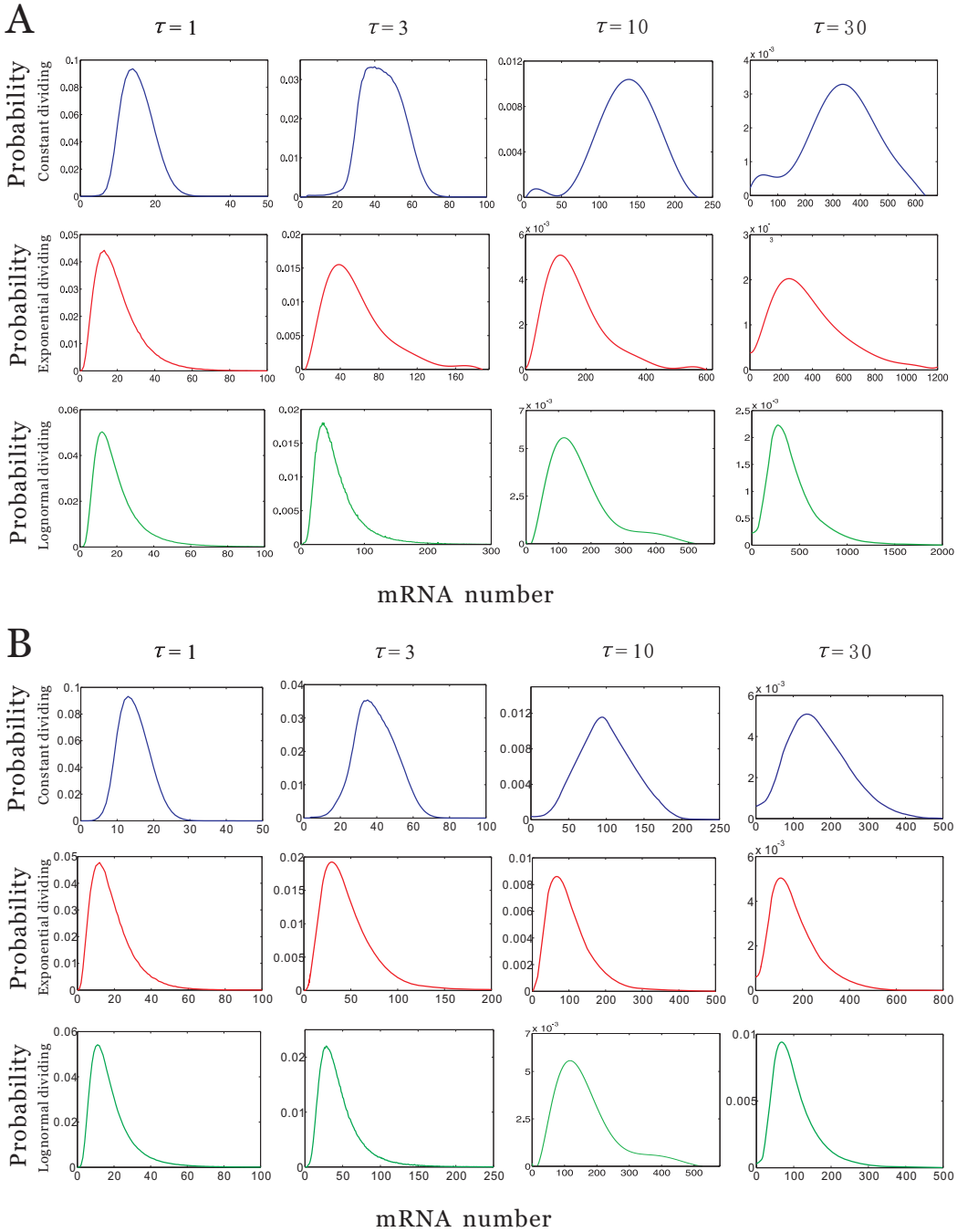
approaches to be an approximately stable value after a long time. Moreover, this stable value is approximately 0.46 in the case of lognormal division but 0.34 in the case of exponential division, indicating that the former is larger than the latter. In both cases, the stable noise intensity remains almost the same size, independent of promoter switching, referring to the diagrams in the first and third columns of Fig. 4 in the case of random division. This is also counterintuitive because the mean mRNA number increases with the increase of the mean cell-cycle length (see Fig. 1), and implies that the cell-cycle variability at a stable level can regulate the mean mRNA number while keeping the expression noise level invariant. Such a mechanism would be used by the cells to adapt to the change of physiology states such as cell growth, division and micro environments through regulating the cell-cycle length.

3.5 Cell-cycle variability induces long-tailed distributions

In contrast to single stochastic trajectories, the distribution can provide more complete information on stochastic properties of the underlying system. It is known that in changing environments, cells need to make decisions,[57] and phenotypic switching provides a way that cells survive in fluctuating environments. Several single-cell experiments have reported bimodal expression,[58,59] providing evidence that gene regulatory interactions can encode distinct phenotypes in isogenic cells. It is also known that gene regulatory networks with slow promoter switching may lead to distinct expression levels characterized by, e.g., bimodal distributions,[58,60] or by more general distributions such as long-tailed distributions[45] and mixed exponential distributions.[4,45] In spite of these, it is unclear how cell-cycle variability affects gene-product distributions.

Here, we first show that cell-cycle variability can change features of the mRNA distribution. The numerical results are shown in Fig. 5. From this figure we observe that in the case of constant division with a small cell cycle, the distribution of the mRNA number is unimodal for arbitrarily fixed promoter switching rates. With the increase of cell cycle, bimodality can appear in the case of symmetrical switching but needs a large cell cycle, implying that constant division tends to enhance bistability. In

the case of random division, however, the mRNA number does not follow a bimodal distribution but a long-tailed distribution. Moreover, this tail becomes longer when the mean cell-cycle length increases. The similar phenomena were also observed.[37,45] This implies that random division plays a role of suppressing bistability, different from the role of constant division. We also observe that when promoter switching rates change from fast to slow or vice versa, the mRNA distribution does



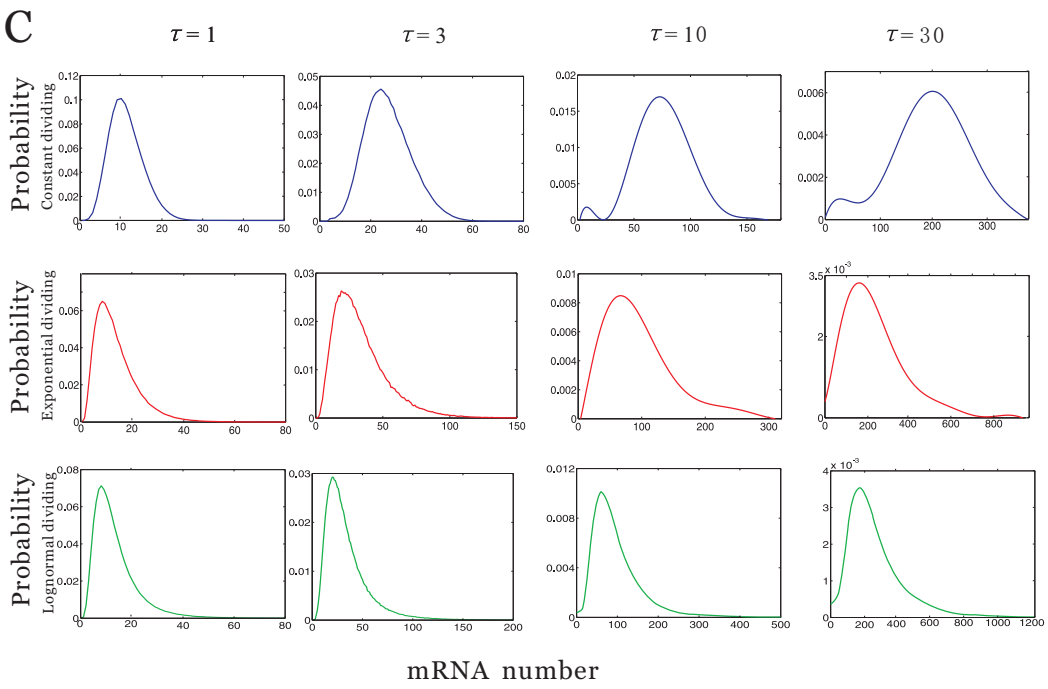


Fig. 5 The dependence of mRNA distribution on mean cell-cycle length: (A) slow switching; (B) asymmetric switching; (C) fast switching. The blue line represents the distribution in the case of constant division; the red line represents the distribution in the case of exponential division; the green line represents the distribution in the case of lognormal division. The first column corresponds to $\tau = 1$; the second column to $\tau = 3$; the third column to $\tau = 10$; and the forth column to $\tau = 30$. In every sub-diagram, the distribution is obtained by averaging 10000 realizations generated by GSSA.

not exhibit bimodality even in the case of constant division (comparing Fig. 5B with Fig. 5A or with Fig. 5C).

By combining Fig. 1 and Fig. 4, we know from Fig. 5 that expression noise is approximately stable in time although the mean mRNA number increases with the increase of the mean cell-cycle length. On the other hand, it is known that at the microscopic level, long-tailed distributions can occur in two cases: if the microscopic processes are inherently non-Gaussian (e.g., if there is multiplicative noise or if the mean numbers are small and there is a lower cutoff); if distributions are expressed as a superposition of Gaussian distributions, where means and variances are slightly

different in each case.[45] In addition, a long-tailed distribution means that the dynamic equilibrium of cellular stability reaches possibly through a multi-step process and the net level of gene expression at every step would be small under random division. Thus, our results imply that the random division strategy can increase the cellular survival probability when the nutrient resources are limited.[27,28,51]

We also investigate the effect of time-based division on promoter activity, and show that time-based division tends to enhance promoter activity. See Fig. S3 in the Supporting Material for details.

3.6 Cell-cycle variability lengthens MFPT

Here, we examine the influence of cell-cycle variability on FPT, where by FPT we mean the time at which a stochastic variable hits some critical threshold. It is known that the FPT can provide the information on stochastic properties of a dynamical system. Given this threshold at a fixed value of the mean cell-cycle length, we calculate MFPTs in three cases of cell division as well as in three cases of promoter switching. The numerical results are shown in Fig. 6.

From Fig. 6, we first observe that the MFPT increases approximately linearly with the increase of the mean cell-cycle length, regardless of promoter switching rates. Then, the MFPT curve in the case of random division is always below that in the case of constant division, indicating that random division shortens the MFPT. However, the distribution of division times has little influence on the MFPT (comparing red lines with green lines). Third, the MFPT is longer in the case of slow switching than in the case of fast switching for a same mean cell-cycle length, possibly since the mean mRNA number is lager in the former case than in the latter case (referring to Fig. 1 or Fig. 2).

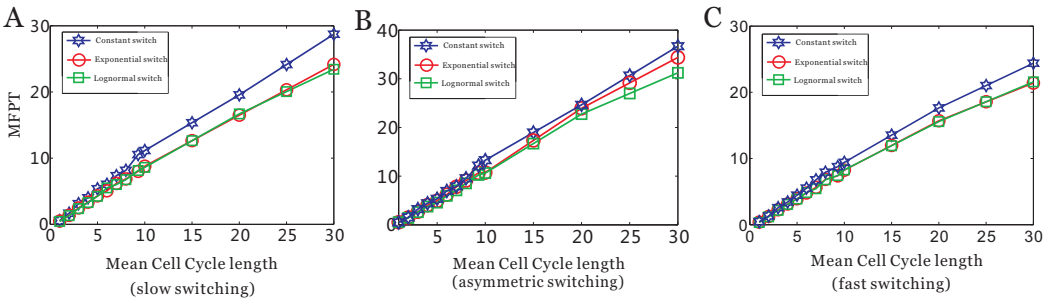


Fig. 6 The influence of mean cell-cycle length on MFPT: (A) slow switching; (B) asymmetric switching; (C) fast switching. The blue line with stars represents the MFPT in the case of constant division; the red line with circles represents MFPT in the case of exponential division; the green line with squares represents MFPT in the case of lognormal division.

By combining Fig. 1, Fig. 4 and Fig. 6, we know that random division can increase the mean mRNA number and lengthens the MFPT at the same time; the difference between the influences of exponential and lognormal divisions on the MFPT is not apparent; In contrast, the difference between the influences of these two kinds of divisions on the mean mRNA number or on the noise strength is apparent, reflecting that random division contributes positively to the expression level or expression noise.

3. Methods

3.1 Model description

To clearly show the effect of cell growth and division on gene expression as well as on cell-to-cell variability, here we consider a simplified model of gene expression at the transcription level (since our main interest is in effects of cell division, we ignore the translation process to simplify our analysis but it should be pointed out that this simplification does not influence the qualitative conclusions obtained in this paper). This model assumes that the gene promoter has two activity states: active (or ON) state at which transcription is very efficient and inactive (or OFF) state at which no transcription takes place (i.e., we do not consider promoter leakage,[53,61] although it is possible that a gene has more than two promoter states.[22,62-64] Furthermore, we suppose that there are stochastic transitions between ON and OFF states. This switching results in the bursty synthesis of mRNA,[20,65] which is a major source of cell-to-cell variability.[15-17,66] In addition, we hypothesize that the mRNA degradation is of one order (i.e., linear degradation) and is a single-step process although a multi-step process is possible in real cases.[67] Figure 7(A) schematically shows our

gene model.

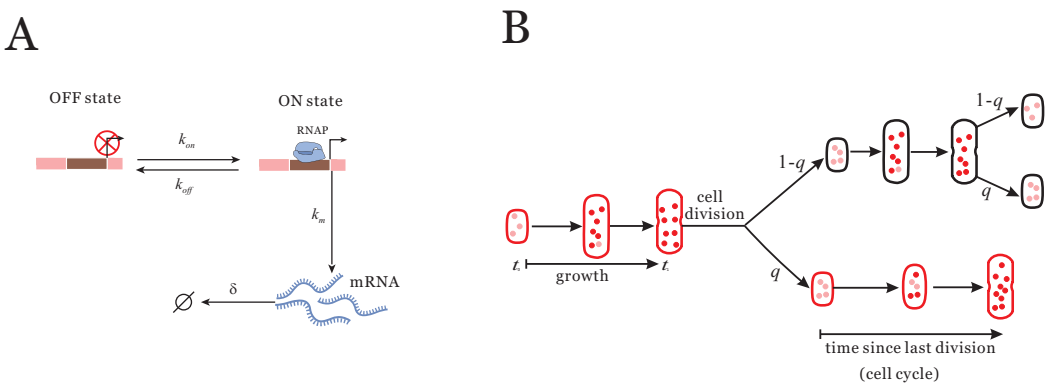


Fig. 7 Schematic diagrams for a two-state gene model considering the times of cell division: (A) the process of gene expression, where the gene promoter is assumed to have two active states (ON and OFF states) and there are transitions between ON and OFF states; (B) the process of cell division, where a mother cell is assumed to divide into two daughters at each time point of division in a probabilistic manner with probability q.

Our gene model also considers cell division. In general, there are two methods to describe cell division: the one is to consider that if the volume of a cell is increasing and finally reaches twice the original volume due to cell growth with a constant rate, then the cell begins to divide[9,48,49] (this description is termed as the volume-based division); the other is to consider that a mother cell divides to two daughters at every division point, and that a finite series of division times are either deterministic or random, depending on cellular environments or cellular physiological states[42-44.48] (this description is called as the time-based division). Here, we will adopt the latter to establish our model. Note that in the description of time-based division, a cell might not divide although its volume reaches the double of the original volume. Thus, stochastic gene models of time-based division are different from those of volume-based cell division. Knowing this difference is important for understanding the results obtained in this paper.

In order to better trace the time evolution of the joint probability of mRNA

(denoted by X), we introduce two factorial probabilities $P_0(m;t)$ and $P_1(m;t)$, which represent that X has m mRNAs at time t when the gene is at OFF and ON states, respectively. Thus, $P = P_0 + P_1$ represents the total probability. Let k_m be the transcription rate from DNA to mRNA. Let δ be the degradation rate of mRNA. Denote by k_{on} and k_{off} transition rates from OFF to ON and from ON to OFF, respectively. Let t_k, $0 \le k \le K$, represent division time points. To that end, if we note that no transcription takes place at OFF state, then the CME corresponding to the reaction network schematized in Fig.7 can be described as

$$\begin{aligned}\frac{\partial P_0(m;t)}{\partial t} &= -k_{on}P_0(m;t) + k_{off}P_1(m;t) + \delta(E-I)\left[mP_0(m;t)\right] \\ \frac{\partial P_1(m;t)}{\partial t} &= k_{on}P_0(m;t) - k_{off}P_1(m;t) + k_m\left(E^{-1}-I\right)\left[P_1(m;t)\right] + \delta(E-I)\left[mP_1(m;t)\right]\end{aligned} \tag{1}$$

where I is the unity operator, and E with the inverse E^{-1} is the common step operator defined as $Ef(m) = f(m+1)$ for any function f. Note that Eq. (1) holds only for $t \in R_+ \setminus \{t_k\}_{0\le k\le K}$, where every t_k represents a time at which the cell divides and R_+ the positive half axis. Therefore, Eq. (1) is a piecewise stochastic system.[31,37,45]

Denote by τ_k the difference between two sequential division times, that is, $\tau_k = t_k - t_{k-1}$ with $t_0 = 0$, which represents a length of cell cycle, $1 \le k \le K$. To better show effects of cell-cycle variability on expression dynamics, we consider two kinds of time-based divisions: constant and random (referring to Fig. 8). For the former, all the τ_k are the same constant (denoted by τ), and the corresponding model is called as constant division model (or denoted by Model A), whereas for the latter, every τ_k is a random variable and the corresponding model is called as random division model (or denoted by Model B).

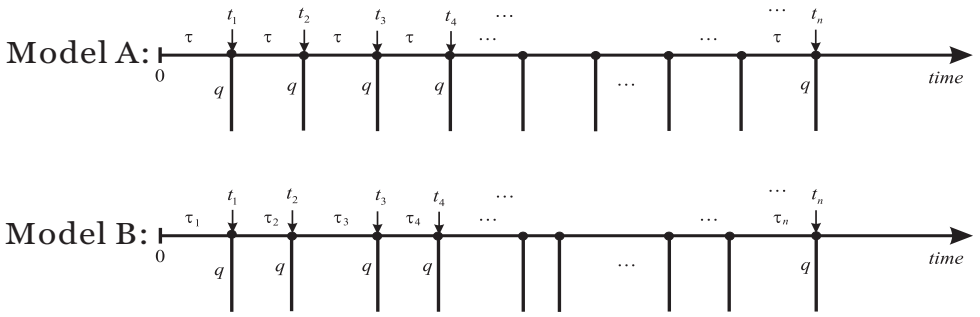


Fig. 8 Schematic diagrams for two kinds of time-based divisions: in Model A, cell cycle is a constant, i.e., $\tau_k = \tau$ for any k, and in Model B, cell cycle is a random variable.

As usual, we assume that a series of cell cycles $\{\tau_k\}_{1\le k\le K}$ follow an independent and identical distribution with finite moments and denote by τ the first-order moment that represents the mean cell-cycle length. To that end, we finish the description of our model.

3.2 Experimental data-based hypotheses of division events

First, we assume that the cell divides symmetrically, i.e., the division probability is $q = 1/2$.[27,28,31,37] This means that the mRNA number is halved at every division time point t_k, i.e., $m(t_k) = m(t_k^-)/2$ with $1 \le k \le K$, where the symbol $m(t_k^-)$ represents the left limit as the time t approaches t_k from the left of t_k.

Second, note that division times (in particular the mean division time) are experimentally measurable and the corresponding distribution can be fitted using experimental data.[29,42,68,69] In fact, the division tracking dye, carboxyfluorescin diacetate succinimidy ester (CFSE) is currently the most informative labeling technique for characterizing the division history of cells in immune systems. Gett and Hodgkin[68] pioneered the quantitative analysis of CFSE data, and used direct fitting, indirect fitting or rescaling methods to estimate the average division time or the mean cell cycle. In addition, mathematical models based on the Smith–Martin concept of two phases of the cell cycle have been successfully applied to division tracking

data.[29,42,68,69] In order to quantify the division-time distribution, Hawkins, et al., investigated B lymphocytes using an experimental model and found that division times follow a skewed or lognormal distribution,[42,44] referring to Fig. 9. In addition, most common distributions to model cell cycle include Erlang, uniform, lognormal and exponential distributions.[29,31,42-44] Here, we will use the latter two distributions for investigation.

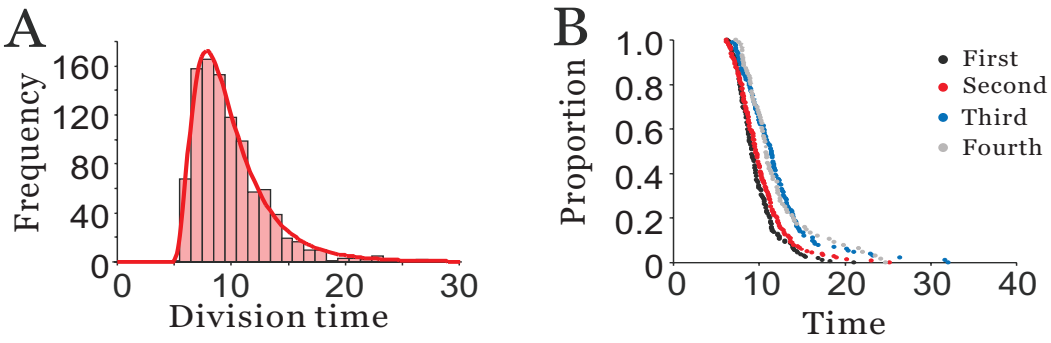


Fig. 9 Experimental data for variation in subsequent B-cell division times represented as a histogram collated in 1-h time intervals: (A) the distribution of division times, where the solid line represents a fitted lognormal distribution with the mean $\tau = 9.3$ and the variance $\sigma = 2.54$; (B) the proportion of division events, where "First" represents the first division (similarly for the others).

To demonstrate effects of cell division on gene expression clearly, we distinguish three cases of promoter switching rates: fast switching (i.e., two equal switching rates are equal but large) and slow switching (i.e., two equal switching rates are equal but small), each belonging to symmetric switching; and asymmetric switching where two switching rates are unequal. This classification considers the characteristic of gene promoter in eukaryotic cells.[22] In simulation throughout this paper, we set $k_{on} = k_{off} = 1$, $k_m = 10$ (fast switching); $k_{on} = 0.01$, $k_{off} = 0.1$, $k_m = 10$ (asymmetric switching); and $k_{on} = k_{off} = 0.01$, $k_m = 10$ (slow switching). All the parameter values that are dimensionless are estimated according to recent experimental data on mammalian gene expression, seeing Refs. (31,42) and references therein.

3.3 Analysis method

As pointed out above, there are two models to describe cell division: volume-based division models, and time-based division models. Note that numerical analysis of the latter like Eq. (1) is not easy since the models are piecewise stochastic systems.[31,37,45,48] Therefore, it is necessary to state our numerical method clearly.

For a given distribution (exponential or lognormal), we first generate a series of division times randomly, $\{\tau_k\}_{1\le k\le K}$, whose mean is denoted by τ (representing the mean cell-cycle length), and variance by σ. Then, we solve the CME (1) using the Gillespie stochastic simulation algorithm (GSSA).[37,53,70] To model the effect of time-based division, we generate 10000 runs at every time point (for convenience, we choose some time step such that the mean cell-cycle length is an integer multiple of the time step). When we calculate every stochastic trajectory, the mRNA number must be halved as long as the advancing time reaches a time division point (see Fig. S1 and Fig. S2 in the Supporting Material). Moreover, half the resulting values at the time end of the last cell cycle are taken as the initial values of solving the CME in the next cell cycle. Third, if the advancing time length reaches some value (e.g., 200 time units), then we view that the resulting mean mRNA number is steady-state or stable. By averaging these sets of trajectory data, we obtain the time-evolutional mean mRNA number.

To show the diluting effect of time-based division clearly, we may set $\delta = 0$ without loss of generality. Thus, we obtain a stable mean mRNA (denoted by $\langle m\rangle_{ss}$) using the above numerical method. Furthermore (see the Supporting Material), we obtain the analytical expression for the equivalent degradation rate ($\tilde{\delta}$) below

$$\tilde{\delta} = \frac{k_m k_{on}}{\langle m\rangle_{ss}\left(k_{on} + k_{off}\right)} \tag{2}$$

Note that $\langle m\rangle_{ss}$ that has been obtained using the above numerical method should depend on the given distribution of division times (denoted by $p_{div-dis}$), in particular on its mean (τ) and variance (σ). Thus, $\tilde{\delta}$ has to depend on $p_{div-dis}$, in particular

on τ and σ. This implies that the effect of extrinsic noise due to randomness of division times can be reflected in the equivalent degradation rate.

In turn, we can give the analytical expression of the time-evolutional mean mRNA number using the known equivalent degradation rate (see the Supporting Material). That is,

$$\langle m\rangle(t)=\frac{k_m k_{on}}{\tilde{\delta}(k_{on}+k_{off})}+\frac{k_m k_{on}}{(k_{on}+k_{off})(k_{on}+k_{off}-\tilde{\delta})}e^{-(k_{on}+k_{off})t}-\frac{k_m k_{on}}{\tilde{\delta}(k_{on}+k_{off}-\tilde{\delta})}e^{-\tilde{\delta}t} \qquad (3)$$

This indicates that the effect of random division is transferred to the time-dependent mean mRNA number through the equivalent degradation rate.

As examples, here we give the values of equivalent degradation rate for some models of cell division with different distributions of division times. See Table 1 for details, where we set the mean cell-cycle length is set as $\tau = 9.3$, which is estimated based on the experimental data of B cells,[42] and the mean cell growth rates are calculated by the formula $\lambda=\ln(V_1/V_0)/\tau$ whereas the equivalent degradation rates ($\tilde{\delta}$) by the formula (2). These examples indicate that different distributions of division times lead to different equivalent degradation rates although the cell growth rates may be the same (comparing slow switching with fast switching), implying that time-based division is different from volume-based division.

Table 1 Cell growth rates (λ) and equivalent degradation rates ($\tilde{\delta}$) for different gene models of cell division.

Mode of Switching	growth rate (λ)	degradation rate ($\tilde{\delta}$) (constant divide)	degradation rate ($\tilde{\delta}$) (exponential divide)	degradation rate ($\tilde{\delta}$) (lognormal divide)
Slow	0.0745	0.0375	0.0292	0.0298
Fast	0.0745	0.067	0.0529	0.0512
Asymmetric	0.0134	0.0098	0.0094	0.0094

Next, we give a formula to calculate the intensity of expression noise defined as the ratio of variance $Var(t,p_{div-dis})$ over the square of mean $\langle m(t,p_{div-dis})\rangle$, denoted

by η_m^2. This formula is given by

$$\eta_m^2(t) = \frac{Var(t, p_{div-dis})}{\langle m(t, p_{div-dis}) \rangle^2} \quad (4)$$

which is in particular a function of both the mean (τ) and the variance (σ) of the distribution of division times. Note that this noise (actually the total noise) contains the promoter noise due to stochastic switching between gene states, the transcriptional noise due to the synthesis and degradation of mRNA and the division noise due to random division times. It has been shown that expression noise can be decomposed into the sum of promoter noise and transcriptional noise if cell division is not considered.[2,15,16] However, it seems difficult to separate division noise from the total noise.

It should be pointed out that we actually consider only the mean cell growth rate described by $\lambda = \ln(V_1/V_0)/\tau$ where τ is the mean cell cycle, rather than the "point" growth rate described by $\lambda = \ln(V_1/V_0)/\tau_k$ where τ_k representing a cell-cycle length is random and follows a distribution. In addition, we emphasize that our time-based division might be different from volume-based division. More precisely, in our case, a cell may begin to divide when its volume does not reach or is beyond the double of the original volume, whereas in the case of volume-based division, a cell begins to divide as long as its volume reaches twice the original volume.

4. Conclusion and discussion

Sources of expression noise may be diverse mainly due to complexity of gene expression processes, but can be simply classified as intrinsic and extrinsic sources.[1,15-17,47] Intrinsic noise results from stochasticity of intracellular chemical kinetics when the numbers of interacting molecules are very small whereas extrinsic noise originates either from extracellular signals regulating intracellular processes or from cellular environmental perturbations resulting in fluctuations in intracellular reaction rates, or from both. The former can be described by a CME, and in essence

represents deviation of known reactions with known rates from their results predicted by classical chemical kinetics.[12,15-17,47] In contrast, the latter may result from any process not represented in the network model itself and therefore may be more complex.[15-17,21,24] Gene expression, in particular expression noise, has been extensively studied, but the effect of random cell division (a ubiquitous process taking place in natural gene systems) on gene expression is rarely considered. Here, we have analyzed a representative model of gene expression, which, although simple, contains noise of three kinds: promoter noise due to stochastic transitions between promoter activity states, transcriptional noise due to the birth and death of mRNA, and division noise due to random division times that result in cell-cycle variability. By analysis, we have shown that time-based division, in particular random division, acts as an amplifier for gene expression, which can not only significantly raise the mean level of gene expression but also always amplify expression noise and hence enhances cell-to-cell variability important for cellular survival in changing environments.

As mentioned in the introduction, time-based division leads to a piecewise (i.e., discontinuity in time) stochastic system. Numerical simulation of such a kind of system is a challenge. Here, we have developed an effective method for capturing the effect of random division times. Although its initial aim is in dealing with the numerical calculation of a simplified model of gene expression with time-based division, this method can be easily extended to more complex cases of gene expression, including RNA nuclear retention,[71-72] stochastic transitions among many promoter states,[62-64] multi-step synthesis or multi-step degradation,[67] alternative splicing,[73] and chromatin remodeling.[22,74,75] This is because these processes, which take place at the inside of cells and are hence intrinsic sources of expression noise, are in general linear, thus being easily incorporated into CMEs.[12,16,54] By calculating the corresponding equivalent degradation rates, we can use them to describe the effects of time-based division on expression dynamics, as done in this paper. This method of capturing the diluting effect of time-based division even can be extended to those cases where a series of cell-cycle lengths may not be independent of one another nor may follow the same distribution, the cell volume is changeable, and the partitioning

of molecules between daughter cells at the time of cell division is stochastic and modeled not necessarily by the $1/2$ probability used here but possibly by a binomial distribution.[27,28,48] In a word, when more complex situations underlying sources of expression noise are considered, the qualitative conclusions obtained here should not be completely ruined but are even model-free although quantitative results would be different.

An exceptional case for our qualitative results would be feedback regulation (a universal mechanism of controlling the information flow in gene regulatory networks), which leads that the underlying piecewise stochastic system is nonlinear. As is well known, positive feedback can amplify expression noise whereas negative feedback can reduce expression noise under conditions of some hypotheses.[53] However, it is not clear whether these conclusions are still maintained in the presence of time-based division, in particular random division. This requires further investigation since in the case that cell division is ignored, feedback has the potential to generate stochastic focusing[76] that like the role of time-based division shown here, can also raise the mRNA number or concentration, thus influencing the diluting effect of cell division.

From a viewpoint of biology, gene expression may follow some design principles for optimal evolutionary fitness, implying that gene expression is locally and globally constrained.[30,50] One constraint may be in efficiency.[30,50] We have shown that the mRNA number is distinctly larger in the case of random division than in the case of constant division, implying that the efficiency is higher in the former case than in the latter case. Similarly, our results indicate that the efficiency is higher in case of slow switching than in the case of fast switching, implying that slow promoter switching would be a predominant mode in the case of random division, consistent with the characteristic of eukaryote cells.[4] Another constraint is in the ability that a gene responds quickly to external stimuli.[30,50] We have shown that the MFPT is shorter in the case of random division than in the case of constant division, and that the MFPT in the case of slow switching is nearly as large as that in the case of fast switching (Fig. 6). These would imply that random division enhances the ability that genes deal

with external signals. When the “efficiency” and the “ability” in the above sense are considered simultaneously, there would be a tradeoff relationship between them, which is worth further investigation.

Author Contributions

Conceived and designed the experiments: HH ZJ PJ TS. Analyzed the data: HH PJ ZJ. Wrote the paper: TS HH.

Acknowledgments

This work was partially supported by Grant Nos. 91230204 (TS), 30973980 (TS), 11005162 (ZJ), 2010CB945400 (TS), 114002 (HH), 31460420 (HH) and 71161008 (HH).

References

1. E.M. Ozbudak, M. Thattai, I. Kurtser, A.D Grossman, and A. van Oudenaarden, *Nat. Genet.*, 2002, **31**, 69-73.
2. M.B. Elowitz, A.J. Levine, E.D. Siggia, and P.S. Swain, *Science*, 2002, **297**, 1183-1186.
3. W.J. Blake, M. Kærn, C.R. Cantor, and J.J. Collins, *Nature*, 2003, **422**, 633-637.
4. J.M. Raser, and E.K. O'Shea, *Science*, 2004, **304**, 1811-1814.
5. S.B.T. de-Leon, and E.H. Davidson, *Dev. Biol*., 2009, **325**, 317-328.
6. S. Komili, and P.A. Silver, *Nat. Rev. Genet*., 2008, **9**, 38-48.
7. J.C. Locke, and M.B. Elowitz, *Nat. Rev. Microbiol*., 2009, **7**, 383-392.
8. D. Muzzey, and A. van Oudenaarden, *Annu. Rev. Cell. Dev. Biol*., 2009, **25**, 301-327.
9. D. Gonze, *J. Theor. Biol.*, 2013, **325**, 22-33.
10. O.G. Berg, *J. Theor. Biol.*, 1978, **71**, 587-603.
11. J. Peccoud, and B. Ycart, *Theor. Popul. Biol.*, 1995, **48**, 222-234.
12. H.H. McAdams, and A. Arkin, *Proc. Natl. Acad. Sci. U.S.A.*, 1997, **94**, 814-819.
13. M. Thattai, and A. van Oudenaarden, *Proc. Natl. Acad. Sci. U.S.A.*, 2001, **98**,

8614-8619.

14. T.B. Kepler, and T.C. Elston, *Biophys. J*., 2001, **81**, 3116-3136.
15. J. Paulsson, *Nature*, 2004, **427**, 415-418.
16. J. Paulsson, *Phys. Life. Rev.*, 2005, **2**, 157-175.
17. J.M. Raser, and E.K. O'Shea, *Science*, 2005, **309**, 2010-2013.
18. N. Friedman, L. Cai, and X.S. Xie, *Phys. Rev. Lett*., 2006, **97**, 168302.
19. V. Shahrezaei, and P.S. Swain, *Proc. Natl. Acad. Sci. U.S.A.*, 2008, **105**, 17256-17261.
20. J.M. Pedraza, and J. Paulsson, *Science*, 2008, **319**, 339-343.
21. A. Sánchez, and J. Kondev, *Proc. Natl. Acad. Sci. U.S.A.*, 2008, **105**, 5081-5086.
22. A. Sanchez, H.G. Garcia, D. Jones, R. Phillips, and J. Kondev, *PLos Comput. Biol.*, 2011, **7**, e1001100.
23. A. Sanchez, S. Choubey, and J. Kondev, *Methods*, 2013, **62**, 13-25.
24. T. Jia, and R.V. Kulkarni, *Phys. Rev. Lett.*, 2011, **106**, 058102.
25. M. Dobrzyński, and F.J. Bruggeman, *Proc. Natl. Acad. Sci. U.S.A*., 2009, **106**, 2583-2588.
26. R. Rinott, A. Jaimovich, and N. Friedman, *Proc. Natl. Acad. Sci. U.S.A*., 2011, **108**, 6329-6334.
27. D. Huh, and J. Paulsson, *Nat. Genet*., 2011, **43**, 95-100.
28. D. Huh, and J. Paulsson, *Proc. Natl. Acad. Sci. U.S.A.*, 2011, **108**, 15004-15009.
29. J.A. Smith, and L. Martin, *Proc. Natl. Acad. Sci. U.S.A.*, 1973, 70, 1263-1267.
30. C. Li, and J. Wang, *Proc. Natl. Acad. Sci. U.S.A*., 2014, **111**, 14130-14135.
31. D. Antunes, and A. Singh, *J. Math. Biol.*, 2014, **64(5)**, 829-854.
32. B. Sun, L.S. Looi, S. Guo, Z. He, E.S. Gan, J. Huang, and T. Ito, *Science*, 2014, **343**, 1248559.
33. M. Scott, C.W. Gunderson, E.M. Mateescu, Z. Zhang, and T. Hwa, *Science*, 2010, **330**, 1099-1102.
34. M. Scott, and T. Hwa, *Curr. Opin. Biotech*., 2011, **22**, 559-565.
35. S. Klumpp, and T. Hwa, *Curr. Opin. Biotech*., 2014, **28**, 96-102.
36. S. Cooper, and C.E. Helmstetter, *J. Mol. Boil*., 1968, **31**, 519-540.

37. J. Jaruszewicz, M. Kimmel, and T. Lipniacki, *Phys. Rev. E,* 2014, **89**, 022710.

38. S.T. Liang, Y.C. Xu, P. Dennis, and H. Bremer, *J. Bacteriol.*, 2000, **182**, 3037-3044.

39. A.H. Roeder, V. Chickarmane, A. Cunha, B. Obara, B.S. Manjunath, and E.M. Meyerowitz, *PLoS Biol*., 2010, **8**, e1000367.

40. A. Zilman, V.V. Ganusov, and A.S. Perelson, *PLoS one*, 2010, **5**, e12775.

41. M. Malumbres, and M. Barbacid, *Nat. Rev. Cancer*, 2009, **9**, 153-166.

42. E.D. Hawkins, J.F. Markham, L.P. McGuinness, and P.D. Hodgkin, *Proc. Natl. Acad. Sci. U.S.A.*, 2009, **106**, 13457-13462.

43. M.L. Turner, E.D. Hawkins, and P.D. Hodgkin, *J. Immunol.*, 2008, **181**, 374-382.

44. E.D. Hawkins, M.L. Turner, M.R. Dowling, C. van Gend, and H.D. Hodgkin, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 5032-5037.

45. S. Krishna, B. Banerjee, T.V. Ramakrishnan, and G.V. Shivashankar, *Proc. Natl. Acad. Sci. U.S.A.*, 2005, **102**, 4771-4776.

46. A. Schwabe, and F.J. Bruggeman, *Biophys. J.*, 2014, **107**, 301-313.

47. A. Sanchez, S. Choubey, and J. Kondev, *Annu. Rev. Biophys.,* 2013, **42**, 469-491.

48. M. Osella, E. Nugent, and M.C. Lagomarsino, *Proc. Natl. Acad. Sci. U.S.A.,* 2014, **111**, 3431-3435.

49. J. Lloyd-Price, H. Tran, and A.S. Ribeiro, *J. Theor. Biol.*, 2014, **356**, 11-19.

50. B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, and M. Selbach, *Nature*, 2011, **473**, 337-342.

51. J.J. Tyson, and B. Novak, *Interface Focus*, 2014, **4**, 20130070.

52. M. Hintsche, and S. Klumpp, *J. Biol. Eng*., 2013, **7**, 22.

53. L.F. Huang, Z.J. Yuan, P.J. Liu, and T.S. Zhou, *Phys. Rev. E*, 2014, **90**, 052702.

54. DT. Maheshri, and E.K. O'Shea, *Annu. Rev. Bioph. Biom.*, 2007, **36**, 413-434.

55. B. Munsky, and M. Khammash, *J. Chem. Phys*., 2006, **124**, 044104.

56. M. Acar, J.T. Mettetal, and A. van Oudenaarden, *Nat. Genet.,* 2008, **40**, 471-475.

57. M. Mann, *Nat. Rev. Mol. Cell. Biol.,* 2006, **7**, 952-958.

58. P. Thomas, N. Popović, and R. Grima, *Proc. Natl. Acad. Sci. U.S.A*., 2014, **111**, 6994-6999.

59. T.L. To, and N. Maheshri, *Science,* 2010, **327**, 1142-1145.

60. H. Qian, P.Z. Shi, and J. Xing, *Phys. Chem. Chem. Phys*., 2009, **11**, 4861-4870.

61. A. Ochab-Marcinek, and M. Tabaka, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 22096-22101.

62. J.J. Zhang, and T.S. Zhou, *Biophys. J.,* 2014, **10**, 479-488.

63. T.S. Zhou, and J.J. Zhang, *SIAM J. Appl. Math.*, 2012, **72**, 789-818.

64. J.J. Zhang, L.N. Chen, and T.S. Zhou, *Biophys. J.,* 2012, **102**, 1247-1257.

65. S. Chong, C. Chen, H. Ge, and X.S. Xie, *Cell*, 2014, **158**, 314-326.

66. J. Jaruszewicz, P.J. Zuk, and T. Lipniacki, *J. Theor. Biol.,* 2013, **317**, 140-151.

67. H. Chen, K. Shiroguchi, H. Ge, and X.S. Xie, *Mol. Syst. Biol.*, 2015, **11(5)**, 808.

68. A.V. Gett, and P.D. Hodgkin, *Nat. Immunol.*, 2000, 1, 239-244.

69. R.J. De Boer, V.V. Ganusov, D. Milutinović, P.D. Hodgkin, and A.S. Perelson, *B. Math. Biol.,* 2006, **68**, 1011-1031.

70. D.T. Gillespie, *J. Chem. Phys.,* 1977, **81**, 2340-2361.

71. R.C. Herman, J.G. Williams, and S. Penman, *Cell*, 1976, **7**, 429-437.

72. J.S. Mattick, *Nat. Rev. Genet*., 2004, **5**, 316-323

73. Q.L. Wang, and T.S. Zhou, *Phys. Rev. E*, 2014, **89**, 012713.

74. M.P. Cosma, T. Tanaka, and K. Nasmyth, *Cell*, 1999, **97**, 299-311.

75. K. Martinowich, D. Hattori, H. Wu, S. Fouse, F. He, Y. Hu, and Y.E. Sun, *Science*, 2003, **302**, 890-893.

76. J. Paulsson, O.G. Berg, and M. Ehrenberg, *Proc. Natl. Acad. Sci. U.S.A*., 2000, **97**, 7148-7153.