Advanced Studies in Biology, Vol. 9, 2017, no. 2, 91 - 100
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/asb.2017.726

Effects of Different Burst Form on Gene Expression Dynamic and First-Passage Time

Xuejie Liu and Qiuying Li

College of Mathematics and Information Science
Guangzhou University, Guangzhou, 510006, P. R. China

Abstract

Timing of events in many cellular processes, such as cell-cycle control, cell differentiation, and so on, depend on regulatory proteins reaching critical threshold levels. The time at which special protein cross up the threshold levels is called as the first-passage time(FPT). Here, we theoretically show that geometric conditional burst has unneglectable effects on gene expression and FPT. More precisely, if the mean of protein is fixed and the same burst size, conditional geometric burst increase the mean and noise of FPT compared with the geometric distribution. If the mean of FPT is fixed and the same burst size, conditional geometric burst increase noise of FPT compared with the geometric distribution.

Keywords: First-passage time; gene expression; burst expression.

1 Introduction

The intrinsic stochastic nature of biochemical reactions, can lead to cell-to-cell variabilities contributing to cell survival across identical cells[9, 12].As a result, the inherent random feature may lead to cell-to-cell variations between identical genes, implying that the timing of a cellular event which triggers at a critical protein level is stochastic in nature. Switching timing of many cellular phenotypes and function, such as sporulation, cell-cycle control, and so on, depend on regulatory proteins reaching critical threshold levels. For example,lysis time for an bacteriophage λ depends on the accumulation of

holin protein in the cell membrane exceeds a threshold [5, 6, 10]. Timing of these protein reaching the critical threshold in single cells is influenced by fluctuations that arise naturally due to noise in gene expression. Therefore, we need to discuss the minimal time that special protein reach a given threshold in the process of gene expression. The minimal time is called as the first-passage time(FPT).

There exist two kinds of experiment views on the gene expression in both prokaryotes and eukaryotes. The classical view of gene expression as a constitutive, Poisson-like accumulation of gene products which is supported by a comprehensive large-scale study[11, 14]. Conversely, several elegant studies showed that specific promoters in bacteria , yeast and human gene(> 8000) express gene products in an episodic process, characterized by pulsatile bursts in transcription[1, 3, 4, 7] . Many studies indicated that the number of protein in a gene expression event may be the following three synthesis forms: one by one, geometric burst, conditional geometric burst[1, 2, 8]. Here, a simple gene burst expression model is investigated to reveal mechanisms of protein dynamic and FPT at any finite time t, quantifying the effects of different burst forms.

2 Gene Expression Models and Analysis

In order to clearly reveal mechanisms of how burst forms affect protein dynamic, we consider a simple gene expression model(see figure1). This paper discusses dynamical stochasticity of the time that cell phenotype switch or cell divide. To obtain exact analytical results, we further assume that the interval

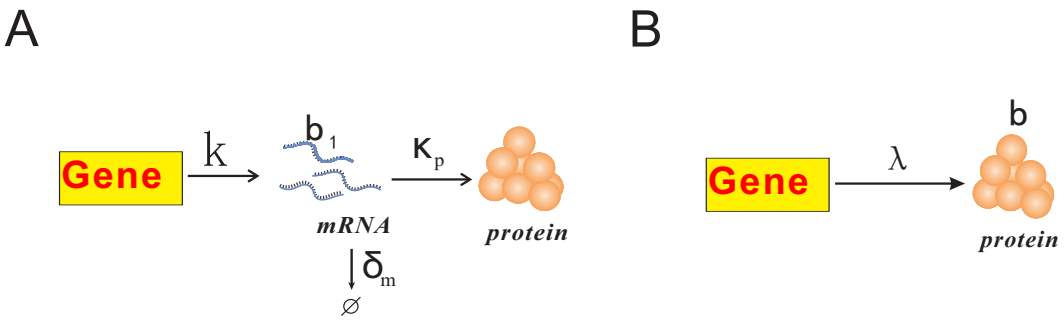

Figure 1: Gene expression models. A: gene expression process B: gene expression model. Here, λ, b is denoted as gene expression frequency and the number of protein from a gene expression event respectively. Protein synthesis is in four different forms:one by one, a fixed number, geometric burst and geometric conditional burst.

time between transcription burst events obey the exponential distribution. In addition, we also assume that time scale of lifetime mRNA is faster than the interval time between transcription burst events and the mRNA degrades instantaneously after producing protein in burst manner.

Thus, the simplified model considers gene expression wherein each protein synthesis event (equivalent to transcription event) occurs at an exponentially distributed time with parameter λ and mean of protein synthesis is denoted as b as seen in of B Figure1.

2.1 For four types of protein synthesis form, probability distribution of protein

The intrinsic randomness of biochemical reactions leads to the protein number being a stochastic process. Further, let P_n be protein count after n gene expression events and $N(t)$ be the random variable, representing the number of protein burst synthesis event in time interval $[0, \quad t]$. Let $P(m,t)$ represent the probability of the event $\{P(t) = m\}$. At the same time, let T_i be the interval time between $(i-1)^{th}$ transcription bursting events and i^{th} transcription bursting events, then $T_i, i = 1, 2, 3, \cdots$ are independent and identically distribution by the assumptions. Further, the probability density function of random variable $\tau_n = \sum_{k=1}^{n} T_i$ follows a Gamma distribution process $\Gamma(n, \lambda t)$. Obviously, $N(t)$ follows a Poisson process $P(\lambda t)$.

Next, we consider production of protein is in the following burst forms. We assume that $\lambda_i, i = 1, 2, 3$ represent the gene expression rate for one by one form, the gene expression rate for geometric burst form, and the gene expression rate for conditional geometric form, respectively.

Protein synthesis in one by one form. If protein synthesis is in one by one form. That is to say that the number Y of protein equally to the times of transcription. So, we have

$$P(m,t) = Pr\{P(t) = m\} = Pr\{N(t) = m\} = e^{-\lambda_1 t}\frac{(\lambda_1 t)^m}{m!}. \tag{1}$$

Protein synthesis in geometric burst form. If the number Y of protein produced in a transcription obey the geometry distribution, $\mu(1-\mu)^k, k = 0, 1, 2, \cdots$, then we can obtain

$$\begin{aligned} P(m,t) &= Pr\{P(t) = m\} = e^{-\lambda_2 t}\sum_{n=0}^{+\infty} Pr\{P_n = m\}\frac{(\lambda_2 t)^n}{n!} \\ &= \begin{cases} e^{-\lambda_2(1-\mu)t}, & m = 0; \\ e^{-\lambda_2 t}\sum_{n=1}^{+\infty}\frac{(\lambda_2 t)^n}{n!}\frac{\Gamma(n+m)}{\Gamma(n)\Gamma(m+1)}\mu^n(1-\mu)^m, & m \neq 0. \end{cases} \end{aligned}$$

Protein synthesis in geometric conditional burst. Consistent with experiments and theoretical analysis, the distribution of the number of protein

molecules produced by a single gene expression event has the following the probability mass function.

$$Q(0) = Pr\{Y = 0\} = \frac{1}{1+b_1} + A. \tag{2}$$

and

$$Q(k) = Pr\{Y = k\} = \frac{b_1}{1+b_1}\frac{(b_1+1)^k b_2^k}{(1+(b_1+1)b_2)^{k+1}} = AB^k, k = 1, 2, 3, \cdots \tag{3}$$

here $A = \frac{b_1}{1+b_1}\frac{1}{1+(b_1+1)b_2}$, $B = \frac{(b_1+1)b_2}{1+(b_1+1)b_2}$. Here, b_1, b_2 represent the mean of burst gene expression. When $m = 0$, we have

$$P(0,t) = Pr\{P(t) = 0\} = e^{-\lambda t}\sum_{n=0}^{+\infty}\frac{1}{(1+b_1)^n}\frac{(\lambda t)^n}{n!}\sum_{j=0}^{n}\frac{\Gamma(n+1)(A+Ab_1)^j}{\Gamma(j+1)\Gamma(n-j+1)}. \tag{4}$$

When $m \neq 0$, we have

$$\begin{aligned} P(m,t) &= Pr\{P(t) = m\} \\ &= e^{-\lambda_4 t}B^m\sum_{n=1}^{+\infty}\frac{z^{n-m}(\lambda_4 t)^n}{m!(1+b_1)^n n!}\frac{\Gamma(-n+m)}{\Gamma(-n)}F(-m, n; 1+n-m; z), \end{aligned} \tag{5}$$

here $z = 1 + A(1 + b_1)$.

3 Main results

3.1 The noise caused by geometric conditional burst more than caused by the other two forms

In previous studies, the effect of geometric conditional burst is frequently neglected. Here, we theoretically show that geometric conditional burst has unneglectable effects on gene expression. More precisely, if the mean of protein is fixed, the noise of protein caused by geometric conditional burst is more than that of the other two. The theoretical results are show in (8),(10) and(12). Next, we discuss the mean and noise of proteins number at any time t. By using the property for conditional expectation, we have

$$E(P(t)) = E[E(P(t)|N(t))] = \sum_{n=0}^{\infty}E[P(t)|N(t) = n]P_r(N(t) = n) \tag{6}$$

Protein synthesis in one by one form. Therefore, in the case that protein synthesis is in the form of one protein from a gene expression event with a fixed frequency λ_1, we can calculate the mean of protein at time t.

$$E_1(P(t)) = \sum_{n=0}^{\infty} E[P(t)|N(t)=n]P_r(N(t)=n) = \lambda_1 t \tag{7}$$

Thus, we can derive the formulae for noise $\phi_1(t)$ of protein at any time t,

$$\phi_1(t) \;=\; \frac{E(P^2(t)) - E^2(P(t))}{E^2(P(t))} = \frac{1}{\lambda_1 t} = \frac{1}{E_1(P(t))} \tag{8}$$

Protein synthesis in geometric burst. In the case that protein synthesis in the form of geometric bursting at a fixed frequency λ_3

$$E_2(P(t)) = \sum_{n=0}^{\infty} e^{-\lambda_3 t}\frac{(\lambda_3 t)^n}{n!}E[P(t)|N(t)=n] = \frac{(1-\mu)\lambda_3 t}{\mu} \tag{9}$$

Further, we can derive the formulae for noise $\phi_3(t)$ of protein at any time t,

$$\phi_2(t) \;= \frac{E(P^2(t))}{E^2(P(t))} - 1 = \frac{b+\frac{1}{\mu}}{E_2(P(t))}, \quad b = \frac{(1-\mu)}{\mu}. \tag{10}$$

Protein synthesis in geometric conditional burst. In the case that protein synthesis is in the form of conditional geometric bursting with a foxed frequency λ_3,

$$E_3(P(t)) = \frac{AB\lambda_3 t}{(1-B)^2} = b_1 b_2 \lambda_3 t, \tag{11}$$

Further, we can derive the formulae for noise $\phi_3(t)$ of protein at any time t,

$$\phi_3(t) \;= \frac{E(P^2(t))}{E^2(P(t))} - 1 = \frac{1+B}{(1-B)E(P(t))} = \frac{1+2(b_1+1)b_2}{E_3(P(t))}. \tag{12}$$

Obviously, if the mean of protein is fixed and the same burst size, conditional geometric burst increase the noise of protein compared with the geometric distribution. When $\frac{\lambda_2}{\lambda_3}(\frac{1}{2b_1}+1) < 1$, we obtain that conditional geometric burst can decrease noise of protein.

3.2 Conditional Geometric burst increase mean and decrease noise of FPT

In previous studies, the effect of geometric conditional burst is frequently neglected. Here, we theoretically show that geometric conditional burst has unneglectable effects on gene expression. More precisely, if the mean of protein

is fixed, geometric conditional burst increase mean and decrease noise of FPT compared with the geometric distribution. If the mean of FPT is fixed, geometric conditional burst increase compared with the geometric distribution. Geometric conditional burst is less than caused by the geometric distribution. The theoretical results are show in (19),(26) and (29).

We define the First-passage time (FPT) at which the protein of interest across the critical thresholds m as F_m impling that F_m, can be written as:

$$F_m = \inf\{t : P(t) \geq m\},$$

for a given threshold m. Obviously the event $\{F_m \leq t\}$ is equivalent to the event $\{p(t) \geq m\}$. Hence, we have

$$Pr\{F_m \leq t\} = Pr\{p(t) \geq m\}.$$

Let N_m be the random variable, representing the minimum number of protein burst event that it takes for the protein count to across the given threshold m, then $N_m = \inf\{n, P_n \geq m\}$. By the definition of the random variables N_m and T_i, we have $F_m = \sum_{i=1}^{N_m} T_i$.

To quantify FPT moments, we adopt a mathematical description. We assume that number Y of protein in a protein bursting obeys the given discrete probability distribution $P_r(Y = k) = f_Y(k)$. Thus, let X_i be the random variable, denoting the number of proteins produced by i^{th} protein bursting. Obviously, X_i are independent, and identically distributed. By the definition of the random variable N_m, the probability for the random event $\{N_m = n\}, n = 1, 2, 3, \cdots$ can be denoted as

$$Pr\{N_m = 1\} = Pr\{X_1 \geq m\} = 1 - Pr\{X_1 < m\} = 1 - \sum_{k=0}^{m-1} f_Y(k) \qquad (13)$$

and

$$Pr\{N_m = n\} \;=\; \sum_{k=0}^{m-1} Pr\{N_{m-k} = n-1\} Pr\{X_1 = k\}, \qquad (14)$$

for $n = 2, 3, 4, \cdots$. Further, we get the recurrence formula for the mean $E(N_m)$ of random variable N_m

$$E(N_m) \;=\; \frac{1}{1 - f_Y(0)} + \frac{1}{1 - f_Y(0)} \sum_{k=1}^{m-1} E(N_{m-k}) f_Y(k) \qquad (15)$$

Particulary, $E(N_1) = 1$ while $f_Y(0) = 0$. It is easy to derived the formulae for $E(N_m)$ by induction for given threshold m.

From the definition of the random variable N_m, we can derive the recurrence formula for $E(N_m^2)$,

$$E(N_m^2) = E(N_1)\sum_{k=1}^{m-1} E(N_{m-k}^2) f_Y(k) + 2E(N_m)E(N_1) - E(N_1) \quad (16)$$

At the same time, we can get

$$E(N_1^2) = E(N_1^2) f_Y(0) + 2E(N_1) - 1 = 2E^2(N_1) - E(N_1), \quad (17)$$

and

$$Var(N_1) = E(N_1)(E(N_1) - 1) \quad (18)$$

It is easy to derived the formulae for $Var(N_m)$ by using induction for the given threshold m. The method of above analysis is independent from the concrete form of the probability distribution for protein synthesis bursting.

Protein synthesis in the one by one form. If protein synthesis is in the one by one form, we can obtain

$$E(F_m) = \frac{m}{\lambda_1}, Var(F_m) = \frac{m}{\lambda_1^2} \quad (19)$$

Protein synthesis in geometric burst form. If number Y of protein burst expression being produced in a transcription event follows the probability mass function that is geometric distribution

$$f_Y(k) = Pr\{Y = k\} = \mu(1-\mu)^k, k = 0, 1, 2, \cdots, \quad (20)$$

with mean $b = \frac{1-\mu}{\mu}$. Substituting (20) into the recurrence equation (15) of mean $E(N_m)$, the formulae of mean $E(N_m)$ can be easily obtained

$$E(N_m) = 1 + \frac{m\mu}{1-\mu} = 1 + \frac{m}{b}, \quad (21)$$

From (17) and (18), we have

$$E(N_1^2) = \frac{1+\mu}{(1-\mu)^2},\ Var(N_1) = \frac{\mu}{(1-\mu)^2}, \quad (22)$$

For $k = 2, 3, 4\cdots$, we can derive the following formulation

$$E(N_m^2) = [m^2\mu^2 - 2m\mu^2 + \mu^2 + 3m\mu - 2\mu + 1]\frac{1}{(1-\mu)^2} \quad (23)$$

Thus, we have

$$Var(N_m) = \frac{m\mu}{1-\mu} + \frac{m\mu^2}{(1-\mu)^2} = \frac{m}{b} + \frac{m}{b^2}. \quad (24)$$

For $i = 1, 2, 3, 4, \cdots$, the corresponding mean interval times and its covariance would be

$$E(T_i) = \frac{1}{\lambda_3}, Var(T_i) = \frac{1}{\lambda_3^2}.$$

Thus, we obtain

$$E(F_m) = \frac{1}{\lambda_3}[1 + \frac{m}{b}] \quad \text{and} \quad Var(F_m) = \frac{1}{\lambda_3^2}[1 + \frac{m}{b}] + \frac{1}{\lambda_3^2}[\frac{m\mu}{1-\mu} + \frac{m}{b^2}]. \quad (25)$$

Thus, the noise intensity of F_m is obtained according to the formula

$$\eta_F = \frac{Var(F_m)}{E^2(F_m)} = \frac{(1+2b)m + b^2}{(b+m)^2}. \quad (26)$$

Protein synthesis in geometric conditional burst form. If number Y of protein burst expression being produced in a transcription event follows conditional geometric distribution, the formulae of mean $E(N_m)$ can be easily obtained

$$E(N_m) = \frac{m}{b_1 b_2} + \frac{1}{b_1} + 1. \quad (27)$$

Similarly to the aforementioned discussion about (24)

$$Var(N_m) = (2a^2 b_2 + a^2 + a)m + ab_2(ab_2 + 1). \quad (28)$$

For $i = 1, 2, 3, 4, \cdots$, the corresponding mean interval times and its covariance would be

$$E(T_i) = \frac{1}{\lambda_3}, Var(T_i) = \frac{1}{\lambda_3^2}.$$

Thus, we obtain

$$E(F_m) = \frac{1}{\lambda_3}[ma + ab_2 + 1],$$

and

$$Var(F_m) = \frac{1}{\lambda_3^2}[(2a^2 b_2 + a^2 + 2a)m + ab_2(ab_2 + 2) + 1].$$

Thus, the noise intensity of F_m is obtained according to the formula

$$\eta_F = \frac{Var(F_m)}{E^2(F_m)} = \frac{\frac{m}{b_1^2 b_2^2} + \frac{2m}{b_1^2 b_2} + \frac{2m}{b_1 b_2} + \frac{1}{b_1^2} + \frac{2}{b_1} + 1}{(\frac{m}{b_1 b_2} + \frac{1}{b_1} + 1)^2}. \quad (29)$$

4 Conclusion

We theoretically show that geometric conditional burst has unneglectable effects on gene expression and FPT. If the mean of protein is fixed and the same burst size, the conditional geometric burst increase the mean and noise of FPT compared with the geometric distribution. When $\frac{\lambda_2}{\lambda_3}(\frac{1}{2b_1} + 1) < 1$, we obtain that conditional geometric burst can decrease noise of protein and FPT.

References

[1] L. Cai, N. Friedman and X.S. Xie, Stochastic protein expression in individual cells at the single molecule level, *Nature,* **440** (2006), 358-362. https://doi.org/10.1038/nature04599

[2] S. Chong, C.Y. Chen, H. Ge and X.S. Xie, Mechanism of transcriptional bursting in bacteria, *Cell,* **158** (2014), 314-326. https://doi.org/10.1016/j.cell.2014.05.038

[3] J.R. Chubb and T.B. Liverpool, Bursts and pulses: insights from single cell studies into transcriptional mechanisms, *Current Opinion in Genetics and Development,* **20** (2010), 478-484. https://doi.org/10.1016/j.gde.2010.06.009

[4] V. Elgart, T. Jia, A.T. Fenley and R. Kulkarni, Connecting protein and mRNA burst distributions for stochastic models of gene expression, *Physical Biology,* **8** (2011), 046001. https://doi.org/10.1088/1478-3975/8/4/046001

[5] A. Gründling, M.D. Manson and R. Young, Holins kill without warning, *Proc. Natl. Acad. Sci. USA,* **98** (2001), 9348-9352. https://doi.org/10.1073/pnas.151247598

[6] J.D. John, L.N. Wang, Factors influencing lysis time stochasticity in bacteriophage λ, *BMC Microbiology,* **11** (2011), 174. https://doi.org/10.1186/1471-2180-11-174

[7] T. Morisaki, K. Lyon and K. F. DeLuca, et al., Real-time quantification of single RNA translation dynamics in living cells, *Science,* **352** (2016), 1425-1429. https://doi.org/10.1126/science.aaf0899

[8] J.M. Pedraza, J. Paulsson, Effects of Molecular Memory and Bursting on Fluctuations in Gene Expression, *Science,* **319** (2008), 339-343. https://doi.org/10.1126/science.1144331

[9] A. Sanchez, S. Choubey and J. Kondev, Regulation of Noise in Gene Expression, *Annual Rev. Biophys.,* **42** (2013), 469-491. https://doi.org/10.1146/annurev-biophys-083012-130401

[10] A. Singh and J.J. Dennehy, Stochastic holin expression can account for lysis time variation in the bacteriophage λ, *Journal of The Royal Society Interface,* **11** (2014), 0140. https://doi.org/10.1098/rsif.2014.0140

[11] Y. Taniguchi, P. J. Choi, G.-W. Li, et al., Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells, *Science*, **329** (2010), 533-538. https://doi.org/10.1126/science.1188308

[12] M. Thattai and A.V. Oudenaarden, Intrinsic noise in gene regulatory networks, *Proc. Acad. Sci. USA.,* **98** (2001), 8614-8619. https://doi.org/10.1073/pnas.151588598

[13] B. Wu, C. Eliscovich, Y. J. Yoon, R. H. Singer, Translation dynamics of single mRNAs in live cells and neurons, *Science,* **352** (2016), 1430-1435. https://doi.org/10.1126/science.aaf1084

[14] S. Yunger, L. Rosenfeld, Y. Garini and Y. Shav-Tal, Single-allele analysis of transcription kinetics in living mammalian cells, *Nature Methods*, **7** (2010), 631-633. https://doi.org/10.1038/nmeth.1482

Received: March 12, 2017; Published: March 31, 2017