Extrinsic vs. Intrinsic Noises in Phage Lambda Genetic Switch

Wei Tian†, Hongyuan Zhu†, Xue Lei and Ping Ao‡
Shanghai Center for Systems Biomedicine, Key Laboratory of Systems Biomedicine of Ministry of Education,
Shanghai Jiao Tong University, 200240, Shanghai, P.R. China
† Co first authors; contribute equally to this work.
‡ To whom correspondence should be addressed. Email: aoping@sjtu.edu.cn

***Abstract*—Noises in biological modeling may be classified into two kinds: intrinsic noise, which derives from the variability in dominant molecular interaction and is responsible for the given phenomenon, and extrinsic noise, which arises from other sources, like fluctuations in the environment and so on. Phage lambda is a simple model organism that exhibits important noisy characteristics. It lives in either lysogenic state or lytic state after infecting a bacterium, that is determined by a genetic switch. The mathematical modeling of this genetic switch typically only considers intrinsic noise, though a previous study by one of present authors suggested the critical role of extrinsic noise. In the present study by comparing theoretical results of phage lambda in lysogeny with experiment data, we first achieve good numerical agreements of five constrains of phage lambda for averaged variables. This success indicates that current dominant molecular agents are right. In addition, we confirm the existence of extrinsic noise in lambda genetic switch and find it surprisingly large. This finding calls for an extension of the current mathematical model to better describe the noises. We also point out some possible sources of extrinsic noise.**

***Index Terms*—phage lambda, stochasticity, intrinsic noise, extrinsic noise, lysogeny**

I. Background

Typically, behaviors of organisms are shaped by their genotype and environment. However, these behaviors sometimes cannot be precisely predicted when examined at a single individual level, even when examinations are conducted with no genotype difference and subjected to the identical conditions[1][2][3]. A well-known example that exhibits "noisy" character is phage lambda which has two different states of growth in its Escherichia coli host. In the lytic state, phage lambda uses molecular-genetic apparatus for large-scale production of progeny phage resulting the lysis of the host. In the lysogenic state, phage DNA is integrated into the host genome and passed to each daughter cell passively as the host divides. The host is maintained in this dormant state by a single protein CI, which represses gene expression in lytic state[4].

In the past five decades, a large amount of efforts have been put on phage lambda, not only because phage lambda is one of the simplest living things on the earth, but also because such bi-stable switch, which performs functions of state transition or of state locking, is one of the basic building blocks to form living organisms. For instance, the similar switch structures have been found in developmental transcription network[5]. In an informatics view, the decision mechanism between two states provides the basic function of biological computation. A deep understanding of the decision making of phage lambda will give us a standpoint for studying other more sophisticated biological organisms and promote our insights into systems biology.

Some efforts to explain lambda genetic switch are made through a deterministic view. The rationale is straightforward: heterogeneity of cell fate may be deterministically determined by some variables different from cell to cell, such as cell volume, which is not under consideration or at least not detected before. These attempts show that certain variables account for cell-fate heterogeneity to some extent. However, the careful quantification of cell-to-cell difference, failing to find a certain threshold to flip the genetic switch, still cannot explain experimental data in a deterministic way, albeit discovering some relations between heterogeneity and individual difference of cells[6].

On the other hand, a more traditional approach can describe the heterogeneity of fate of cells to inherent stochasticity of chemical reactions, which bases solidly on physical-chemical nature of the life process. Dramatic progress has been made in the past decades from such a point of view. Shea and Ackers formulated an elegant physical-chemical model which provided a base for later study[7]. Arkin et al. performed simulation and pointed out that the random cell-fate decisions result from the inevitable fluctuations of the chemical reactions in gene expression[1]. Zhu et al. analyzed stochastic dynamics of phage lambda system, where a successful numerical agreement between experiments and model was achieved, as well as pointing out that potential landscape is an insightful tool in understanding biological systems[8].

Although stochasticity plays an important role in lambda genetic switch, still little is known about noises in such system. Noises in gene expression can be divided into two kinds: intrinsic noise and extrinsic noise. Intrinsic noise is usually termed as noise that derives from natural variability in the way that molecules interact in biochemical reactions, whereas extrinsic noise is usually termed as the fluctuation outside the system (like cell growth[9][10]), or the effects of other cellular regulatory inputs[11]. The above definitions are of many liberties, resulting in that the two terms have different forms and different mathematic definitions in works of different research groups. Here, to make the meanings

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

terminologically clear, we will use the operational definitions of intrinsic noise and extrinsic noise: Intrinsic noise is defined as the fluctuation that derives from the factors we considered in the model; any noise beyond that is included into extrinsic noise. Since we considered only biochemical reactions in the system, such definition is consistent with the biological meanings of the usual definitions mentioned above.

An accurate way to capture the intrinsic noise in such system is using Chemical Master Equation (CME), which describes exact the evolution of chemical processes of a system[12]. CME model has successfully explained some phenomena of phage lambda[13]. However, there is no obvious way to include extrinsic noise into a CME model, while the counterpart Chemical Langevin Equation (CLE) shows much larger flexibility to depict noise, for CLE is a good enough approximation to CME in the situation we concerned[14]. So as to capture both intrinsic noise and extrinsic noise, we used CLE to model. In this work, we compared the theoretical steady distributions of CI in lysogeny with the distributions of five strains of phage lambda measured in experiments[15], achieved a good agreements of the mean values and confirmd the existence of extrinsic noise. We found that strength of extrinsic noise is over 10 times as large as that of intrinsic noise of the system. This result suggests a need for refining of current model, and calls for more both experimental and theoretical efforts to study exact sources of extrinsic noise and their impacts.

II. Results

In a host in lysogeny, regulatory protein CI plays a crucial role in maintaining the dormant state[4]. CI stabilizes the lysogenic state by binding to the operators O_R and O_L, repressing the promoter that leads to lytic growth (Fig 1). O_R and O_L each contains three binding sites where CI dimers can be bound to with different affinities. As for O_R, the affinity order is $O_R1 > O_R2 = O_R3$. When CI dimers occupy O_R1, which partially overlap with the lytic promoter P_R, RNA polymerase (RNAp) is prevented from initiating transcription for the lytic gene. Cooperativity between adjacent CI dimers increases affinity of CI dimers to O_R2, so that the second site is readily occupied. When CI dimers are bound to O_R2, which is adjacent to the own promoter of CI, cooperativity between CI dimers and RNAp activates CI transcription. CI molecules binding to O_R and O_L can polymerise and form a long range DNA loop and the looping effect activates CI transcription to different degrees according to different binding configurations of O_R and O_L[15].

Two chemical reaction channels we concerned are the synthesis and decay of CI:

$$\begin{aligned} \mathrm{DNA} &\rightarrow \mathrm{CI} \quad \mathrm{R1} \\ \mathrm{CI} &\rightarrow \phi \quad \mathrm{R2} \end{aligned}$$

In reaction R1, we abbreviate the process of transcription from DNA to mRNA and of translation from mRNA to protein, regarding the synthesis of CI as a one-step reaction from DNA to proteins directly. The simplification was first

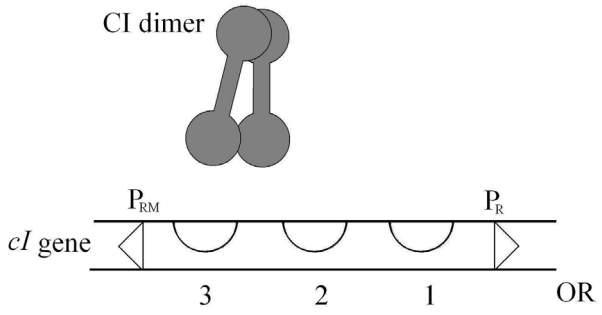

Fig. 1. Operator O_R. Three binding sites for CI dimers comprise the Operator O_R. CI dimers bind to the two highest affinity sites cooperatively and repress P_R. Production of cI is simultaneously activated when CI dimers bound to O_R2

used by Aurell and Sneppen[16], and similar simplification is widely adopted when modeling other gene circuits. The reason behind this simplification is that the time scale associated with the dynamics of proteins is usually much longer than one associated with the dynamics of mRNA. Thus, a quasi steady state assumption can be made, so mRNA is no longer relevant except for including an overall constant representing the rate of gene expression. As for reaction R2, the concentration of CI decreases with cell division, since protein CI does not degrade actively in lysogeny.

The system of reaction R1 and R2 can be described by CLE

$$\frac{dN_{\mathrm{CI}}(t)}{dt} = f_{\mathrm{CI}}(N_{\mathrm{CI}}) - N_{\mathrm{CI}}/\tau_{\mathrm{CI}} + \zeta(N_{\mathrm{CI}}, t) \tag{1}$$

Here N_{CI} is the protein number for CI. The rate of CI production f_{CI} is a function of CI number in the host. The decay constant τ_{CI} is an effective lifetime proportional to the bacterial lifetime, given in Table III. The random term $\zeta(N_{\mathrm{CI}}, t)$ describes the stochastic effects in the system, consisting of intrinsic noise and extrinsic noise, which we assume are two Gaussian white noises with variance D_{in} and D_{ex} separately.

$$\zeta(N_{\mathrm{CI}}, t) = \zeta_{\mathrm{in}}(N_{\mathrm{CI}}, t) + \zeta_{\mathrm{ex}}(N_{\mathrm{CI}}, t). \tag{2}$$

D_{in} and D_{ex} represent the strength of intrinsic and extrinsic noises. As we assumed intrinsic noise came from production and decay rates of the biochemical reactions[17], the stochastic nature of the system, D_{in} can be derived from the reaction rate constants of the chemical reactions[14] based on the thermodynamic rule. Using a powerful method proposed by Ao[18], we can get the potential function of the system

$$\Phi(N_{\mathrm{CI}}) = \int_0^{N_{\mathrm{CI}}} \frac{f_{\mathrm{CI}}(N'_{\mathrm{CI}})}{D_{\mathrm{in}} + D_{\mathrm{ex}}} dN'_{\mathrm{CI}} \tag{3}$$

Then a steady distribution of CI can be expected from the potential using Gibbs measure

$$P(N_{\mathrm{CI}}) \propto e^{-\Phi(N_{\mathrm{CI}})} \tag{4}$$

First, we assume that total stochasticity in the system is contributed by intrinsic noise, i.e. $\zeta_{\mathrm{ex}}(N_{\mathrm{CI}}, t) \equiv 0$. The

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

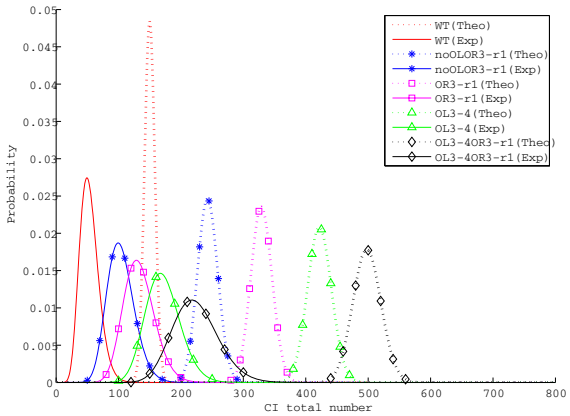

Fig. 2. The solid curves are experiment data, from [15], and dashed curves are theoretical results. Plotted in the same figure, the units of abscissa of experiment data and of theoretical results are different. The unit for experiment data (dashed curves)is the fluorescence intensity, and theoretical results (solid curves) the total numbers of CI proteins.

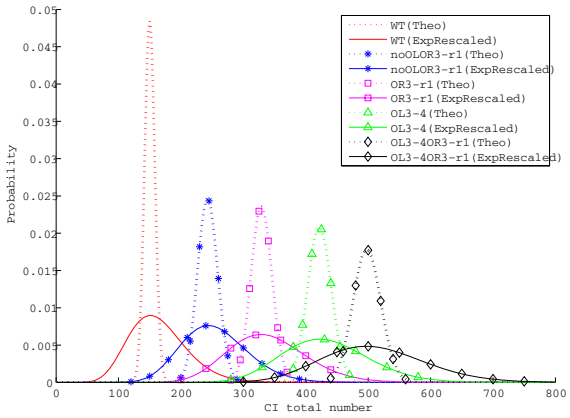

Fig. 3. The experiment data [15] after rescaling. The rescaling factors for the five constrains are 3.06, 2.45, 2.55, 2.53 and 2.28 in the order from left to right. All these are around 2.57.

theoretical steady distribution of CI of five strains can be got from Eq. 1 (Fig. 2). To make the experimental results of the work of Anerson and Yang[15] (cf. Fig. 2) comparable with the theoretical results, the fluorescence intensity needs to be converted to the number of CI. We assumed that fluorescence intensity is proportional to the number of CI, and rescaled experimental distributions to make each of them has the same mode (the value occurs most frequently) with corresponding theoretical distribution. The recaling factors of four mutants are around 2.45, while the one of wild type is sort of deviated, which we will discuss later. Fig. 3 shows that the experimental results contain more variation than the results we calculated, suggesting extrinsic noise is needed to explain the experiments beyond intrinsic noise.

Then we introduced extrinsic noise in two ways: the extrinsic noise is independent from and dependent on the intrinsic noise. Adopting simple approaches, we assume the extrinsic noise is a constant

$$D_{\mathrm{ex}} = c \tag{5}$$

when independent from the intrinsic noise, and is proportional to the intrinsic noise

$$D_{\mathrm{ex}} = qD_{\mathrm{in}} \tag{6}$$

when dependent on the intrinsic noise.

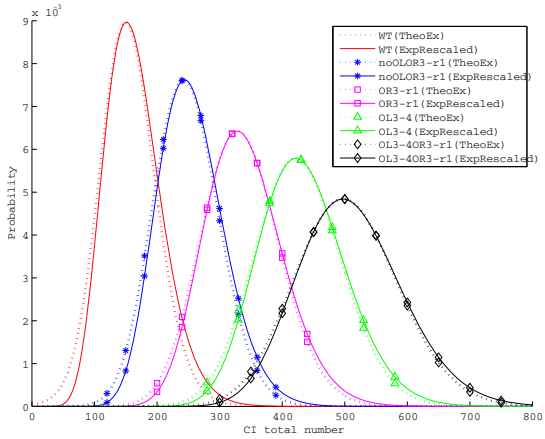

Fig. 4. The solid curves are the rescaled experiment data, while the dashed curves are the theoretical results with extrinsic noise considered. The extrinsic noise is of Eq. 6 form. The figure for extrinsic noise of Eq. 5 is similar.

constrain	q	$c/D_{\mathrm{in}}(\mathrm{mode})$
wt	27.58	27.92
noolor3r1	10.41	9.34
or3r1	13.40	12.47
ol34	12.89	11.91
ol34or3r1	13.56	12.57

Table I

THE SECOND AND THE THIRD COLUMNS SHOW RELATIVE STRENGTH OF EXTRINSIC NOISE TO INTRINSIC NOISE. $D_{\mathrm{in}}(\mathrm{mode})$ DENOTES THE INTRINSIC NOISE STRENGTH WHEN THE TOTAL NUMBER OF CI PROTEINS IN THE SYSTEM IS EQUAL TO THE CORRESPONDING MODE VALUES.

Certain extrinsic noise added, theoretical results agrees quite well with experiments (cf. Fig. 4). Surprisingly, we found that large extrinsic noise (10 times or more as large as the intrinsic noise, cf. Table I) is needed to explain the experiments, and different constrains exhibit different strength of extrinsic noise. The noOL,OR3-r1 mutant, DNA of which cannot form a loop due to lost of left operator, has the smallest relative extrinsic noise, which may suggest looping effect being a source of extrinsic noise. The large relative extrinsic noise and the large rescaling factor of wild type may be ascribed to some effects unknown so far. It is intriguing to find that noise in wild type is relatively much larger than that in mutants, which may examplify the tolerance of biological organisms to noise, given that wild type phage lambda has much more fitness than the mutants after millions of years of evolution.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

III. Discussion

After decades of accumulation of knowledge on phage lambda, there remains questions around this genetic switch. One of these questions is the existence and impacts of noise onto the bi-stable switch. The theoretical results have already agreed well with experiment data in a sense of average, hence a deterministic description would be good. Nevertheless, in the present work we found a large deficit of variation in theoretical results in comparison with experimental data, which indicates the existence of extrinsic noise. This finding suggests some shortcomings of current mathematic model. The current model, which is simple but insightful, is far from enough when we turn to the variation, or the noise issue, albeit being good enough when we focus on some mean values of the system. One reason for this deficiency may be that we did not fully realize the important role noise plays in organism before. Fortunately, the idea that noise is important within biological systems has been generally accepted, ever since it was realized not only is it an inevitable by-product of any inherent biochemical processes, but also noise is in fact essential for development and evolution. Some simple game analysis demonstrates that noise help phage lambda to survive in the volatile environment[19].

Another reason of the deficiency of the model may be that the model is oversimplified. In the widely used model[16], proteins are considered to be synthesized from DNA in one step(cf. R1) for transcription to and translation from mRNA is omitted. The simplification is based on the fact that dynamics of mRNA is associated with a shorter time scale compared with dynamics of proteins, so that a quasi steady state assumption, in a sense, can be made without affecting the mean values of the system. On the other hand, the copy number of mRNA is rather small, which suggests that mRNA could be an important noise source of the switch system. We have some evidences(unpublished) that when transcription to and translation from RNA is considered, the variation loss, or the extrinsic noise, is much smaller than we currently found in this work. Beyond that, CI production per translation and cell growth of E. coli might be also sources of extrinsic noise. We believe more about extrinsic noise will be known when we have accumulated more experiment data and have further refined model, and the precise role noise plays in biological phenomena will be understood more quantitatively.

IV. Materials and methods

The system of reaction R1 and R2 can be described by a CLE

$$\frac{dN_{\rm CI}(t)}{dt} = f_{\rm CI}(N_{\rm CI}) - N_{\rm CI}/\tau_{\rm CI} + \zeta(N_{\rm CI}, t) \tag{7}$$

Here $N_{\rm CI}$ is the protein number for CI. The decay constant $\tau_{\rm CI}$ is an effective lifetime proportional to the bacterial lifetime. The rate of CI production $f_{\rm CI}$ is a function of CI number in the host. The CI production does not depend on another regulatory protein Cro, which competes with CI for binding sites O_R and O_L, because Cro transcription in a lysogenic bacterium is transient and rare[4]. The rate of transcription of CI when activated by CI dimers bound to O_R2 is denoted T_{RM}, and when not activated T^U_{RM}. There are three levels of activation $A1$, $A2$ and $A3$ according to different configurations of O_R and O_L[15]. The probability of each configuration can be calculated from CI using the canonical approach of Shea and Ackers[7]

$$P(i|[CI]_{tot}) \propto e^{-\Delta G_i/RT}[CI_2]^{n_i} \tag{8}$$

where ΔG_i (Table II) is the Gibbs free energy of the ith configuration, R (Table III) is the gas constant, T (Table III) is the absolute temperature, $[CI_2]$ is the concentration of CI dimers given the total concentration of CI $[CI]_{tot}$. Summarily, the rate of production of CI is

$$\begin{aligned} f_{\rm CI}(N_{\rm CI}) =& ET_{RM}(A_1 \sum_{i\in A1} P_i + A_2 \sum_{i\in A3} P_i \\ &+ A_3 \sum_{i\in A3} P_i) + ET^U_{RM} \sum_{i\notin A1,A2,A3} P_i \end{aligned} \tag{9}$$

where E (Table III) denotes the number of CI molecules produced per transcript.

Our intrinsic noise strength, same as diffusion coefficient in physics, is obtained based on thermodynamics[14]. The intrinsic noise is like,

$$\zeta_{in}(N_{\rm CI}, t) = \sqrt{f_{\rm CI}(N_{\rm CI})}\Gamma_1 + \sqrt{N_{\rm CI}/\tau_{\rm CI}}\Gamma_2 \tag{10}$$

Γ_1 and Γ_2 are Gaussian white noise caused by CI synthesis and degradation fluctuations respectively, which is defined as

$$\Gamma_j = \lim_{dt\to 0} N(0, \frac{1}{dt}) \tag{11}$$

And its correlation function is

$$< \Gamma_j(t)\Gamma'_j(t') >= \delta(j, j')\delta(t - t') \tag{12}$$

The first delta function is Kronecher's and the second is Dirac's. Then definite form of the intrinsic noise's covariance in mathematic form is,

$$D_{\rm in} = f_{\rm CI}(N_{\rm CI}) + N_{\rm CI}/\tau_{\rm CI} \tag{13}$$

Parameters

Parameter	ΔG(kcal/mol)	Parameter	ΔG(kcal/mol)
OR1	-12.5	OR3R1	-6.6
OR2	-10.5	OL1	-13.0
OR3	-9.5	OL2	-11.2
OR12coop	-2.7	OL3	-12.0
OR23coop	-2.9	OL3-4	-4.1
OL12coop	-2.7	ΔGoct	-0.5
OL23coop	-2.0	ΔGtet	-3.0

Table II
GIBBS FREE ENERGY USED IN OUR MODEL. THE VALUES ARE AFTER [15]. ORI(OLI) IS THE BINDING ENERGY FROM CI DIMER TO THE ITH OPERATOR SITE OF OR(OL) AND ORIJCOOP(OLIJCOOP) IS COOPERATIVITY ENERGY WHEN CI DIMER BINDS TO BOTH ORI(OLI) AND ORJ(OLJ). (HERE I,J=1,2,3) OR3R1 AND OL3-4 ARE FOR ONE MUTATED IN OR3 AND OL3. ΔGOCT AND ΔGTET ARE FREE ENERGY FOR OCTAMER AND TETRAMER FORMED IN DNA LOOPING.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE

Parameter	value
v	2e-15 L
R	1.985877e-3 kcal/K/mol
T	273.16+37 K
MOL	6.02e23 /mol
T_{RM}	0.115 /s
T^U_{RM}	0.0105 /s
τ_{CI}	2251s
E	1

Table III
PARAMETERS IN OUR CLE. THE VALUES ARE AFTER [8]. V IS EFFECTIVE CELL VOLUME OF BACTERIA AND MOL IS THE AVOGADRO CONSTANT. R IS GAS CONSTANT AND T IS TEMPERATURE IN KELVIN. T_{RM} AND T^U_{RM} ARE TRANSCRIPTION RATE WHEN OR2 IS OCCUPIED AND NOT. E IS PROTEIN NUMBER PER TRANSCRIPTION. τ_{CI} IS PROTEIN LIFETIME.

ACKNOWLEDGMENT

The authors would like to thank Youfang Cao for the insightful discussion on phage lambda in the weekly meeting, and Tianqi Chen for the discussion on the possible source of extrinsic noise. This work was supported in part by the National 973 Projects No. 2007CB914700 and No. 2010CB529200 (P.A.).

REFERENCES

[1] A. Arkin, J. Ross, and H. McAdams, "Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected escherichia coli cells," *Genetics*, vol. 149, no. 4, p. 1633, 1998.

[2] W. Blake, M. Kærn, C. Cantor, and J. Collins, "Noise in eukaryotic gene expression," *Nature*, vol. 422, no. 6932, pp. 633–637, 2003.

[3] H. Maamar, A. Raj, and D. Dubnau, "Noise in gene expression determines cell fate in bacillus subtilis," *Science*, vol. 317, no. 5837, p. 526, 2007.

[4] M. Ptashne, *A genetic switch: phage lambda revisited*. Cold Spring Harbor Laboratory Pr, 2004.

[5] E. Davidson, J. Rast, P. Oliveri, A. Ransick, C. Calestani, C. Yuh, T. Minokawa, G. Amore, V. Hinman, C. Arenas-Mena, O. Otim, C. Brown, C. Livi, P. Lee, R. Revilla, A. Rust, Z. Pan, M. Schilstra, P. Clarke, M. Arnone, L. Rowen, R. Cameron, D. McClay, L. Hood, and H. Bolouri, "A genomic regulatory network for development," *Science*, vol. 295, p. 1669, 2002.

[6] F. St-Pierre and D. Endy, "Determination of cell fate selection during phage lambda infection," *Proceedings of the National Academy of Sciences*, vol. 105, no. 52, p. 20705, 2008.

[7] M. Shea and G. Ackers, "The or control system of bacteriophage lambda* 1:: A physical-chemical model for gene regulation," *Journal of molecular biology*, vol. 181, no. 2, pp. 211–230, 1985.

[8] X. Zhu, L. Yin, L. Hood, and P. Ao, "Robustness, stability and efficiency of phage lambda genetic switch: dynamical structure analysis," *J Bioinform Comput Biol*, vol. 2, no. 4, pp. 785–817, 2004.

[9] D. Volfson, J. Marciniak, W. Blake, N. Ostroff, L. Tsimring, and J. Hasty, "Origins of extrinsic variability in eukaryotic gene expression," *Nature*, vol. 439, no. 7078, pp. 861–864, 2005.

[10] S. Tsuru, J. Ichinose, A. Kashiwagi, B. Ying, K. Kaneko, and T. Yomo, "Noisy cell growth rate leads to fluctuating protein concentration in bacteria," *Physical Biology*, vol. 6(3), p. 036015, 2009.

[11] M. Elowitz, A. Levine, E. Siggia, and P. Swain, "Stochastic gene expression in a single cell," *Science*, vol. 297, no. 5584, p. 1183, 2002.

[12] J. Liang and H. Qian, "Computational cellular dynamics based on the chemical master equation: a challenge for understanding complexity," *Journal of computer science and technology*, vol. 25, p. 154, 2010.

[13] Y. Cao, H. Lu, and J. Liang, "Probability landscape of heritable and robust epigenetic state of lysogeny in phage lambda," *Proceedings of the National Academy of Sciences*, vol. 107, p. 18445, 2010.

[14] D. Gillespie, "The chemical langevin equation," *The Journal of Chemical Physics*, vol. 113, p. 297, 2000.

[15] L. Anderson and H. Yang, "Dna looping can enhance lysogenic ci transcription in phage lambda," *Proceedings of the National Academy of Sciences*, vol. 105, no. 15, p. 5827, 2008.

[16] E. Aurell and K. Sneppen, "Epigenetics as a first exit problem," *Physical review letters*, vol. 88, no. 4, p. 48101, 2002.

[17] X. Zhu, L. Yin, L. Hood, D. Galas, and P. Ao, "Efficiency, robustness, and stochasticity of gene regulatory networks in systems biology: λ switch as a working example," *Introduction to Systems Biology*, pp. 336–371, 2007.

[18] P. Ao, "Potential in stochastic differential equations: novel construction," *Journal of Physics A: Mathematical and General*, vol. 37, p. L25, 2004.

[19] M. Avlund, I. Dodd, S. Semsey, K. Sneppen, and S. Krishna, "Why do phage play dice?" *Journal of virology*, vol. 83, no. 22, p. 11416, 2009.

978-1-4577-1666-9/11/$26.00 ©2011 IEEE