REVIEW / SYNTHÈSE

What all the noise is about: the physical basis of cellular individuality

Daniel A. Charlebois and Mads Kærn

Abstract: Noise has been traditionally viewed as undesirable in biology, resulting in disorder, distortion, and disruption, and ultimately as something that needs to be filtered and removed. More recently, it has been shown that noise can also be beneficial. We briefly review historical developments pertaining to noise in biological physics, and some of the current research in the field of molecular and cellular biophysics.

PACS Nos: 87.18.Tt, 05.40.Jc, 87.10.Mn, 87.10.Rt, 87.16.Yc, 87.23.Kg

Résumé : Traditionnellement, le bruit a été vu comme étant indésirable en biologie, résultant en désordre, distorsion et perturbation, et ultimement comme quelque chose qui doit être filtré et éliminé. Dernièrement, la recherche a démontré que le bruit peut être utile. Dans ce travail nous passons brièvement en revue les développements historiques touchant le bruit dans le domaine de la physique biologique et la recherche actuelle en biophysique moléculaire et cellulaire.

A historical perspective

In 1827 a botanist named Robert Brown observed pollen particles moving about randomly in a fluid [1, 2]. He observed this with all sorts of nonliving materials including minerals, woods, and century-old dried out plants, and concluded that the erratic movements were not a property of living organisms.

Many years later, physicist Georges Gouy [3] conceived that the motion observed by Brown was a result of the irregular thermal fluctuations of the molecules in the liquid. This phenomenon is now called Brownian motion (BM). It was a young Albert Einstein who worked out the now famous result describing the mean-square displacement (in one dimension) of a particle undergoing BM

$$\langle x^2 \rangle = 2Dt \tag{1}$$

where t is the time and D is the diffusion constant. The key idea behind Einstein's equation is that the random motion of a large particle occurs because it is being constantly bombarded by other "invisible" smaller particles in the fluid. Ultimately, Einstein used the concept of such "noise" to predict the existence of atoms [4–7]. Paul Langevin arrived at the same result for the mean-square displacement a few years later using a different approach, namely, a differential equation with a random force term, or as it is now known in statistical physics, a Langevin equation [8].

In the previous framework by Einstein, the position of a Brownian particle undergoing BM is nowhere differentiable and its instantaneous velocity is correspondingly undefined [9]. To avoid this, Ornstein and Uhlenbeck described the velocity of a Brownian particle, instead of the position, as the main random quantity using a Langevin equation [10]

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = \frac{1}{\tau}[\mu - x(t)] + c^{1/2}\xi_t \tag{2}$$

where μ is the mean, c the diffusion constant, τ the relaxation time, and ξ_t describes a Gaussian white noise process with zero mean and fixed variance.

In classical biology, genetically identical cells in an identical environment are expected to have identical phenotypes (i.e., observable chemical and physical properties). Any observed difference is attributed to experimental error. However, in 1945 a biophysicist named Max Delbrück found that the number of virus particles released from infected bacteria showed reproducible variations and, accordingly, is best described by a probability distribution rather than a single value [11]. These experiments were inspired by earlier theoretical work where Delbrück wrote down a master equation (ME; see later in text) describing the statistical

Received 10 May 2012. Accepted 25 June 2012. Published at www.nrcresearchpress.com/cjp on 21 August 2012.

D.A. Charlebois. Ottawa Institute of Systems Biology, University of Ottawa, 451 Smyth Road, Ottawa, ON K1H 8M5, Canada; Department of Physics, University of Ottawa, 150 Louis Pasteur, Ottawa, ON K1N 6N5, Canada.
M. Kærn. Ottawa Institute of Systems Biology, University of Ottawa, 451 Smyth Road, Ottawa, ON K1H 8M5, Canada; Department of Cellular and Molecular Medicine, University of Ottawa, 451 Smyth Road, Ottawa, ON K1H 8M5, Canada; Department of Physics, University of Ottawa, 150 Louis Pasteur, Ottawa, ON K1N 6N5, Canada.

Corresponding author: Daniel Charlebois (e-mail: daniel.charlebois@uottawa.ca).

fluctuations in the number of particles for an autocatalytic chemical reaction [12]. When Delbrück solved these equations, he obtained the well-known noise scaling relationship

$$\eta = \frac{1}{\sqrt{N}} \tag{3}$$

in which the magnitude of the fluctuations or noise, η, is equal to the reciprocal of the square root of the number of particles, N, that initiate the reaction. Correspondingly, Delbrück hypothesized that the variation he observed arose from the variation in the number of the initial infection. About four decades later, Spudich and Koshland demonstrated that bacterial cells grown in homogeneous conditions showed characteristic behavioural differences that persisted over their lifespans [13]. They attributed this nongenetic individuality to poissonian fluctuations in the small numbers of generator molecules, and suggested that it may also apply to other processes, such as differentiation and asynchrony of cell cultures.

Noise and biological systems

It is not altogether surprising that variation exists in biological systems when one considers the random (stochastic) nature of biochemical reactions. These reactions are stochastic as they result from collisions between Brownian particles, which lead to the nondeterministic timing of individual reactions and an inherently noisy time evolution of molecular population levels [14, 15]. The relative amplitudes of these fluctuations are effectively averaged out of systems, such as test tubes, with a large number of molecules (see (3)). These systems are appropriately described using deterministic equations.

A noisy system can formally be described using the so-called ME approach [16]. A ME is a set of first-order differential equations governing the time evolution of the probability of a system to occupy each one of a discrete set of states. A ME usually takes the form

$$\frac{\mathrm{d}p_k(t)}{\mathrm{d}t} = \sum_{k'} \{W_{k'\to k} p_{k'}(t) - W_{k\to k'} p_k(t)\} \tag{4}$$

Here, p_k is the time-dependent probability associated with state and $W_{k\to k'}$ is the transitional probability per unit time from k' to k. In this form, it is clear that the ME is a gain–loss equation for the probabilities of the separate states, k. The gain of state k due to the transitions from other states k' is represented by the first term, and the loss due to transitions from k into other states k' is represented by the second term.

The time evolution of the probability distribution for a continuous variable can be described using a Fokker–Plank equation. The Fokker–Plank equation for a single variable x has the form

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}[\gamma(x)p(x,t)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[D(x)p(x,t)] \tag{5}$$

where γ is the deterministic drift term and D is the stochastic diffusion term. The Fokker–Plank equation is often used as an approximation of the ME (4). We refer the reader to ref. 16 for a thorough introduction to the subject.

A chemical master equation (CME) accounts for the random timing in the birth and death of individual molecules caused by the nondeterministic timing of individual reactions. As elegant as the CME formalism is, it usually cannot be solved analytically. Consequently, one either has to resort to approximations or simulate every individual state transition occurring in the system. Daniel Gillespie's stochastic simulation algorithm is a Monte Carlo simulation of the very process that the CME describes, and is the gold standard for simulating biochemical reaction systems [14, 15].

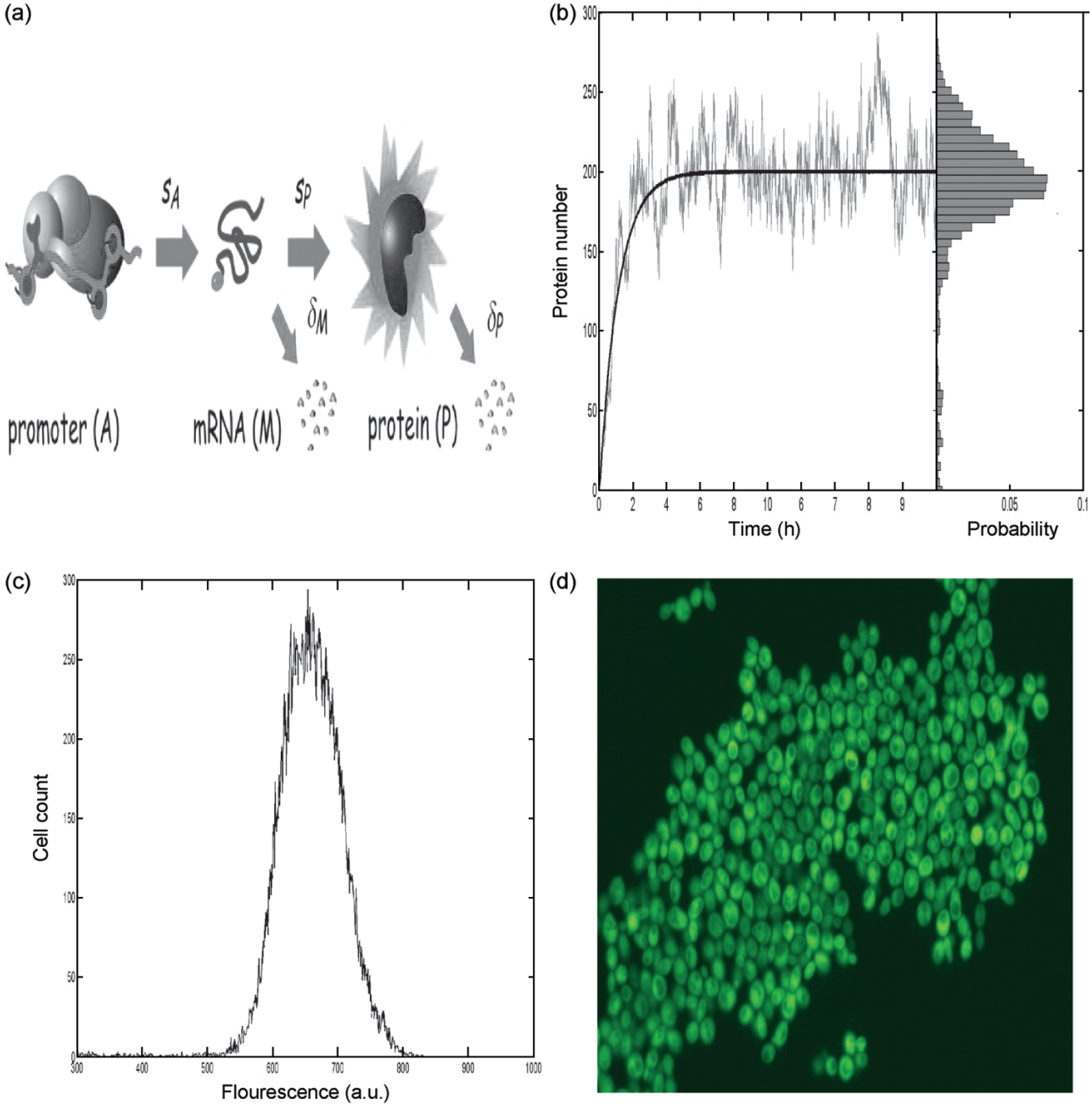
The Gillespie algorithm is often applied to simulate the gene expression process inside living cells. This process is fundamental to all life and is one of the most actively researched topics in science today. Physicists have long been interested in genetics; in his famous book *What is Life?*, Erwin Schrödinger introduced the idea of a gene as an aperiodic structure that stored genetic information in its configuration of covalent chemical bonds [17]. He also predicted that due to the order present in living organisms, DNA must be made up of a large number of atoms to counter the property of increasing randomness with smaller numbers of atoms. Within a decade, Watson and physicist Francis Crick deduced the double helical model for the structure of DNA [18]. Crick subsequently proposed the central hypothesis of molecular biology, namely that the gene expression process involves copying DNA into mRNA (transcription) and the production of a protein from this mRNA template (translation) (Fig. 1*a*) [19, 20]. Importantly, gene expression involves the collisions of small numbers of particles. Usually only one or two copies of DNA are found in a cell along with small numbers of mRNAs and transcription factors [21], and thus gene expression is an inherently noisy process (see (3) and Fig. 1*b*) like the processes observed by Delbrück and Spudich and Koshland [11, 13].

Gene expression can be measured experimentally using fluorescent proteins. More precisely, the gene coding for the fluorescent protein is placed beside a gene of interest such that they are transcribed and translated together. The degree of fluorescence, which indicates the level of gene expression, can then be measured in individual cells by flow cytometry to produce a population "snapshot" in the form of a gene expression distribution (Fig. 1*c*), or by time-lapse microscopy to produce a time series (Fig. 1*d*).

Current Research

The noise in gene expression allows for variation to exist among genetically identical cells in the same environment (for a comprehensive review see refs. 22 and 23). This is of particular interest because it can allow some members of a population to survive while others perish [24] (Fig. 2*a*). For instance, Blake et al. [25] observed that genetically identical yeast populations engineered to have higher noise (more cell-to-cell variation) reproduced faster than low noise populations when exposed to high levels of an antibiotic. Noise in gene expression also allows for "elastic adaptation", which occurs when the noise-generated distribution of a phenotype changes reversibly due to an environmental stress such that the reproductive fitness of a population in the new environment is optimized. For example, populations of yeast cells have been observed to adapt to long-term exposure to a drug

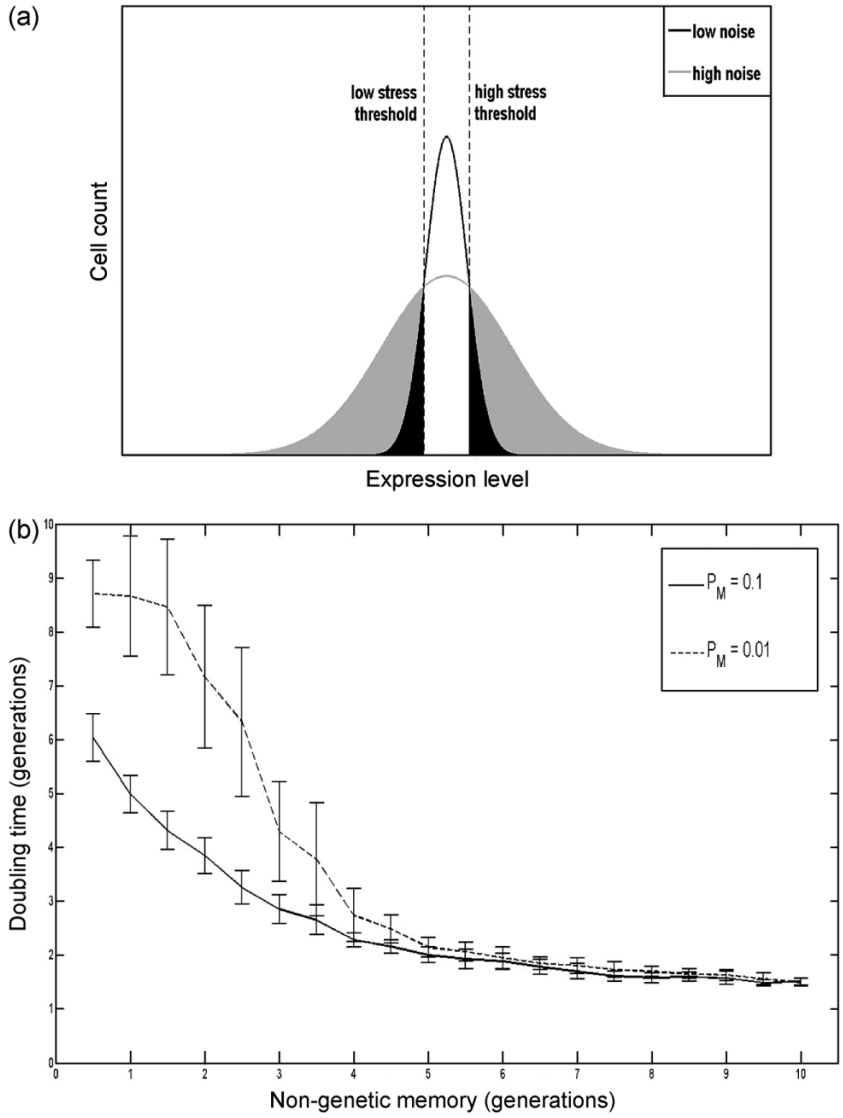
Fig. 1. Gene expression is a stochastic process. (*a*) A simple two-step model of gene expression. The schematic shows the synthesis of mRNA (M) from a gene with an active promoter (A) at a rate S_A, and the synthesis of protein (P) from an M template at a rate S_P, and the decay of M and P molecules at rates δ_M and δ_P, respectively. Reprinted with permission from (Scott et al. *Chaos*, **16**, 026107-2, (2006)). Copyright 2006, American Institute of Physics. (*b*) Time series of protein number generated by deterministic (solid black line) and stochastic (gray line) simulations. The histogram in the right-hand panel corresponds to the stochastic simulation and shows the probability that a cell will have a given intracellular protein level. Reprinted by permission from Macmillan Publishers Ltd. (Kaern et al. *Nat. Rev. Genet.* **6**, 453, copyright 2005. (*c*) Experimental green fluorescent protein (GFP) distribution for a clonal population of budding yeast obtained via flow cytometry (unpublished data). (*d*) GFP expression for a clonal population of budding yeast obtained using a microfluidics device (unpublished data).



by shifting the gene expression distribution in the direction that minimizes the impact of a drug [26]. When the drug is removed, the shift in gene expression can revert back to the distribution observed before the drug was applied [26]. This phenomenon has been attributed to nongenetic memory, as opposed to genetic memory where a mutation in the DNA would result in a permanent shift in gene expression. The term "nongenetic memory" can generally be defined as any mechanism that produces an enduring phenotype without altering the DNA sequence.

Genetic networks can store nongenetic memory in two or more discrete, stable states of network activity (see ref. 27

Fig. 2. Gene expression noise confers survival in clonal cell populations. (*a*) Schematic illustration of distributions for a low- and high-noise population. A greater number of cells in the high-noise population express above (cell survival) and below (cell death) the high- and low-stress thresholds, respectively. Reproduced with permission from (Fraser et al. *Mol. Microbiol.* **71**, 1335 (2009)). In the case shown here, the high-noise population has a higher fitness than the low-noise population when the stress is high, vice versa when the stress is low. (*b*) Effect of nongenetic memory and probability of mutation (P_M) per generation on the time for a simulated cancer cell population undergoing prolonged drug treatment to double. Note that the doubling time is more or less unaffected by P_M when the nongenetic memory is roughly above four generations and that in both cases a drug-resistant cell population develops. Reprinted figure with permission from (Charlebois et al. *Phys. Rev. Lett.* **107**, 218101-4 (2011)). Copyright 2011 by the American Physical Society.



for a review). For instance, in yeast, nongenetic memory can be enhanced by reducing the rate of stochastic transitions between two stable gene expression states [28]. In human cells, Brock et al. [29] proposed that nongenetic memory conferring temporary drug resistance contributes to tumour development by increasing the chance that some cells acquire a mutation conferring permanent immunity to the treatment regime. Nongenetic memory can also be stored in the lifetime of the gene expression fluctuations. That is, the lower the frequency of the noise the higher the level of the nongenetic memory, as the previous state is "remembered" by the cell for a longer period of time than at higher frequency noise. This was shown using an Ornstein and Uhlenbeck process [10] (see (2)) to be sufficient for the development of long-term drug resistance, independent of genetic memory conferring resistance [30] (Fig. 2*b*). This hypothesis is currently being investigated experimentally.

Conclusion

This is a new era for biology, one where more and more physicists are playing leading roles and driving the field to become more quantitative. The mathematical models being developed are helping to better explain the data gathered in the laboratory and to predict novel behaviour. In particular, stochastic models are being used increasingly in preference to deterministic models to describe biochemical networks and elucidate dynamics at the single-cell level [21]. Due to

the randomness inherent in living systems, an understanding of the source of this randomness and its effects is of fundamental importance. Fortunately, due to the foundations laid by early physicists, and the familiarity of many physical scientists today with the theory of stochastic processes, we can obtain a deeper understanding of biological systems.

Acknowledgments

The authors would like to acknowledge D. Jedrysiak for carrying out the experiments together with DC to obtain the data for Figs. 1*c* and 1*d*. We would also like to thank N. Abdennur for helpful discussions.

References

1. R. Brown. Philos. Mag. **4**, 161 (1828).
2. R. Brown. Philos. Mag. **6**, 161 (1829).
3. L.-G. Gouy. C.R. Hebd. Seances Acad. Sci. **CIX**, 102 (1889).
4. A. Einstein. Ann. Phys. **322**, 549 (1905). doi:10.1002/andp.19053220806.
5. A. Einstein. Ann. Phys. **324**, 371 (1906). doi:10.1002/andp.19063240208.
6. A. Einstein. Angew. Phys. Chem. **13**, 41 (1907).
7. A. Einstein. Zeitschrift für Elektrochemie und Angewandte Physikalische Chemie, **14**, 235 (1908). doi:10.1002/bbpc.19080141703.
8. P. Langevin. C.R. Acad. Sci. (Paris), **146**, 530 (1908).
9. W. Horsthemke and R. Lefever. Noise-Induced Transitions: Theory and Applications in Physics, Chemistry, and Biology. Springer-Verlag, Berlin. 2006.
10. G. Uhlenbeck and L. Ornstein. Phys. Rev. **36**, 823 (1930). doi:10.1103/PhysRev.36.823.
11. M. Delbrück. J. Bacteriol. **50**, 131 (1945).
12. M. Delbrück. J. Chem. Phys. **8**, 120 (1940). doi:10.1063/1.1750549.
13. J.L. Spudich and D.E. Koshland, Jr. Nature, **262**, 467 (1976). doi:10.1038/262467a0. PMID:958399.
14. D.T. Gillespie. J. Comput. Phys. **22**, 403 (1976). doi:10.1016/0021-9991(76)90041-3.
15. D.T. Gillespie. J. Phys. Chem. **81**, 2340 (1977). doi:10.1021/j100540a008.
16. N.G. van Kampen. Stochastic Processes in Physics and Chemistry. North-Holland Personal Library, Amsterdam, North-Holland. 1992.
17. E. Schrödinger. What is Life?: The Physical Aspect of the Living Cell, Based on lectures delivered under the auspices of the Dublin Institute for Advanced Studies at Trinity College, Dublin. February 1943. 1944.
18. J.D. Watson and F.H.C. Crick. Nature, **171**, 737 (1953). doi:10.1038/171737a0. PMID:13054692.
19. F.H.C. Crick. Symp. Soc. Exp. Biol. **XII**, 138 (1958).
20. F. Crick. Nature, **227**, 561 (1970). doi:10.1038/227561a0. PMID:4913914.
21. D.J. Wilkinson. Nat. Rev. Genet. **10**, 122 (2009). doi:10.1038/nrg2509. PMID:19139763.
22. M. Kaern, T.C. Elston, W.J. Blake, and J.J. Collins. Nat. Rev. Genet. **6**, 451 (2005). doi:10.1038/nrg1615. PMID:15883588.
23. M.S. Samoilov, G. Price, and A.P. Arkin. Sci. STKE, **2006**, re17 (2006). doi:10.1126/stke.3662006re17. PMID:17179490.
24. D. Fraser and M. Kaern. Mol. Microbiol. **71**, 1333 (2009). doi:10.1111/j.1365-2958.2009.06605.x. PMID:19220745.
25. W.J. Blake, G. Balázsi, M.A. Kohanski, F.J. Isaacs, K.F. Murphy, Y. Kuang, C.R. Cantor, D.R. Walt, and J.J. Collins. Mol. Cell, **24**, 853 (2006). doi:10.1016/j.molcel.2006.11.003. PMID:17189188.
26. D. Zhuravel, D. Fraser, S. St-Pierre, L. Tepliakova, W.L. Pang, J. Hasty, and M. Kaern. Syst. Synth. Biol. **4**, 105 (2010). doi:10.1007/s11693-010-9055-2. PMID:20805931.
27. S. Huang. Prog. Biophys. Mol. Biol. Available online 10 May 2012. doi:10.1016/j.pbiomolbio.2012.05.001.
28. M. Acar, A. Becskei, and A. van Oudenaarden. Nature, **435**, 228 (2005). doi:10.1038/nature03524. PMID:15889097.
29. A. Brock, H. Chang, and S. Huang. Nature, **10**, 336 (2009).
30. D.A. Charlebois, N. Abdennur, and M. Kaern. Phys. Rev. Lett. **107**, 218101 (2011). doi:10.1103/PhysRevLett.107.218101. PMID:22181928.