Listening to the noise: random fluctuations reveal gene network parameters

Brian Munsky1,*, Brooke Trinh2 and Mustafa Khammash3,*

1 Computer, Computational, and Statistical Sciences Division (CCS), and the Theoretical (T) Division at Los Alamos National Laboratory, Los Alamos, NM, USA, 2 Department of Molecular, Cellular and Developmental Biology, University of California, Santa Barbara, CA, USA and 3 Department of Mechanical Engineering and Center for Control, Dynamical Systems and Computations, University of California, Santa Barbara, CA, USA

* Corresponding authors. B Munsky, CCS-3 and CNLS, Los Alamos National Lab, Los Alamos, NM 87545, USA. Tel.: +1 505 665 6661; Fax: +1 505 665 7652; E-mail: brian.munsky@gmail.com or M Khammash, Department of Mechanical Engineering, University of California, Engineering II Building Room 2333, Santa Barbara, CA 93106, USA. Tel.: +1 805 893 4967; Fax: +1 805 893 8651; E-mail: khammash@engr.ucsb.edu

Received 20.2.09; accepted 8.9.09

The cellular environment is abuzz with noise originating from the inherent random motion of reacting molecules in the living cell. In this noisy environment, clonal cell populations show cell-to-cell variability that can manifest significant phenotypic differences. Noise-induced stochastic fluctuations in cellular constituents can be measured and their statistics quantified. We show that these random fluctuations carry within them valuable information about the underlying genetic network. Far from being a nuisance, the ever-present cellular noise acts as a rich source of excitation that, when processed through a gene network, carries its distinctive fingerprint that encodes a wealth of information about that network. We show that in some cases the analysis of these random fluctuations enables the full identification of network parameters, including those that may otherwise be difficult to measure. This establishes a potentially powerful approach for the identification of gene networks and offers a new window into the workings of these networks.

Molecular Systems Biology 5: 318; published online 13 October 2009; doi:10.1038/msb.2009.75

Subject Categories: computational methods; simulation and data analysis

Keywords: gene regulatory networks; stochastic biological processes; system identification

This is an open-access article distributed under the terms of the Creative Commons Attribution Licence, which permits distribution and reproduction in any medium, provided the original author and source are credited. This licence does not permit commercial exploitation or the creation of derivative works without specific permission.

Introduction

Computational modeling in biology seeks to reduce complex systems to their essential components and functions, thereby arriving at a deeper understanding of biological phenomena. However, measuring or estimating key model parameters can be difficult when measurement noise corrupts experimental data. Thus, when cellular variability or ‘noise’ (Elowitz et al., 2002) leads to measurement fluctuations, it may appear deleterious. However, this is not the case. Just as white noise inputs help to identify dynamical system parameters (Ljung, 1999), so too can characterization of noise dynamics elucidate natural mechanisms. For example, steady state noise characteristics can distinguish between different logical structures, such as AND or OR gates (Warmflash and Dinner, 2008). At the same time, temporal measurements of transient dynamics can aid in the construction of reaction pathways (Arkin et al., 1997). In combination, noise and temporal analyses yield powerful tools for parameter identification. For example, the averages of correlations in cell expression at many time points reveal feed-forward loops in the galactose metabolism genes of Escherichia coli (Dunlop et al., 2008). Similarly, manipulating certain gene network transcription rates while observing the response of statistical cumulants can help to identify reaction rates for some gene regulatory networks (Raffard et al., 2008).

In this study, we examine the possibility of identifying system parameters and mechanisms directly from single-cell distributions, such as those obtainable with flow cytometry, without time-varying control and at only a handful of different time points. We prove that the analysis of variability provides more information than the mean behavior alone. Furthermore, we illustrate potential of our approach using numerical and experimental analyses of common gene regulatory networks.

Results and discussion

Gene expression model

We adopt the gene expression model used in the study carried out by Thattai and van Oudenaarden (2001), which is characterized by random integer numbers of mRNA and protein molecules: R and P, respectively. Transcription, translation, and degradation events change the system state.
by altering these numbers. mRNA changes are modeled as random events that occur according to exponentially distributed waiting times that depend on the transcription and degradation rates \(k_r\) and \(\gamma_r\). Thus, given the state of \(r\) mRNA molecules, the probability that a single mRNA molecule is degraded within the time increment \(dr\) is given by \(r \cdot (\gamma_r \cdot dr)\). Similarly, translation and degradation of proteins are dictated by rates \(k_p\) and \(\gamma_p\). The resulting stochastic model is represented by a continuous time, discrete state Markov process. The probability of finding the system in a given state \((R(t)=r, P(t)=p)\) is fully characterized by the system’s master equation from which the evolution of moments \(E[R(t)], E[P(t)], E[R^2(t)], \ldots\) can be described (see Supplementary Section 1).

Our first finding is that all parameters of this model are identifiable from cell population distributions of mRNA/protein measured at least at two time points. In contrast, two time point measurements of mRNA/protein population averages are never sufficient for identifiability. To show this, the use of first and second-order moments, or equivalently means, variances, and covariances of proteins and mRNAs is sufficient, instead of the use of full distributions. At a given time point, \(t\), each such measurement yields a vector: \(v(t):=\{E[R(t)], E[P(t)], E[R(t)^2], E[P(t)^2], E[R(t)P(t)]\}\). Given \(v(t_0)\) and \(v(t_1)\) at two distinct time points \(t_0 < t_1\), there generically exists a set of parameters \(k_r, k_p, \gamma_r, \gamma_p\) that uniquely gives these measurements—all other parameter sets yield different measurements (see Figures 1E and 2A).

We illustrate this here only for transcription (Supplementary Section 3 provides an implicit expression for the parameters of the full model). Suppose that \(\{\mu_R, \sigma_R^2\}\) and \(\{\mu_P, \sigma_P^2\}\) represent the measured mRNA mean and variance at two time points \(t_0 < t_1 < \infty\). Then the parameters, \(\{k_r, \gamma_r\}\) are fully identifiable, and

\[
\gamma_r = \frac{1}{2\tau} \log \left( \frac{\sigma_R^2}{\sigma_R^2} - \frac{\mu_1}{\mu_0} \right),
\]

\[
k_r = \gamma_r - \mu_1 \exp \left( -\gamma_r \tau \mu_0 \right),
\]

where \(\tau := t_1 - t_0\).

Thus, the statistics, \(\{\mu_R, \sigma_R^2, \mu_P, \sigma_P^2\}\), contain sufficient information to identify the model parameters. However, measurement of just the population averages, for example, \(E[R]\), is insufficient for identifiability, and there exists an infinite set of parameters \(\{k_r, \gamma_r\}\), that is consistent with the same two mean measurements \(\mu_R\) and \(\mu_P\).

Although parameters are identifiable from transient moment measurements, we find that it is impossible to identify all parameters from stationary moments. Measuring means, variances, and other statistics after all the transients have died represents a lost opportunity to peek into the cell’s inner workings and to recover the network parameters. For example, two different parameter sets may produce very different protein distributions after a short interval time (Figure 1E), but indistinguishable distributions after a longer interval (Figure 1D). Supplementary Section 2 provides a proof that stationary moments of any arbitrary order are insufficient to uniquely identify the model parameters \(k_r, k_p, \gamma_r, \gamma_p\). Such stationary distributions will only enable the determination of relative parameter values, but any positive scaling of these values would produce the exact same measurements for \(v_p\). We note that stationary correlations, for example, \(E[R(t)R(t+\tau)]\) for small time intervals, \(\tau\), could also provide the necessary dynamic information (Cinquemani et al, 2009), but taking such measurements is more difficult and requires the tracking of individual cells between measurement times.

After having determined that full identification is achievable using two measurements of all first and second order moments, we now explore the effect of partial moment measurements. We consider two new scenarios: (a) only \(\{E[R], E[P]\}\) measurements are available; and (b) only \(\{E[P], E[R^2]\}\) measurements are available. For each scenario, Figure 2A shows the number of measurements needed for parameter identifiability and demonstrates the advantage of using full second order statistics. Furthermore, the performance with partial information depends on which partial information is being used. When protein and mRNA mean

![Figure 1](image_url) (A) Simple gene expression model representing gene transcription and translation. (B, C) Simulations of mRNA (green) and protein (blue) populations. The solid red lines denote the mean values and the dashed lines are one s.d. value above and below that mean. (D, E) mRNA (green) and protein (blue) distributions at (D) \(t=5000\) s and (E) \(t=1000\) s for two different parameter (solid or dotted lines) sets but with the same initial conditions.
The gene expression model has four unknown parameters that are needed to uniquely solve for the unknown parameters. The above results can be understood by exploring the number of such constraints algebraic constraints on model parameters. The above results (Section 4). When only protein mean and variance measurements are used, full parameter identifiability is never enough (see Figure 2A). However, for a model that describes only protein mean and variance measurements, at least five time measurements are needed for full parameter identifiability. In this case, the dynamics of \{E[P], E[P^2]\} are coupled to those of \{E[R], E[R^2], E[RP]\}, and additional measurements are needed to identify the initial values for these. Finally, we note that in these cases, the number of measurements needed for parameter identification are far fewer than the $2p + 1$ measurements that were shown (Sontag, 2002) to be sufficient for the identification of the $p$ unknown parameters of a general nonlinear dynamical system.

The results above establish the principle that transient measurements of full second order moments carry information that allows one to identify all model parameters, at least, assuming noise-free measurements. If the measurements are corrupted by noise, it is often possible to compensate with a larger number of measurements. To illustrate this, we have conducted 100 simulated identification studies in which the unknown parameters were taken from a broad lognormal distribution (Figure 2B). For these, we supposed that $v_j=\gamma_j/\kappa_j$ could be measured at $m$ equally separated time points $\{t_0, ..., t_m\}$, and that each measurement had unknown errors of $\pm 10\%$. To explore the effect of incomplete measurements, we performed the identification method for the three data scenarios considered earlier: (1) all moments; (2) only the means; and (3) only the protein means and variances. For each scenario, we investigated the impact on parameter identification of using an increasing number of noisy measurements obtained from a different number of independent experiments (with different randomly chosen unknown initial conditions).

As more data were gathered, the effects of measurement error were overcome and the probability of successful identification increased for every strategy (see Figure 2C). Using many measurements, the parameters and the unknown initial conditions of mRNAs and proteins could be resolved even from inaccurate protein data alone—provided that it included information on the protein variance. All of the above numerical experiments were conducted assuming that the initial conditions were unknown; for known or specified initial conditions, we found that the identification was even more successful (see Supplementary Figure 5). We have thus shown that for the simple gene expression model, cellular noise enhances the opportunity for system parameter identification, whereas measurement noise impedes it. The deleterious effects of measurement noise can be overcome by increasing the number of measurements.

**Experimental identification of lac induction**

Among the most studied gene regulatory elements is the lac operon of *Escherichia coli*. This mechanism has been used to construct toggle switches (Gardner *et al.*, 2000; Kobayashi *et al.*, 2000).
2004), genetic oscillators (Elowitz and Leibler, 2000; Atkinson et al., 2003) and logical circuits (Weiss, 2001). Despite its ubiquitous use, precise in vivo single-cell quantification of the system remains insufficient. Indeed, most such quantification attempts have come from in vitro experiments or population level studies. For example, the lac repressor dissociation constant has been estimated to be \( K_d = 10^{-11} \text{M} \) (Oehler et al., 1990). In an E. coli cell with a volume of 10\(^{-15}\) l, such dissociation constants mean the occupancy of the lac promoter is 94–99.9% when there are ten such molecules. At best, such measurements have only a probabilistic meaning at the level of single cells; at worst, they have no relevance at all as other mechanisms, such as nonspecific binding (Kao-Huang et al., 1977), take on much greater significance.

We used flow cytometry experiments and computational analyses to identify a parameter set to describe the in vivo single-cell dynamics of green fluorescent protein (GFP) controlled by the lac operon under isopropyl-\(\beta\)-D-thiogalactoside (IPTG) induction (see Figure 3A and Materials and methods section). We explored the response of the system at several IPTG levels and at multiple time points. Although many mechanistic models may capture the available data, we focused on the simplest consistent model, which consists of diffusion of IPTG into the cell, \([\text{IPTG}]_\text{in} = [\text{IPTG}]_\text{out}\).

**Figure 3** Experimental identification of a simple construct (A) in which IPTG induces the production of GFP. (B) Experimentally measured histograms of gfp expression on two different days (solid blue and green lines—in arbitrary units) and the best determined parameter fit (red dashed lines). Here, each column corresponds to a different measurement time (0, 3, 4, or 5 h) after induction, and each row corresponds to a different level of extracellular IPTG induction (5, 10, or 20 \(\mu\)M). In the parameter fits, different weights were applied to each experimental condition, shown as the values \(q\) in the histograms. (C) Predicted (red) and then measured (blue and green) fluorescence at 40 and 100 \(\mu\)M.
analyses. First, the production and degradation rates of LacI are measured. Even so, it is possible to make some multiple events (folding dynamics, elongation, etc.), these unknown positive real parameters for the regulatory system, which characterizes uncertainties in the flow cytometry measurements (see Materials and methods). In total, there are ten parameters to values in the literature, we have used the parameter set identified from [5, 10, 20] μM IPTG induction to predict the fluorescence under [40, 100] μM IPTG. Figure 3C shows that these predictions match the subsequent experimental measurements very well despite the vastly different shapes observed at the high induction levels.

Using single-cell experimental techniques, it has become possible to efficiently measure fluctuations in cell constituents. When properly extracted and processed with rapidly improving computational tools, these measurements contain sufficiently rich information as to enable the unique identification of parameters. We have shown that transient dynamics are important to this effort, and in principle, identification can be accomplished when accurate distributions are measured at only two distinct time points. More time points are needed if the distributions are poorly measured, but the idea remains the same. We have show the potential of our approach by experimentally identifying a predictive model of lac regulation from flow cytometry data. Hence, the proposed integration of single-cell measurements and stochastic analyses establishes a promising approach that offers new windows into the workings of cellular networks.

Materials and methods

Medium and reagents

Cells were grown in Luria–Bertani (LB) medium supplemented with 1% tryptone, 0.5% yeast extract, and 0.4% NaCl and containing IPTG at the concentrations noted. To select for plasmid maintenance, antibiotics were used at the following concentrations: 100 μg/ml ampicillin (amp); 40 μg/ml kanamycin (kan); and 12.5 μg/ml tetracycline (tet).

Bacterial strains and plasmids

The E. coli strain used was DL5905—E. coli K-12 (isolate MC4100) containing F’ proAB lacZM15 Trn1 (Tet') from strain XL-1 Blue (Stratagene) and plasmid pDLS12. To construct plasmid pDLS12, GFP(LVA) (Anderson et al., 1998) was PCR amplified from plasmid pRK9 (a gift from John Cronan) using the forward primer (5’-CAACA AAGATCATTAAAGGAGAAATAAGGATGATGAAAGGAAGAAC TTCTCA-3’) that includes a BgII site and removes an SpHl site from the original pRK9 sequence, and the reverse primer (5’-CAACAGACATGCA TTAAGCCTAATAAAAAGGATGTTTTCGTTTCC-3’) that adds an SpHl site. This fragment was digested using BgII and SpHl and cloned into BgII and SpHl sites of pLAC33 (Warren et al., 2000), removing a portion of the Tet’ cassette.

Fluorescence induction experiments

Twenty-four separate cell cultures were allowed to grow in LB broth containing the appropriate antibiotics to an approximate ODoo of 0.2, and were then induced with [0, 5, 10, 20, 40, 100] μM concentrations of IPTG at 5, 4, 3, and 0 h before flow cytometry measurements. Flow cytometry was carried out using a BD Biosciences FACSaria instrument with a 100-μm sorting nozzle at low pressure. GFP(LVA) was excited using a 488-nm blue laser and detected using 530/30-nm filter. For each sample, 1,000,000 events were collected. To ensure repeatability, experiments were conducted twice, each on a separate day.
**GFP induction model**

The stochastic model for the IPTG–GFP induction is composed of four nonlinear production/degradation reactions given in the main text. The rates of these reactions depend on the integer populations of the proteins LacI and GFP, as well as the set of nonnegative parameters, \( \{k_a, k_c, \delta(I), \delta_c, \alpha, \tau, \eta\} \in \mathbb{R}^6 \). For the stochastic system modeled here, the joint (LacI, GFP) probability distributions of both proteins evolve according to the infinite dimensional chemical master equation (CME; van Kampen, 2007). This can in turn be expressed as an infinite set of linear ordinary differential equations—\( \dot{P}(t, A) = \mathcal{A}(A) \cdot P(t, A) \). Unlike in the simple transcription/translation model, the toggle reactions are nonlinear, and the CME has no known exact solution. We use a finite state projection approach (Munsky and Khammash, 2006) that makes it possible to approximate the solution to any degree of accuracy. For any error tolerance \( \varepsilon > 0 \), we systematically find a finite-dimensional projected system—\( \dot{\mathbf{P}}_{\text{SP}}(t, A) = \mathcal{A}_f(A, \cdot) \cdot \mathbf{P}_{\text{SP}}(t, A) \)—the solution for which is within the desired tolerance. More precisely, 

\[
\left| \begin{array}{c}
\mathbf{P}(t, A) - \mathbf{P}_{\text{SP}}(t, A) \\
\mathbf{P}(t, A) - \mathbf{P}_{\text{SP}}(0, A)
\end{array} \right| \leq \varepsilon, \quad \text{and} \quad \mathbf{P}_{\text{SP}}(0, A) = \mathbf{P}_0(0, A),
\]

where the index vector \( J \) denotes the set of states included in the projection, \( \mathbf{P}_i \) is the corresponding probability of those states, and \( \mathcal{A}_f \) is the corresponding principle sub-matrix of \( \mathcal{A} \) (Munsky and Khammash, 2006). The one-norm measure is used to ensure that absolute sum of the probability density error is guaranteed to lie within the tolerance. The solution of each projected master equation is found using the stiff ode solver ode23s in MathWorks Matlab.

**Modeling flow cytometry data**

In addition to modeling the regulatory dynamics of the system, one must also account for the inherent uncertainty within measured levels of fluorescence activity. The process used to account for this uncertainty has three components. First, in an effort to remove outliers in cell volume and density, and thereby reduce the effects of unmodeled dynamics, each cell population was gated separately using forward and side scatter data. Specifically, the forward and side scatter measurements were used to form a two-dimensional joint histogram with 50 x 50 logarithmically distributed bins (see Supplementary Figure 6). The maximum point in this histogram was recorded and then the gating region was chosen to include every bin that had, at least, one third as many counts as the maximal bin. Second, flow cytometry measurements in the absence of IPTG have been used to calibrate the background fluorescence of cell populations at various instances in time, and it has been assumed that the background fluorescence distribution, \( \mathbf{f}_{\text{BG}}(x) \), is independent of the levels of IPTG, LacI, and GFP. Third, each GFP molecule is assumed to emit a normally distributed random amount of fluorescence with unknown mean, \( \mu_{\text{GFP}} \), and variance, \( \sigma_{\text{GFP}}^2 \); both of which are to be identified. Thus, if \( \mathbf{p}_i = p_i(t, A, \text{[IPTG]}) \) denotes the probability of having exactly \( n \in \{0, 1, 2, \ldots\} \) molecules of GFP, then the probability density of having exactly \( x \) arbitrary units of fluorescence because of GFP is computed as:

\[
f_{\text{GFP}}(x) = \sum_{n=0}^\infty p_n \frac{1}{\sqrt{2\pi n} \cdot \sigma_{\text{GFP}}} \exp\left(-\frac{(x-n \cdot \mu_{\text{GFP}})^2}{2n \cdot \sigma_{\text{GFP}}^2}\right).
\]

Finally, the total observable fluorescence is the sum of the GFP fluorescence plus the background noise, and the distribution of total fluorescence is found using the convolutions:

\[
f_{\text{Tot}}(x) = \frac{d}{dx} \int_{-\infty}^{x} f_{\text{GFP}}(x-s) \cdot f_{\text{BG}}(s) \cdot ds \approx \int_{0}^{x} f_{\text{GFP}}(x-s) \cdot f_{\text{BG}}(s) \cdot ds.
\]

**Identification procedure**

With the FSP solution and the computation of the expected fluorescence, the identification procedure is carried out by finding the parameter vector \( \mathbf{A}^* \) that minimizes the one norm difference between the experimentally measured distribution \( f_{\text{Tot}}(t, \text{[IPTG]}) \) and the numerical solution of that distribution:

\[
\mathbf{A}^* := \arg\min_{\mathbf{A}} \sum_{i=1}^n q_i \cdot \left| \left| f_{\text{Tot}}(t, \mathbf{A}) - f_{\text{Tot}}(t, \mathbf{A}^*) \right| \right|_1,
\]

where the summation is taken over all of the different experimental conditions of different induction times and IPTG levels, and the weight \( q_i \) specifies the relative importance to each of these measurements. These weights have been chosen such that each IPTG level has the same total importance and so that greater importance is placed on measurements that differ the most from the background fluorescence. The values for these weights are given in Figure 3. The parameter identification is accomplished by starting with an initial parameter guess, \( \mathbf{A}_0 \), and then this set is updated iteratively using gradient-based and simulated annealing searches until the computed distribution matches the experimental distribution as closely as possible. The optimization procedure is repeated for multiple, randomly generated initial parameter guesses. An optimal parameter set is regarded as unique if the given solution yields the smallest achieved value for the objective function, and if that parameter has been achieved during many such identification runs each beginning with different parameter guesses.

**Supplementary information**

Supplementary information is available at the Molecular Systems Biology website (www.nature.com/msb).

**Acknowledgements**

The authors thank John Little and Hana El-Samad for thoughtful discussions and David Low for advice, assistance, and materials required in the experimental aspects of this study. This study is funded by the LANL LDRD program and NSF through Grants ECCS-0835847 and ECCS-0802008, and by ICB Grant DAAD19-03-D-0004 from the US Army Research Office.

**Conflict of interest**

The authors declare that they have no conflict of interest.

**References**

Andersen J, Sternberg C, Poulsen L, Bjorn S, Givskov M, Molin S (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. Appl Environ Microbiol 64: 2240–2246

Arkin A, Shen P, Ross J (1997) A test case of correlation metric construction of a reaction pathway from measurements. Science 227: 1275–1279

Atkinson M, Savageau M, Myers J, Ninfa A (2003) Development of genetic circuitry exhibiting toggle switch or oscillatory behaviour in E. coli. Cell 113: 597–607

Cinquemani E, Milias-Argeitis A, Summers S, Lygeros J (2009) Local identification of piecewise deterministic models of genetic networks, lecture notes in computer science. Springer 5469: 105–119

Dunlop M, Cox III R, Levine J, Murray R, Elowitz M (2008) Regulatory activity revealed by dynamic correlations in gene expression noise. Nat Genet 40: 1493–1498

Elowitz M, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. Nature 403: 335–338

Elowitz M, Levine A, Siggia E, Swain P (2002) Stochastic gene expression in a single cell. Science 297: 1183–1186

Gardner T, Cantor C, Collins J (2000) Construction of a genetic toggle switch in escherichia coli. Nature 403: 339–342
Kao-Huang Y, Revzin A, Butler AP, O’Conner P, Noble DW, von Hippel PH (1977) Nonspecific DNA binding of genome-regulating proteins as a biological control mechanism: Measurement of DNA-bound E. coli lac repressor in vivo. Proc Natl Acad Sci USA 74: 4228–4232
Kobayashi H, Kaern M, Araki M, Chung K, Gardner T, Cantor C, Collins J (2004) Programmable cells: interfacing natural and engineered gene networks. Proc Natl Acad Sci 101: 8414–8419
Ljung L (1999) System Identification, Theory for the User. Upper Saddle River, NJ, USA: Prentice-Hall
Munsky B, Khammash M (2006) The finite state projection algorithm for the solution of the chemical master equation. J Chem Phys 124: 044104
Oehler S, Eismann E, Krämer H, Müller-Hill B (1990) The three operators of the lac operon cooperate in repression. EMBO J 9: 973–979
Raffard R, Lipan O, Wong W, Tomlin C (2008) Optimal discovery of a stochastic genetic network. Proc 2008 Amer Contr Conf, Vol. 1, pp 2773–2779, 11–13 June 2008, Seattle, WA, USA
Sontag E (2002) For differential equations with r parameters, 2r + 1 experiments are enough for identification. J Nonlinear Sci 12: 553–583
Thattai M, van Oudenaarden A (2001) Intrinsic noise in gene regulatory networks. Proc Natl Acad Sci USA 98: 8614–8619
van Kampen N (2007) Stochastic Processes in Physics and Chemistry, 3rd edn. Amsterdam: Elsevier
Warmflash A, Dinner A (2008) Signatures of combinatorial regulation in intrinsic biological noise. Proc Nat Acad Sci USA 105: 17262–17267
Warren J, Walker J, Roth J, Altman E (2000) Construction and characterization of a highly regulable expression vector, pLAC11, and its multipurpose derivatives, pLAC22 and pLAC33. Plasmid 44: 138–151
Weiss R (2001) Cellular Computation and Communications using Engineered Genetic Regular Networks. PhD thesis, MIT, 2001

Molecular Systems Biology is an open-access journal published by European Molecular Biology Organization and Nature Publishing Group.
This article is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 Licence.