Variability in cellular gene expression profiles and homeostatic regulation

Author: Yuriy Mishchenko

Affiliation: Toros University

Address: Toros University, 45 Evler Campus, Mersin 33140 Turkey

Running title: Variability in cellular gene expression and homeostatic regulation

Corresponding Author:

Dr. Yuriy Mishchenko

45 Evler Campus Toros University Mersin 33140 Turkey

Email: yuriy.mishchenko@gmail.com

Phone:+90 (530) 244 62 40

Abstract

One of surprising recent discoveries in biology is that the gene and protein expression profiles in cells with identical genetic and environmental makeup can exhibit large variability. The nature and the significance of this variability had been posed as one of the current fundamental questions of biology. In this letter, we argue that the observed variability in cellular gene and protein expression can be understood as an outcome of homeostatic regulation mechanisms controlling the gene and protein expression profiles.

INTRODUCTION

One of recent findings in biology is that the gene and protein expression profiles in cellular systems with identical genetic and environmental makeup can often exhibit a large stochastic variability (*1–16*). The nature and the significance of this variability had been stated as one of the fundamental problems of biology. In this letter, we argue that the observed gene and protein expression variability emerges naturally as the outcome of the regulation of the gene and protein expression levels in biological cells, in the context of the cellular homeostatic regulation mechanisms.

One of the main challenges faced by biological systems is maintaining a stable physiological state in the face of the fluctuations in the parameters of internal and external environments. One of the best mechanisms available to biological systems for countering such disruptive influences is homeostatic regulation. Homeostasis is the property of biological systems to maintain a stable physiological state by means of various feedbacks mechanisms sensitive to the changes in that state (*17*). Homeostatic regulation responds directly to the changes of the controlled physiological parameter and, by doing so, can be very effective in maintaining the necessary value of that parameter. In this letter, we show that the property of the homeostatic regulation to be sensitive directly to a system's physiological state also leads with necessity to a large variability observed in the internal configurations of the system. This conclusion is general and is based broadly on two properties of biological systems – the possibility of implementing the same physiological state via different internal configurations and the reliance of homeostatic regulation on that physiological state for feedback. We present a general argument towards that point and demonstrate the emergence of this phenomenon in a similated model, in-silico.

MATERIALS AND METHODS

We consider a simple model here with the goal of inspecting the impact of the homeostatic regulation on the internal configuration states of a population of biological cells. Specifically, we consider a simple model of the expression control for a single protein in a biological cell, whereas (importantly) the protein production can be regulated via multiple pathways. The time-evolution of the protein concentration in the cell in that case is given by the following relationship,

$$\frac{dX_k}{dt} = -aX_k + ru_k + e\varepsilon_k, \quad (1)$$

where X_k is the protein concentration in the cell due to the regulation pathway *k*, aX_k is the rate of natural degradation of the protein, ru_k is the rate of the protein production in the pathway *k*, and ε_k is the standard normal noise variable with zero mean and unit variance. a, r, e are constants and *k* is the index talking on the values from 1 to N, to enumerate different protein regulation pathways, where N is the total number of such pathways. u_k is understood as an internal control variable used by the cell to adjust the

protein production in specific pathway *k*. The key value of interest here is the total protein concentration in the cell, as given by $X = \sum_{k=1}^{k=N} X_k$.

To complete the model (1), it is necessary to define how the control variables u_k will depend on the internal state of the cell $\{X_k, k = 1 \dots N\}$. We inspect three possibilities here: (i) feed-forward regulation, (ii) individual regulation, and (iii) homeostatic regulation. In feed-forward regulation, the values of u_k are fixed at a constant value $u_k = u = a\bar{X}/Nr$, implying that the rate of protein production is fixed in a feed-forward manner, whereas $\bar{X}$ is the desired protein concentration. The equilibrium is then reached when $aX_k = ru_k$ or $X_k = \bar{X}/N$ and, thus, $X = \bar{X}$. In the individual regulation, the pathways are regulated by the cell using individual feedbacks modeled by $u_k = \bar{X} - NX_k$. In this case each pathway is driven towards a fixed point $X_k = \bar{X}/N$ also resulting in the equilibrium total protein concentration $X = \bar{X}$. Finally, in the homeostatic regulation, the pathways are regulated via a common homeostatic feedback originating from the actual concentration of the protein in the cell, X, $u_k = u = \bar{X} - X$. In this model, the regulating pathways are driven to a suitable configuration point by relying directly on the realized protein concentrations.

RESULTS

We argue that the large variability in the gene and protein expression levels observed in otherwise identical biological cells can be a direct consequence of homeostatic regulation mechanisms affecting the gene and protein networks of such cells. One of the main challenges faced by biological systems in maintaining their physiological state is the disruptive fluctuations in such systems' internal and external environments. One of the best mechanisms available to biological systems for countering such fluctuations is the homeostatic regulation.

Homeostasis is the property of biological systems to maintain a stable internal physiological condition by means of a feedback sensitive to the changes directly in the systems' physiological state (*17*). By responding directly to the changes in the systems' physiological state, homeostatic mechanisms can be extremely effective in maintaining the physiological state necessary for biological systems to perform their functions. When compared to other regulation strategies, homeostatic regulation both allows achieving a highly specific regulation of the physiological state in the face of noise as well as offers to the biological systems the capability of recovering from otherwise fatal internal failures, Figure 1.

The capacity of homeostatic regulation to respond directly to the changes of the controlled physiological parameter, which confers to it these important properties, also makes homeostatic regulation insensitive to such parameters' perturbations that do not produce a significant change in a system's physiological state. These can be, for example, random up-regulations of one and simultaneous down-regulations of other

pathways regulating a physiological parameter. Whenever such a perturbation occurs, no net change may result in the physiological state thus procuring no corrective actions from homeostatic feedback.

To consider this peculiar point in a greater detail, we investigate more thoroughly this phenomenon using a simple model of cellular regulation of a protein's expression level X controlled via two complementary regulation pathways X_1 and X_2, so that $X = X_1 + X_2$. If the production of such protein is affected by a significant amount of noise, as can be expected in real cells due to spontaneous protein degradation or thermodynamic noise during protein transcription, there can be multiple approaches for the cell to ensure that the concentrations of the protein in the cytoplasm remain at the levels required for normal functioning. One of the most basic such approaches is to maintain the rate of the protein production in each regulation channel at a fixed level, whereas a given final concentration $\bar{X}$ is achieved in equilibrium depending on the rate of production and the rate of degradation of the protein. Of course, such a simple open-loop regulation approach can be heavily affected by variability due to noise. A better approach may involve a closed-loop feedback that works within each protein regulation pathway and ensures that the protein production in that pathway remains at specified levels, such as $(\bar{X}_1, \bar{X}_2)$, that ensure the required total protein concentration $\bar{X}$. Yet another possibility involves regulating the protein production using feedbacks that rely directly on the actual protein concentration in the cytoplasm, measured in some way. The latter, of course, is the homeostatic regulation. Of the three approaches, the homeostatic regulation is the one that offers the highest degree of robustness.

It is clear now that, should we choose the homeostatic regulation to control the protein production in our cell, such regulation will be insensitive to the changes in the expression profiles of the individual pathways X_1 and X_2 such that do not affect the protein's total concentration $X = X_1 + X_2$. This basically corresponds to up-regulating the protein production in one pathway while simultaneously down-regulating the other, leaving the sum $X_1 + X_2$ intact. This insensitivity can have dramatic consequences for the cell's realized internal configurations (X_1, X_2). In Figure 2, we show the distribution of the internal configurations (X_1, X_2) in such a model cell, simulated over time. One can clearly see in these figures the spread of the cell's internal states along the direction $X_1 + X_2 = const$, where homeostatic regulation provides no corrective feedback. This spread is caused by the accumulation of small random perturbations of the cell's internal state that are "neutral", that is, such that produce no net change in the total protein concentration X and thus cause no corrective action from homeostatic feedback. This accumulation results in a large variability in the internal states of our model cells even if such cells share otherwise identical initial conditions and environments.

The above simple model notwithstanding, this phenomenon can have dramatic effects when the number of regulating pathways becomes large. In Figure 3, we show the

example of the same above model but now with $N = 10$ concurrent protein regulation pathways. The protein expression levels in individual pathways, X_k, in this case spread widely, filling entire regions of space and barely carrying any resemblance to the actually unique physiological state that these model cells live in. The variation in the individual pathways' expression levels is up to 10 times greater than such that would be expected in that model from simple feed-forward or non-homeostatic feedback regulation, Figure 2E and 2F.

As a conclusion, we find that the reliance of homeostatic regulation mechanisms on biological systems' actual physiological state can make such mechanisms insensitive to the perturbations in such systems' internal state that do not result in net change of such systems' physiological parameters. This makes it possible for such "neutral" perturbations to accumulate over time and result in a large spread of the internal state of such systems without any appreciable differences in such systems' phenotype, genotype, or external environment.

DISCUSSION AND CONCLUSIONS

We argue that the puzzling stochastic variability observed in the gene and protein expression of otherwise identical biological cells can be understood as a result of the homeostatic feedbacks in such cells' internal regulation mechanisms. Homeostatic regulation is ubiquitous in biology. Compared with the other regulation mechanisms, homeostatic regulation confers superior ability towards reducing stochastic fluctuations in physiological state as well as recovery from failures. The key property of homeostatic regulation is its responding to the changes in the actual physiological state of a biological system. This makes homeostatic regulation insensitive to the perturbations of biological systems' internal state that do not produce a net change in the controlled physiological parameters. Accumulation of such "neutral" perturbation over time can lead to large variations in the internal states of otherwise identical biological systems.

The wide spread of homeostatic regulation in biology implies that this phenomenon is likely to emerge not only in the gene and protein expression in populations of biological cells, but much more widely as a generic property of biological systems. Indeed, we note other possible examples of this phenomenon reported in axonal and dendritic ion-channel composition and the structure of the central pattern generator neural circuits in lobster somatogastric ganglion (*17–22*). One can expect more examples of similar phenomena to emerge in the future.

If homeostatic regulation is the primary cause of the stochastic variability observed in cellular gene and protein expression, then the discussion above allows making certain experimentally testable predictions. For instance, we should expect that higher amounts of such stochastic variability will be recorded in the systems having a larger number of internal degrees of freedom, that is, where more complementary pathways contribute to the regulation of the same physiological state. Furthermore, such stochastic variability

should be nearly completely decoupled from the variability in the phenotype while the variation of individual gene and protein expression levels should be, on the contrary, quite highly correlated. That is, large variations in gene or protein expression levels should not correlate with the variations in the observed phenotype, while the variations in the individual expression levels should not be independent and should exhibit a significant degree of correlation. More interestingly, however, is the prediction that the gene and protein expression profiles of individual cells should change with time. That is, the stochastic variability can be observed over a population of biological cells, but also using a single cell with the gene and protein expression profiles measured at different time points. Finally, we expect the examples of stochastic variability to be found in many different biological systems ranging from single cells to multicellular organisms, as well as complex biological systems such as neuronal circuits.

Acknowledgement: This work had been supported by the Bilim Akademesi-The Science Academy under the BAGEP program (Turkey), Toros University BAP grant number TUBAP135001 (Turkey), and TUBITAK ARDEB 1001 grant number 113E611 (Turkey).

References:

1. S. J. Altschuler, L. F. Wu, Cellular heterogeneity: do differences make a difference? *Cell*. **141**, 559–563 (2010).
2. A. Sigal, R. Milo, A. Cohen, N. Geva-Zatorsky, Y. Klein, Variability and memory of protein levels in human cells. *Nature*. **444**, 643–646 (2006).
3. S. Frank, Genetic variation of polygenic characters and the evolution of genetic degeneracy. *J. Evol. Biol.* **16**, 138–142 (2003).
4. M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, Stochastic gene expression in a single cell. *Science (80-.).* **297**, 1183–1186 (2002).
5. A. Bar-Even *et al.*, Noise in protein expression scales with natural protein abundance. *Nat. Genet.* **38**, 636–643 (2006).
6. J. Newman *et al.*, Single-cell proteomic analhysis of S. cerevisiae reveals the architecture of biological noise. *Nature*. **441**, 840–846 (2006).
7. E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, A. van Oudenaarden, Regulation of noise in the expression of a single gene. *Nat. Genet.* **31**, 69–73 (2002).
8. J. M. Pedraza, A. van Oudenaarden, Noise propagation in gene networks. *Science (80-.).* **307**, 1965–1969 (2005).
9. N. Rosenfeld, J. W. Young, U. Alon, P. S. Swan, M. B. Elowitz, Gene regulation at the single-cell level. *Science (80-.).* **307**, 1962–1965 (2005).
10. D. Austin *et al.*, Gene network shaping of inherent noise spectra. *Nature*. **439**, 608–611 (2006).
11. I. Mihalcescu, W. Hsing, S. Leibler, Resilient circadian oscillator revealed in individual cyanobacteria. *Nature*. **430**, 81–85 (2004).

12. W. J. Blake, M. Kaern, C. Cantor, J. Collins, Noise in eukaryotic gene expression. *Nature*. **422**, 633–637 (2003).
13. M. Acar, A. Becskei, A. van Oudenaarden, Enhancement of cellular memory by reducing stochastic transitions. *Nature*. **435**, 228–232 (2005).
14. A. Tobin, N. Cruz-Bermudez, E. Marder, D. Schulz, Correlations in ion channel mRNA in rhythmically active neurons. *PLoS One*. **4**, e6742 (2009).
15. D. Schulz, J. Goaillard, E. Marder, Variable channel expression in identified single and electrically coupled neurons in different animals. *Nat. Neurosci.* **9**, 356–362 (2006).
16. D. Schulz, J. Goaillard, E. Marder, Quantitative expression profiling of identified neurons reveals cell-specific constraints on highly variable levels of gene expression. *Proc. Natl. Acad. Sci.* **104**, 13187–13191 (2007).
17. W. B. Cannon, Organization for physiological homeostasis. *Physiol. Rev.* **9**, 339–431 (1929).
18. M. J. Berridge, M. D. Bootman, H. L. Roderick, Calcium signalling: dynamics, homeostasis and remodelling. *Nat. Rev. Mol. Cell Biol.* **4**, 517–529 (2003).
19. E. C. Butcher, L. J. Picker, Lymphocyte homing and homeostasis. *Science (80-.).* **272**, 60–67 (1996).
20. I. Lerner, *Genetic homeostasis* (Oliver and Boyd, 1954).
21. E. Marder, M. Goaillard, Variability, compensation and homeostasis in neuron and network function. *Nat. Rev. Neurosci.* **7**, 563–674 (2006).
22. T. O'Leary, A. Williams, J. Caplan, E. Marder, Corrleations in ion channel expression emerge from homeostatic tuning rules. *Proc. Natl. Acad. Sci.* **110**, E2645–54 (2013).

FIGURE 1

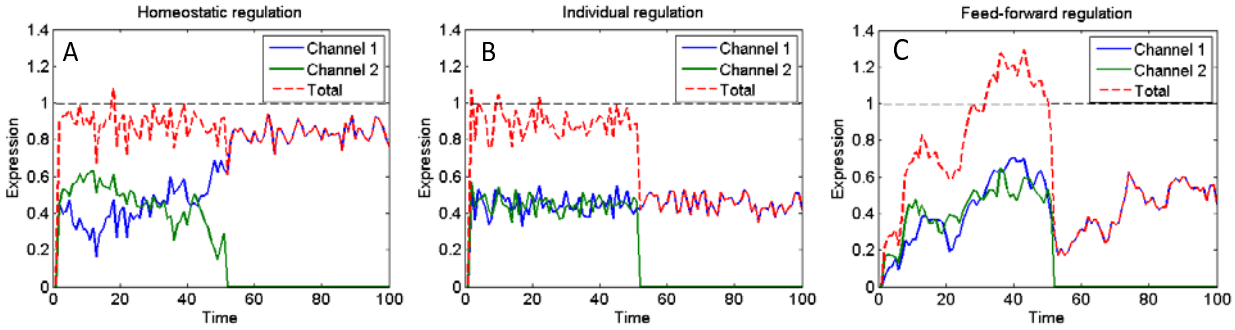

Figure 1: Time evolution of the total protein concentration in model of protein regulation in a cell, controlled by two complementary regulation pathways as discussed in the paper. A) Homeostatic regulation offers a high level of noise suppression in the final protein concentrations as well as allows recovering from failures in one of the pathways (such as pathway 2 blockage introduced here at time 50 units). B) Individual control of the protein production pathways, driving pathways towards a fixed operating point individually, also offers a high level of resistance to noise but fails to recover from the failure of one of the pathways. C) Simple feed-forward regulation both is highly vulnerable to noise and fails to recover from the failure in one of the pathways. Simulation parameters: $e = 0.05, a = 0.1, r = 0.5$. The target concentration $\bar{X} = 1$, as indicated by the dashed line.

FIGURE 2

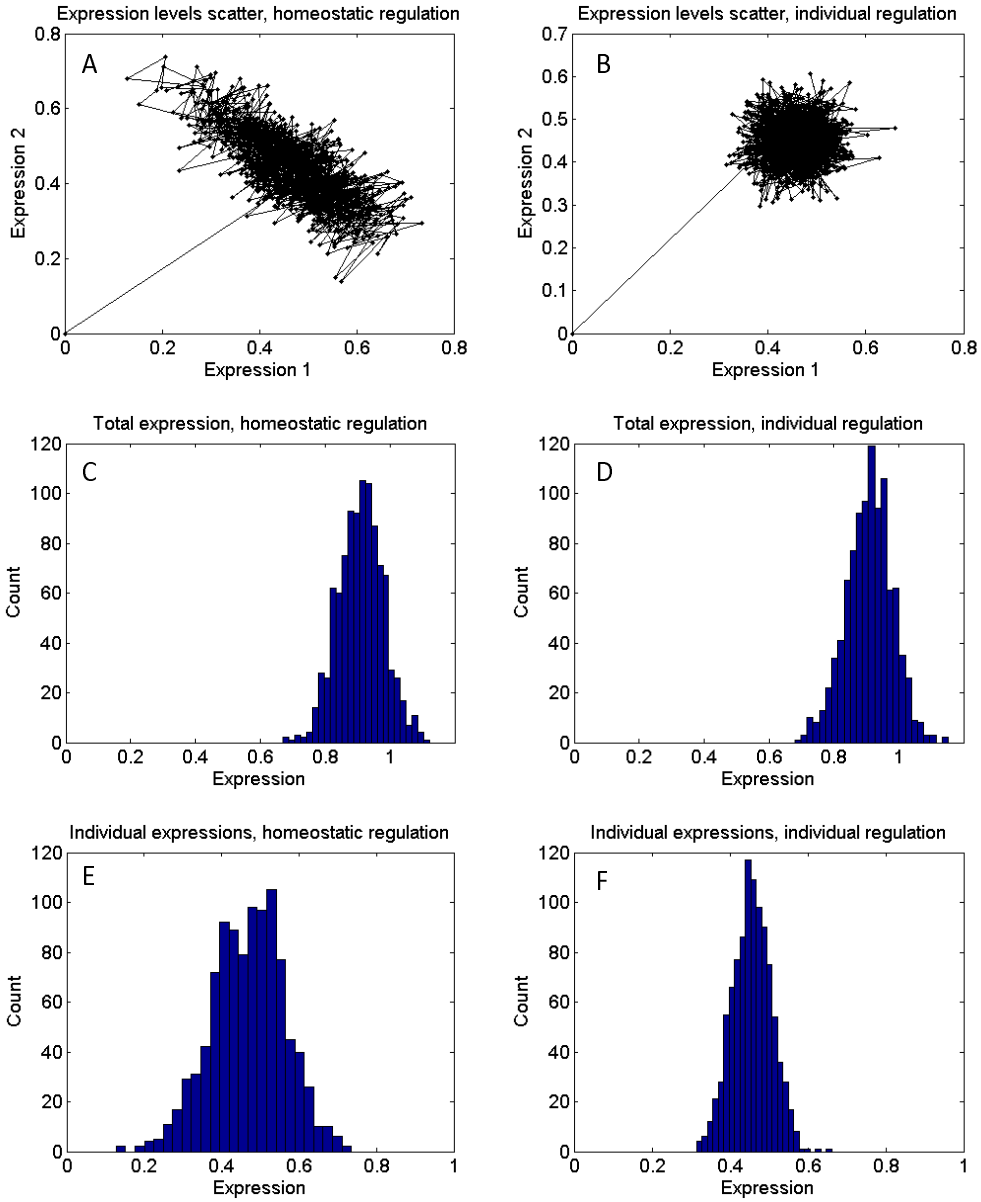

Figure 2: The insensitivity of homeostatic regulation to the perturbations of a biological system leaving the final physiological state unaffected results in a large spread of such systems' internal states along the dimension of "insensitivity" of homeostatic regulation. A) Realized internal configuration states (X_1,X_2) in the model (1) under homeostatic regulation; the spread of the internal states along the direction of X_1+X_2= X=*const* is clearly visible. B) Realized internal configuration states for the model (1) with individually controlled pathways X_1 and X_2. C-D) In both examples, the degree of variation in the final protein concentration in either homeostatic regulation (C) or individual regulation (D) is the same. E-F) Under homeostatic regulation the spread in the expression levels of the individual pathways X_1 and X_2 (left, E) is at least two times larger than that under individual regulation (right, F).

FIGURE 3

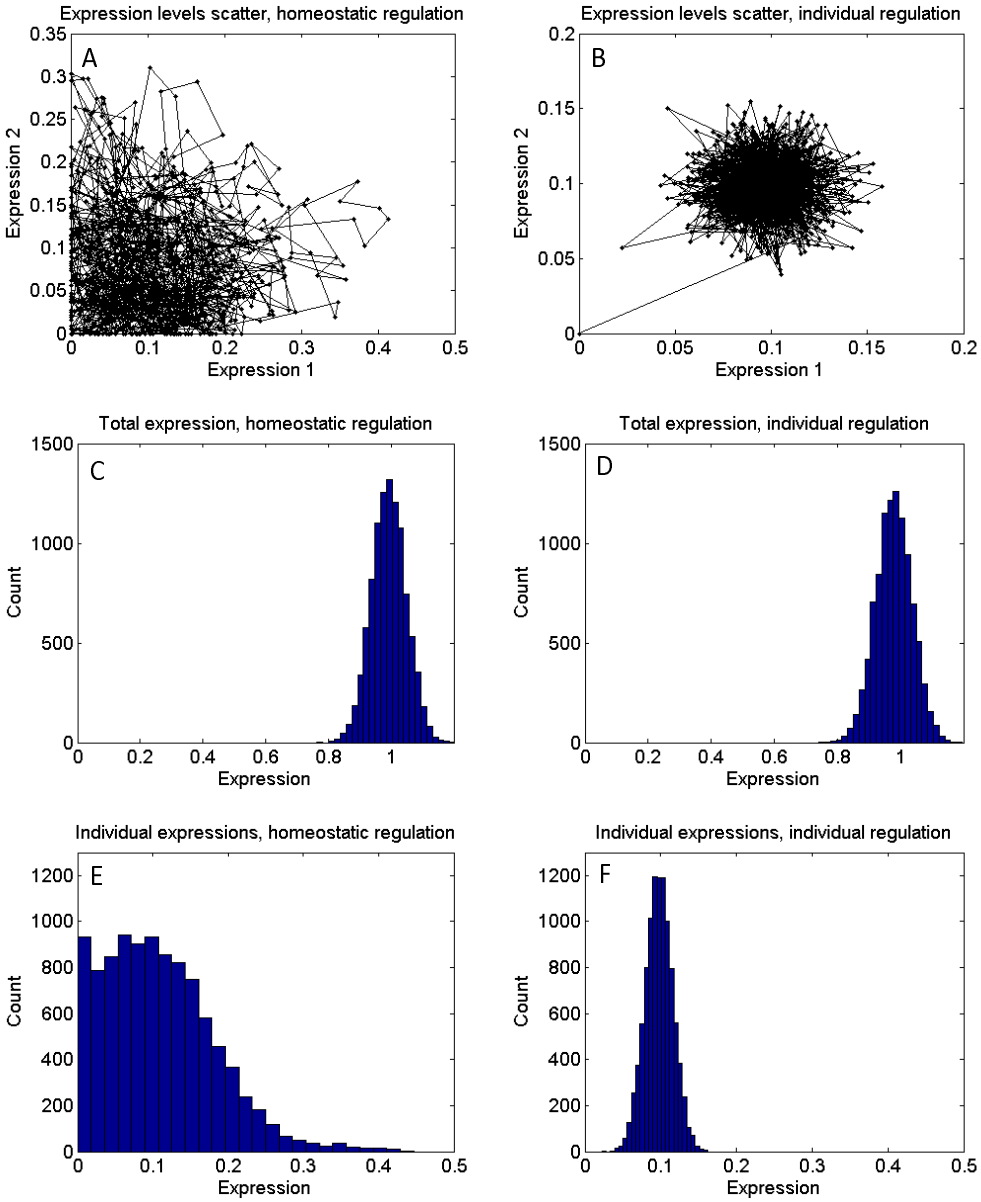

Figure 3: The stochastic variation in biological systems' internal configurations may reach dramatic levels when the number of internal degrees of freedom, that is, the pathways regulating the same physiological parameter, is large. A) Realized internal states (X_1,X_2) of the model system (1) under homeostatic regulation with *N*=10 regulation pathways. B) Realized internal states for the model system (1) with *N*=10 under individually controlled pathways. C-D) The variation of the final total protein concentration is the same under either homeostatic (C) or individual regulation (D). E-F) The spread in the expression levels of the individual pathways X_k under homeostatic regulation (E) is dramatically larger than that under individual regulation (F).