Summing up the noise in gene networks

Johan Paulsson

Department of Molecular Biology, Princeton University, Washington Road, Princeton, New Jersey 08544-1014, USA, and Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, Wilberforce Road, University of Cambridge, Cambridge CB3 0WA, UK

Random fluctuations in genetic networks are inevitable as chemical reactions are probabilistic and many genes, RNAs and proteins are present in low numbers per cell. Such 'noise' affects all life processes and has recently been measured using green fluorescent protein (GFP). Two studies show that negative feedback suppresses noise, and three others identify the sources of noise in gene expression. Here I critically analyse these studies and present a simple equation that unifies and extends both the mathematical and biological perspectives.

Intracellular randomness has long been predicted from basic physical principles[1] and observations of phenotypic heterogeneity[2,3]. In the last few years it has also been visualized directly using fluorescent probes. The first quantitative studies[4–8] collectively examined the noise associated with the principal steps of the central dogma of molecular biology; that is, replication, gene activation, transcription, translation and the enslaving intracellular environment. They also suggested how autorepression of replication and transcription suppresses noise, and how eukaryotes differ from prokaryotes. This analysis connects the different studies to a simple variant of the fluctuation-dissipation theorem and uses the experimental controls to extend or reinterpret many of the conclusions.

Theory

The fluctuation-dissipation theorem

All cell components display intrinsic noise due to random births and deaths of individual molecules, and extrinsic noise due to fluctuations in reaction rates. This has been modelled extensively using both detailed computer simulations and analytically tractable Markov processes, each tailored to a particular system. Here, I take a different approach and see how far a first approximation of a simple but generic model goes in explaining previous theory and experiments. The probabilities of having n_1 and n_2 molecules per cell of chemical species X_1 and X_2 (for instance messenger RNAs and proteins, but interpretations vary with application) are described by a birth-and-death Markov process with events

$$n_1 \xrightarrow{R_1^{\pm}(n_1)} n_1 \pm 1 \quad \text{and} \quad n_2 \xrightarrow{R_2^{\pm}(n_1,n_2)} n_2 \pm 1$$

Because n_1 affects rate R_2 but n_2 does not affect R_1, this is an example of dynamic disorder[9,10] where species X_1 provides the randomly fluctuating environment for X_2, as mRNA fluctuations randomize protein synthesis. To collectively approximate all such processes I use the Ω-expansion[11,12] where the first- and second-order terms reproduce the macroscopic rate equations and the fluctuation-dissipation theorem[13,14] respectively. The latter is then interpreted in terms of the logarithmic gains $H_{ij} = \partial \ln(R_i^-/R_i^+)/\partial \ln(n_j)$ that measure how the balance between production and elimination of X_i is affected by X_j. These scale-free parameters are closely related to the elasticities of metabolic control analysis[15,16] and the apparent kinetic orders of biochemical systems theory[17,18], and can often be estimated directly from the reaction rates. For instance, if R_i^+ and R_i^- are of first and second kinetic order in n_j respectively, then $H_{ij} = 2 - 1 = 1$. For the process described above, using σ_i for standard deviations, $\langle n_i \rangle$ for averages and τ_i for average lifetimes, stationary fluctuations around a stable fixed point follow

$$\frac{\sigma_2^2}{\langle n_2 \rangle^2} \approx \overbrace{\frac{1}{\langle n_2 \rangle H_{22}}}^{\text{Intrinsic } n_2 \text{ noise}} + \overbrace{\underbrace{\frac{\sigma_1^2}{\langle n_1 \rangle^2}}_{n_1 \text{ noise}} \underbrace{\frac{H_{21}^2}{H_{22}^2}}_{\text{Susceptibility}} \underbrace{\frac{H_{22}/\tau_2}{H_{11}/\tau_1 + H_{22}/\tau_2}}_{\text{Time-averaging}}}^{\text{Extrinsic } n_2 \text{ noise}} \quad (1)$$

where $\sigma_1^2/\langle n_1 \rangle^2 \approx (\langle n_1 \rangle H_{11})^{-1}$. Intrinsic noise depends on the average number of molecules and how systematic adjustments (rate H_{22}/τ_2) quench spontaneous fluctuations (rate $1/\tau_2$). The normalized adjustment rate H_{22} can also be interpreted as the statistical bias to return to the average rather than deviate further: a 1% increase in n_2 gives a H_{22} per cent increase in R_2^-/R_2^+. Extrinsic noise instead depends on the magnitude of n_1 fluctuations and how strongly n_1 affects n_2. The normalized susceptibility factor H_{21}/H_{22} reflects that a 1% increase in n_1 gives a H_{21} per cent increase in R_2^-/R_2^+, which makes n_2 adjust towards a H_{21}/H_{22} per cent lower average quasi-steady-state. When n_1 changes rapidly (high H_{11}/τ_1) or n_2 adjusts slowly (low H_{22}/τ_2), n_2 does not have time to reach its quasi-steady-state before n_1 changes anew. Consecutive ups and downs in n_1 then cancel out and n_2 time-averages over the recent history of n_1 fluctuations. The effect of cell growth and division is qualitatively accounted for by adding first-order elimination terms to R_1^- or R_2^-. The method behind equation (1) can be extended to any chemical system, providing a basis for a stochastic Biochemical Systems Theory (J.P., manuscript in preparation).

Noise in the central dogma

The basic principles of noise in genetic networks can be understood by applying equation (1) to the central dogma of molecular biology. Unregulated replication combined with first-order elimination is dynamically unstable as DNA acts as a template for its own synthesis: if each molecule on average replicates more or less than once per cell cycle, the average concentration increases until resources become limiting or decreases until all templates are gone. At intermediate replication frequencies, where synthesis and elimination are delicately balanced, random fluctuations still accumulate in an almost unrestrained fashion. With X_2 as self-replicators, this has been modelled[19] similarly to equation (1) with $H_{22} \ll 1$, and the same principles apply to dynamic instability of microtubules[20] and many other cell processes[12,19,21]. Unregulated transcription and translation are not unstable when combined with first-order elimination, as mRNAs and proteins are not templates for their own synthesis. Over the last 35 years, models of stochastic gene expression[19,22–31] have instead focused on dynamically disordered linear processes, corresponding to $H_{11} = H_{22} = -H_{21} = 1$ in equation (1) with X_2 as proteins and X_1 as genes[27,28], mRNAs[22–26] or the intracellular environment[19]. To

suppress the noise in any of these systems, cells commonly use autorepression that increases and decreases synthesis at low and high concentrations respectively. This has been studied extensively using macroscopic models[17,32], and the stochastic principles are closely related. Autorepression can raise the effective H_{22}, and thus: (1) increase the adjustment rate H_{22}/τ_2 relative to the rate $1/\tau_2$ of spontaneous randomization and thereby suppress intrinsic noise around a given average number of molecules; (2) increase the adjustment rate H_{22}/τ_2 relative to the rate H_{11}/τ_1 of environmental changes and thereby amplify extrinsic noise by preventing time-averaging; (3) decrease the susceptibility H_{21}/H_{22} of the quasi-steady-state, which typically overcompensates for the impaired time-averaging and produces a net decrease in extrinsic noise. This may explain the popularity of autorepression in transcription networks[32–34] and its ubiquity in replication control of chromosomes and plasmids where it has been similarly described[5,19]. Equation (1) thus unifies models of autoreplication, constitutive transcription and translation, and the stabilizing effect of autorepression, all in disordered environments. This makes it ideally suited also to unify the GFP studies that examine these aspects experimentally.

Terminology and measures

Comparing different systems requires consistency in definitions and measures. The terms 'intrinsic' and 'extrinsic' generically distinguish between the origin and propagation of noise, and their biological meaning is always defined in conjunction with a specified component or process. For example, if a gene for a transcriptional repressor spontaneously switches on and off, thereby enslaving the encoded protein and transmitting the fluctuations to repressed genes, the noise is intrinsic to the number of active repressor genes and extrinsic to all affected components. If the corresponding noise in a repressed protein was instead assigned to its transcription, just because transcription transmits the noise from the repressor gene, then by the same logic it must also be assigned to translation or to any other step in the cascade. This relates directly to the mathematical measures. For intrinsic noise it is convenient to use $\sigma_2^2/\langle n_2\rangle$ for a size-independent comparison, but this artificially forces extrinsic noise to increase with $\langle n_2\rangle$, as in $\sigma_2^2/\langle n_2\rangle \approx H_{22}^{-1} + E\langle n_2\rangle$ where E is the second term in equation (1). Unless the measure matches the noise, scale artefacts may thus completely distort interpretations of dynamics. For instance, if all protein (X_2) noise came from fluctuations in protease levels, $\sigma_2^2/\langle n_2\rangle^2$ would typically be independent of transcription and translation rates. But because both of these processes affect $\langle n_2\rangle$, measuring noise strength by $\sigma_2^2/\langle n_2\rangle$ would make it appear as if they contributed noise, even though they have nothing to do with either its production or transmission. Fortunately, these scaling principles also provide an opportunity to trace the noise: increasing the number of repressor genes in the example above should reduce the relative standard deviations of the repressed proteins, whereas other changes in transcription or translation should have little effect. This approach must be used carefully as the susceptibilities or time constants may change as well, but it at least provides useful indications.

Autorepression

Two groups have used plasmid-expressed *gfp* to examine, for the first times, how autorepression of replication[4] and transcription[5] affects noise levels. The replication study used GFP as a reporter for gene dosage, and compared natural and synthetic plasmids that have intact and impaired autorepression respectively. The transcription study instead used plasmids as cloning vectors to study protein fluctuations, and engineered an autorepression loop by placing a fusion of GFP and the tetracycline repressor (TetR) downstream of a TetR-repressed promoter. Both studies showed that autorepression substantially reduces relative standard deviations, which has been suggested to come from the more rapid adjustments to steady state in plasmid[19] and protein[5] concentrations respectively. But even if rapid adjustments reduce intrinsic noise around a given average, the effect is the opposite for extrinsic noise (compare points (1) and (2) above). This reveals an interesting discrepancy between the studies: if the noise came from gene expression, GFP would not measure plasmid copy numbers, and if it came from fluctuations in plasmid copy numbers, a protein that adjusted more rapidly to its meandering steady state would inherit more noise, not less. Thanks to the many experimental controls, this issue can be at least partially settled using equation (1).

Replication

With X_2 as plasmids, the simplest models of plasmid copy number control[35] assume $R_2^+ \propto n_2(K + n_2^h)^{-1}$. Factor n_2 reflects that each copy can self-replicate, and factor $(K + n_2^h)^{-1}$ reflects autorepression, for instance approximating the effect of short-lived plasmid-encoded inhibitors[19] that act cooperatively with effective Hill coefficient h. The statistics of plasmid segregation in growing cells depends on the accuracy of partitioning plasmids at cell division, but the qualitative effect is similar to first-order elimination[19], $R_2^- \propto n_2$. When $K + n_2^h$ is insensitive to n_2, so that autoreplication is poorly constrained by autorepression, both R_2^- and R_2^+ are thus approximately proportional to n_2 and plasmids are almost unstable, $H_{22} \ll 1$. This explains the large copy number variation observed for plasmids with impaired replication control[4].

The random environment X_1 could represent DNA polymerase, ribosomes, chaperones or any other factor that affects the replication control system. But because equation (1) is generic, the qualitative effects of these hypothetical fluctuations can be understood even if the kinetic details are left unspecified. As in points (1) to (3) in the 'Noise in the central dogma' section, autorepression should suppress intrinsic plasmid noise by expediting normalized plasmid adjustments, and suppress extrinsic plasmid noise by decreasing the net influence of environmental variation. The six natural plasmids studied[4] displayed similar $\sigma_2/\langle n_2\rangle$ despite large differences in $\langle n_2\rangle$—typically the signature of extrinsic noise. However, low- and high-copy plasmids are different altogether, and the former are thought to compensate by more efficient partitioning and replication control. Some of the similarities also reflect that GFP is an inexact reporter of single-cell plasmid copy numbers (Supplementary Information). This makes it impossible to say how the noise is checked by autorepression, but because plasmids are present in low numbers per cell, most studies[19] have focused on intrinsic noise and normalized adjustment rates.

Transcription

With X_2 instead representing proteins, constitutive transcription (R_2^+ independent of n_2) combined with first-order elimination ($R_2^- \propto n_2$) is far from unstable ($H_{22} = 1$). Because GFP was also present in high numbers per cell ($\langle n_2\rangle \gg 1$), intrinsic protein noise should thus be negligible, as confirmed by experimental controls[5] where $\sigma_2/\langle n_2\rangle$ was unaffected by large changes in $\langle n_2\rangle$. Similarly, experiments[5] where $\sigma_2/\langle n_2\rangle$ was unaffected by 20-fold changes in total gene activity and mRNA levels suggest that spontaneous small-number fluctuations in these components do not contribute any substantial protein noise, although they may still transmit fluctuations from plasmids, ribosomes or RNAses. The transcriptional autorepression loop was described[5] by $R_2^+ \propto K(K + n_2)^{-1} \approx Kn_2^{-1}$, which would stabilize the system twofold ($H_{22} \approx 2$). The protein now affects the number of active genes and mRNAs, but because these components were not the source of the extrinsic noise, equation (1) still applies (R_1 is unaffected by n_2) and predicts points (2) to (3) in the 'Noise in the central dogma' section with the environment X_1 left unspecified. Rapid protein adjustments thus increase noise by preventing time-averaging, and the observed decrease instead probably comes from an overcompensating reduction in the susceptibility of the steady state to changes in some extrinsic factor. For instance, cells that by chance have more

plasmid copies accumulate more protein, but transcriptional autorepression then partially compensates by reducing the expression per gene copy. Experimental controls[5] measuring the average protein level $\langle n_2\rangle$ as a function of the average plasmid level $\langle n_1\rangle$ suggest an expected $\langle n_2\rangle \propto \langle n_1\rangle$ and $\langle n_2\rangle \propto \sqrt{\langle n_1\rangle}$ for the unregulated and autoregulated system respectively, corresponding to $H_{21}^2/H_{22}^2 = 1$ and $H_{21}^2/H_{22}^2 \approx 1/4$ in equation (1). This does not prove that all protein noise comes from fluctuations in plasmid copy numbers—the results are also consistent with a decreased susceptibility to ribosomes or proteases—but plasmids are likely contributors because of their low copy numbers and because their fluctuations are slow and difficult to time-average. Many natural plasmids in fact use transcriptional autorepression[36] to ensure that their replication proteins are unsusceptible to perturbations in both plasmid copy numbers and the intracellular environment. The transcription study[5] is the first direct assessment of this principle.

Chromosomal gene expression

Plasmid-expressed *gfp* does not reflect the randomness of chromosomal gene expression because transcription and translation are averaged over more gene copies and plasmids themselves fluctuate randomly. The first group[6] to identify the sources of noise in gene expression therefore cloned *gfp* into the chromosome of *Bacillus subtilis* and measured single-cell fluorescence at varying rates of transcription and translation. The result of that study is consistent with a long-standing hypothesis[19,22–26] that protein fluctuations depend on the burst b of proteins (henceforth denoted X_2) made per mRNA transcript, $\sigma_2^2/\langle n_2\rangle \approx 1 + \langle b\rangle$. A second group[7] instead inserted two types of *gfp* into the *Escherichia coli* chromosome and used correlations between them to infer where the fluctuations came from. Finally, a third group revisited the one-gene strategy for *Saccharomyces cerevisiae* and again reported $\sigma_2^2/\langle n_2\rangle$ as a function of transcription and translation rates. They suggested that eukaryotes differ from prokaryotes because promoter fluctuations and transcriptional reinitiation produce a non-monotonous transcription noise.

Transcription and translation noise in *B. subtilis*

With X_1 and X_2 as mRNAs and proteins, and k_1 and k_2 as the transcription and translation rates, the model behind the *B. subtilis* experiments[6,26] assumes $R_1^+ = k_1, R_1^- = n_1/\tau_1, R_2^+ = k_2 n_1$ and $R_2^- = n_2/\tau_2$, so that $H_{11} = H_{22} = -H_{21} = 1$ in equation (1). Further assuming $\tau_1 \ll \tau_2$ gives $\sigma_2^2/\langle n_2\rangle^2 \approx \langle n_2\rangle^{-1} + \langle n_1\rangle^{-1}\tau_1/\tau_2$, and multiplying by $\langle n_2\rangle = \langle n_1\rangle k_2\tau_2$ reproduces the burst prediction $\sigma_2^2/\langle n_2\rangle \approx 1 + k_2\tau_1 = 1 + \langle b\rangle$. The experiments confirmed that $\sigma_2^2/\langle n_2\rangle$ is independent of k_1 and increases linearly with k_2, which is quite remarkable as there are no free parameters and they tested a real prediction rather than making *ad hoc* explanations. This suggested that noise is determined translationally and that strongly translated genes are implicated in particularly noisy processes[6]. However, the burst term $\langle b\rangle$ in $\sigma_2^2/\langle n_2\rangle$ does not come from the randomness of translation, but from random births and deaths of mRNA transcripts. It would disappear completely if transcription and mRNA degradation were deterministic, and would remain unchanged if translation was deterministic. By construction, $\sigma_2^2/\langle n_2\rangle$ is thus independent of the contribution from random translation events, and using the scale-free $\sigma_2^2/\langle n_2\rangle^2$ provides a quite different perspective[19,31]: a higher translation rate has a marginal effect because it only reduces the intrinsic protein noise $\langle n_2\rangle^{-1}$, which should be small throughout the experiments, whereas a higher transcription rate additionally reduces the much greater noise contribution from having few mRNA transcripts, $\sigma_1^2/\langle n_1\rangle^2 = \langle n_1\rangle^{-1}$.

The effect of other random influences was not estimated, and noise was measured at varying rates—tracking changes rather than absolute values—to draw conclusions independently of such effects[6]. But extrinsic processes would typically make superimposable contributions to $\sigma_2^2/\langle n_2\rangle^2$ rather than to $\sigma_2^2/\langle n_2\rangle$, and using $\sigma_2^2/\langle n_2\rangle$ can then be misleading (see 'Terminology and measures' section). In particular, if any substantial noise changed as transcription and translation were varied, total noise should behave in a more complicated manner and the experiments would be hard to interpret at all. If it instead remained constant, it could be calculated from the data. The acid test is how $\sigma_2^2/\langle n_2\rangle^2$ appears to vary with $\langle n_2\rangle^{-1}$ in the transcription experiments. The mRNA–protein model above predicts a perfect proportionality, a constant extrinsic noise adds a displacement, and a varying extrinsic noise creates a distortion. Replotted this way, the data fit well to a straight line with a small displacement, which suggests that most noise indeed comes from having few transcripts per cell, at least at low rates of transcription. This strengthens the authors' model[6,26], although from a different biological perspective, and shows that their data additionally account for other possible sources of noise. Noise in promoter activity or fluctuations in the transcriptional repressor should also have less effect at higher transcriptional induction, and could in principle produce the same result; however, there is then no reason to expect a quantitative fit without fine-tuned parameters (Supplementary Information).

Intrinsic and extrinsic noise in *E. coli*

The kinetic analysis[31] accompanying the two-gene *E. coli* experiments uses separate models for the noise that is intrinsic and extrinsic to the expression of an individual gene. The intrinsic part is similar to the mRNA–protein model above, and the extrinsic part is equivalent to the second term in equation (1) with arbitrary $\sigma_2^2/\langle n_2\rangle^2$ and H_{21}, but $H_{11}/\tau_1 = 0$. The difference reflects that equation (1) is derived from stationary stochastic processes rather than fixed probability distributions, thereby extending the analysis from static ($H_{11}/\tau_1 = 0$) to dynamic ($H_{11}/\tau_1 > 0$) disorder. This is biologically relevant as GFP degrades so slowly that few cell processes are static in comparison.

Experimentally, Elowitz *et al.*[7] measured the concurrent expression of two identically regulated but fluorescently distinguishable *gfp* genes in the *E. coli* chromosome. Spontaneous small-number mRNA fluctuations then generate uncorrelated fluctuations in the two GFPs as each allele makes its own transcript, whereas noise in RNA polymerase levels perturbs both GFPs in the same way. To quantify these data without relying on any particular kinetic model, Swain *et al.*[31] described generic protein fluctuations in terms of distributions of underlying (static) variables that were either intrinsic (for instance mRNA) or extrinsic (for instance RNA polymerase) to the expression of an individual gene. This use of 'intrinsic' and 'extrinsic' makes a non-trivial distinction between system and environment, but is qualitatively supported by equation (1) where the second term can be interpreted as the normalized covariance $\sigma_{23}/(\langle n_2\rangle\langle n_3\rangle)$ between two identical and independent components X_2 and X_3 embedded in the same environment X_1 (Supplementary Information). Nonlinear mechanisms would violate the requirement that the GFPs are independent (Supplementary Information), but that was not a problem in the experiments, and this clever strategy then directly separated the noise that is intrinsic and extrinsic to the expression of an individual gene. Their intrinsic contribution behaved almost like the total noise in the one-gene *B. subtilis* study and was similarly explained in terms of small-number mRNA fluctuations. However, their extrinsic contribution from the intracellular environment was by contrast substantial or even dominant. This is a central discrepancy. If most noise comes from the low numbers of mRNAs or proteins, each gene is under individual selection for high or low noise production, but if it comes from the intracellular environment, numerous genes are affected similarly regardless of transcription and translation rates. Discrepancies can always be ascribed to differences in experimental conditions, but in this case the devil seems rather to be in the genetic details: most noise extrinsic to gene expression in the two-gene study was traced back[7,31] to the transcriptional repressor LacI that was gradually inactivated by isopropyl-β-D-thiogalactoside to

vary the transcription rate. The two-gene study used wild-type, oscillating or plasmid-borne *lacI*, all of which are expected to be noisier than the highly expressed chromosomal *lacI* of the one-gene study. This more complex behaviour of the noise underscores the importance of using two GFPs, especially as $\sigma_2^2/\langle n_2\rangle^2$ was a non-monotonic function of transcription.

Transcription and translation noise in *S. cerevisiae*

The detailed computer simulations accompanying the *S. cerevisiae* study[8] account for numerous steps in eukaryotic gene expression, and the results perfectly match experiments where noise strength, as measured by $\sigma_2^2/\langle n_2\rangle$, responded non-monotonically to transcriptional induction and increased linearly with translation rate. The translation response was much stronger at partially induced rather than maximally induced transcription. This suggested that, in addition to translation noise, eukaryote-specific chromatin remodelling and transcriptional reinitiation produce quantile transcription bursts. However, because $\sigma_2^2/\langle n_2\rangle$ introduces a scale-dependence, the experiments should be re-evaluated in terms of $\sigma_2^2/\langle n_2\rangle^2$. The published transcription data then show a monotonically decreasing $\sigma_2^2/\langle n_2\rangle^2$, but the authors' full data set is non-monotonic and displays an initial increase in $\sigma_2^2/\langle n_2\rangle^2$ at low induction (W. Blake, personal communication). The translation experiments basically reiterate this result: the linear increase in $\sigma_2^2/\langle n_2\rangle$ indicates an extrinsic source of noise, and the stronger translation response was obtained at intermediate transcriptional induction where $\sigma_2^2/\langle n_2\rangle^2$ was larger. This is still consistent with slow chromatin remodelling and transcriptional reinitiation, so their interpretation may be correct, but both monotonic and non-monotonic responses in either $\sigma_2^2/\langle n_2\rangle$ or $\sigma_2^2/\langle n_2\rangle^2$ are also perfectly consistent with a variety of other noise sources, such as repressors[31] and activators or their random association and dissociation to DNA (Supplementary Information). Other types of experiments are thus needed to distinguish between the many alternatives, and at this point there are no reasons to favour one interpretation over another. So far, no principal differences between protein fluctuations in pro- and eukaryotes have been demonstrated. The claim that there is a difference was based on a comparison with the *B. subtilis* experiments[6] that used the same methods and interpretations, whereas the *E. coli* experiments[7] used other methods and interpretations but reported almost identical results. Instead, a comparison of the studies shows how dependent protein noise is on rate constants and upstream inputs; that is, differences in network design have overshadowed organismic differences. Notably, all of them indicate very little noise from, for example, RNA polymerase or ribosomes that would affect gene expression on a global level. In the one-gene experiments, fully induced genes displayed relative protein variation as low as 15% in both *B. subtilis* and *S. cerevisiae*, and much of this may reflect cell-size variation and measurement errors (Supplementary Information). In the two-gene *E. coli* $lacI^-$ experiments, the noise extrinsic to gene expression was so small (5%) that it cannot be convincingly distinguished from zero.

Outlook

Noise in genetic and metabolic networks can be detrimental to fitness—randomizing developmental pathways, disrupting cell cycle control or forcing metabolites away from their optimal levels. It can also be exploited for non-genetic individuality or even for more reliable and deterministic control[25,37,38]. But even if the pervasive influence of noise has been recognized for decades, it is only recently that quantitative measurements have become feasible. A few pioneering groups have now measured stationary fluctuations over populations of cells, and the results so far are consistent with a few basic principles. The next challenge is to monitor stochastic signals and responses in real time and expose the true cell dynamics buried in population averages. □

Received 8 September; accepted 2 December 2003; doi:10.1038/nature02257.

1. Schroedinger, E. *What is Life?* (Cambridge Univ. Press, Cambridge, 1944).
2. Novick, A. & Weiner, M. Enzyme induction as an all-or-none phenomenon. *Proc. Natl Acad. Sci. USA* **43,** 553–566 (1957).
3. Maloney, P. C. & Rotman, B. Distribution of suboptimally induced β-D-galactosidase in *Escherichia coli*. The enzyme content of individual cells. *J. Mol. Biol* **73,** 77–91 (1973).
4. Løbner-Olesen, A. Distribution of minichromosomes in individual *Escherichia coli* cells: implications for replication control. *EMBO J.* **18,** 1712–1721 (1999).
5. Becskei, A. & Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* **405,** 590–593 (2000).
6. Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nature Genet.* **31,** 69–73 (2002).
7. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297,** 183–186 (2002).
8. Blake, W. J., Kærn, M., Cantor, C. R. & Collins, J. J. Noise in eukaryotic gene expression. *Nature* **422,** 633–637 (2003).
9. Xie, X. S. Single-molecule approach to dispersed kinetics and dynamic disorder: Probing conformational fluctuation and enzymatic dynamics. *J. Chem. Phys.* **117,** 11024–11032 (2002).
10. Zwanzig, R. *Nonequilibrium Statistical Mechanics* (Oxford Univ. Press, New York, 2001).
11. van Kampen, N. G. *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).
12. Elf, J., Paulsson, J., Berg, O. G. & Ehrenberg, M. Near-critical phenomena in intracellular metabolite pools. *Biophys. J.* **84,** 154–170 (2003).
13. Keizer, K. *Statistical Thermodynamics of Nonequilibrium Processes* (Springer, Berlin, 1987).
14. Elf, J. & Ehrenberg, M. Fast evaluations of fluctuations in biochemical networks with the linear noise approximation. *Genome Res.* **13,** 2475–2484 (2003).
15. Kacser, H. & Burns, J. A. The control of flux. *Symp. Soc. Exp. Biol.* **27,** 65–104 (1973).
16. Westerhoff, H. V. & Van Dam, K. *Thermodynamics and Control of Biological Free Energy Transduction* (Elsevier, Amsterdam, 1987).
17. Savageau, M. A. *Biochemical Systems Analysis* (Addison-Wesley, Reading, 1976).
18. Savageau, M. A. Parameter sensitivity as a criterion for evaluating and comparing the performance of biochemical systems. *Nature* **229,** 542–544 (1971).
19. Paulsson, J. & Ehrenberg, M. Noise in a minimal regulatory network: plasmid copy number control. *Q. Rev. Biophys.* **34,** 1–59 (2001).
20. Dogterom, M. & Leibler, S. Physical aspects of the growth and regulation of microtubule structures. *Phys. Rev. Lett.* **70,** 1347–1350 (1993).
21. Berg, O. G., Paulsson, J. & Ehrenberg, M. Fluctuations and quality of control in biological cells: zero-order ultrasensitivity reinvestigated. *Biophys. J.* **79,** 1228–1236 (2000).
22. Rigney, D. R. & Schieve, W. C. Stochastic model of linear, continuous protein-synthesis in bacterial populations. *J. Theor. Biol.* **69,** 761–766 (1977).
23. Berg, O. G. A model for statistical fluctuations of protein numbers in a microbial-population. *J. Theor. Biol.* **173,** 307–320 (1978).
24. McAdams, H. H. & Arkin, A. Stochastic mechanisms in gene expression. *Proc. Natl Acad. Sci. USA* **94,** 814–819 (1997).
25. Paulsson, J., Berg, O. G. & Ehrenberg, M. Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl Acad. Sci. USA* **97,** 7148–7153 (2000).
26. Thattai, M. & van Oudenaarden, A. Intrinsic noise in gene regulatory networks. *Proc. Natl Acad. Sci. USA* **98,** 8614–8619 (2001).
27. Peccoud, J. & Ycart, B. Markovian modelling of gene-product synthesis. *Theor. Popul. Biol.* **48,** 222–234 (1995).
28. Kepler, T. B. & Elston, T. C. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.* **81,** 3116–3136 (2001).
29. Sasai, M. & Wolynes, P. G. Stochastic gene expression as a many-body problem. *Proc. Natl Acad. Sci. USA* **100,** 2374–2379 (2003).
30. Singh, U. N. Polyribosomes and unstable messenger RNA: a stochastic model of protein synthesis. *J. Theor. Biol.* **25,** 444–460 (1969).
31. Swain, P. S., Elowitz, M. B. & Siggia, E. D. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl Acad. Sci. USA* **99,** 12795–12800 (2002).
32. Savageau, M. A. Comparison of classical and autogenous systems of regulation in inducible operons. *Nature* **252,** 546–549 (1974).
33. Thieffry, D., Huerta, A. M., Perez-Rueda, E. & Collado-Vides, J. From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in *Escherichia coli*. *Bioessays* **20,** 433–440 (1998).
34. Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genet.* **31,** 64–68 (2002).
35. Paulsson, J. Multileveled selection on plasmid replication. *Genetics* **161,** 1373–1384 (2002).
36. Chattoraj, D. K. Control of plasmid DNA replication by iterons: no longer paradoxical. *Mol. Microbiol.* **37,** 467–476 (2000).
37. Vilar, J. M., Kueh, H. Y., Barkai, N. & Leibler, S. Mechanisms of noise-resistance in genetic oscillators. *Proc. Natl Acad. Sci. USA* **99,** 5988–5992 (2002).
38. Howard, M. & Rutenberg, A. D. Pattern formation inside bacteria: fluctuations due to the low copy number of proteins. *Phys. Rev. Lett.* **90,** 128102 (2003).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements This work was supported by a Lewis-Thomas fellowship from Princeton University and by the Cambridge-MIT Institute. I thank A. Becskei, E. Cox, M. Ehrenberg, J. Elf, M. Elowitz, I. Golding, S. Leibler, A. Løbner-Olesen, C. Peterson, S. Sawai, P. Swain, M. Thattai and A. van Oudenaarden for discussions and comments on the manuscript. I also thank W. Blake, M. Kaern, C. Cantor and J. Collins for sharing unpublished data.

Competing interests statement The author declares that he has no competing financial interests.

Correspondence and requests for materials should be addressed to J.P. (J.Paulsson@damtp.cam.ac.uk).