Published in final edited form as:

Anal Bioanal Chem. 2009 January ; 393(1): 73–80. doi:10.1007/s00216-008-2431-z.

Determining biological noise via single cell analysis

Edgar A. Arriaga

Department of Chemistry, University of Minnesota, Minneapolis, MN 55455, USA

arriaga@umn.edu

Abstract

Single cell analysis techniques describe the cellular heterogeneity that originates from fundamental stochastic variations in each of the molecular processes underlying cell function. The quantitative description of this set of variations is called biological noise and includes intrinsic and extrinsic noise. The former refers to stochastic variations directly involved with a given process, while the latter is due to environmental factors associated with other processes. Mathematical models are successful in predicting noise trends in simple biological systems, but it takes single cell techniques such as flow cytometry and time lapse microscopy to determine and dissect biological noise. This review describes several approaches that have been successfully used to describe biological noise.

Keywords

Biological noise; Extrinsic noise; Gene regulation function; Flow cytometry; Intrinsic noise; Single cell analysis; Time lapse microscopy

Biological noise

Biological systems are intrinsically heterogeneous. Underlying their heterogeneity are stochastic fluctuations of each molecular process (e.g., the production of RNA molecules from DNA transcription, the production of peptides from RNA translation, the degradation of RNA molecules, etc.) needed for cell function. The quantitative description of such fluctuations is termed biological noise [1,2].

Biological noise is divided into intrinsic and extrinsic noise [3]. Intrinsic noise usually refers to the combination of stochastic variations associated with each of the molecular processes directly involved with the production of a gene product (i.e., a protein) or the biological function associated with the gene product (e.g., enzymatic activity). In contrast, extrinsic noise refers to the variations caused by other molecular processes such as cell-to-cell and cell-to-extracellular matrix contacts, mechanical stress, thermal stress, or variations in the concentrations of other surrounding molecules (e.g., regulators, inhibitors, or metabolites) that do not directly participate in gene expression, but have the potential to modify gene expression via molecular interactions. It is the combined effects of intrinsic and extrinsic noise that lead to cellular heterogeneity or phenotypic variation.

The emergence of single cell analysis techniques has made possible the description of heterogeneity (e.g., phenotypic variations) in biological systems. Comprehensive reviews on single cell analysis have recently appeared [4–8], and some examples of single cell techniques are listed in Table 1. However, the emphasis of this review is on those single cell techniques that have been used as tools for investigating biological noise. More specifically, this review provides examples of how mathematical models, flow cytometry and time-lapse fluorescence microscopy have been used to describe biological noise.

Defining biological noise

An analysis of biological noise begins by taking into account the stochastic variations associated with each of the processes along the path of genomic information flow, a.k. a. gene regulation function (GRF) [9]. Thus, stochastic variations in a GRF will include those associated with (i) gene activation through the binding of a transcription factor to a promoter site in the DNA; (ii) DNA transcription into RNA; (iii) RNA translation into a peptide; (iv) peptide maturation into a functional protein; (v) mRNA degradation; (vi) protein degradation; (vii) molecular autoregulation; (viii) segregation due to cell division, and; (ix) incorrect transcription [3,10,11].

Ideally, the stochastic variation in the output from each of these processes is expected to have a Poisson distribution. For example, if the mean number of protein molecules is $\langle p \rangle$, the variance δp^2 equals $\langle p \rangle$ [3]. Thus,

$$v = \delta p^2 / \langle p \rangle \quad (1)$$

where v is known as the Fano factor and represents noise strength. Similarly, the coefficient of variation, $CV = \delta p / \langle p \rangle$, is also used to represent noise. For ideal Poisson behavior, $v = 1$ and $CV^2 = 1 / \langle p \rangle$.

As expected for Poisson distributions, molecular processes involving few molecules result in high noise levels. However, processes with large numbers of molecules do not necessarily have reduced noise levels [9,12], because extrinsic noise or global noise also contributes to the total noise in the GRF output. In fact, this is the main source of phenotypic variation in *E. coli* and *S. cerevisiae* [13,14].

In order to take the various noise sources into account, total biological noise is calculated as

$$CV_{\text{tot}}^2 = CV_{\text{in}}^2 + CV_{\text{ex}}^2 \quad (2)$$

where the subscripts “tot,” “in,” and “ex” in each coefficient of variation represent the total, intrinsic and extrinsic noise, respectively [12]. This expression makes it possible to determine the extrinsic noise after measuring the total biological noise (e.g., CV_{tot}) and the intrinsic noise (e.g., CV_{in}).

GRF outputs are usually time-dependent. It is the temporal variation in the GRF outputs that ultimately defines cell size, cell cycle frequency or status, developmental cell-fate decisions [15], the production of biomolecules such as mRNA [16], and the release of species such as Ca^{2+} [17]. The form of the GRF output varies among biological systems and they may appear as a steady function, a gradual change, on–off switching [15,18], bursts [19], oscillations [20,21], or circadian clocks [22,23]. Therefore, the noise associated with each GRF output is similarly time-dependent. Some of the examples listed below will demonstrate the importance of the time dependence of biological noise.

Examples of biological noise measurements

Most studies investigating biological noise have used either microscopy or flow cytometry to monitor variants of green fluorescent protein [12,24–26]. This review describes four approaches used to characterize biological noise. These approaches are: numeric simulations, the regulation of two reported genes by equivalent promoters, the simultaneous

monitoring of a gene regulator and the corresponding gene product, and high-throughput biological noise analysis.

Numeric simulations

Following earlier simulations [27–29], Thattai et al. used a numeric simulation to predict the production of proteins from the expression of the *lacZ* and *lacA* genes [3]. In this simulation, the authors incorporated reaction rates and the concentrations of species involved in transcription, translation, RNA degradation and protein degradation, and used kinetic constant values previously reported in the literature for the molecular processes. Their simulations predicted Fano factors for protein expression under conditions of steady state, nonequilibrium, autoregulation, and cell division. The reader is referred to the original article for a detailed explanation of the mathematical derivations that were needed [3]. Here, only key features are discussed.

The first simulation assumed a steady state number of mRNA molecules having a Poisson distribution with $\langle r \rangle = k_R/\gamma_R$ and $v = \delta^2/\langle r \rangle = 1$ (c.f. Eq. 1), where $\langle r \rangle$ is the average number of mRNA molecules, k_R is the mRNA production rate, and γ_R is the mRNA degradation rate. The simulation predicted bursts of protein expression every time an mRNA molecule is transcribed (indicated by the vertical arrows along the time axis in Fig. 1A). When the rate of protein degradation is much smaller than that of mRNA degradation, each burst has on average b protein molecules, with a predicted Fano factor larger than that of a Poisson distribution (i.e., $v \sim b + 1$; c.f. Eq. 1). For the genes *lacZ* and *lacA*, the bursts consist of 40 and 5 protein molecules and the predicted Fano factors are ~ 41 and ~ 6 , respectively.

Figure 1B shows the expected Fano factor (y-axis; c.f. Eq. 1) as a function of the average protein expression (x-axis) when all of the parameters are kept constant except for one of protein burst size (circles), transcription rates (inverted triangles), or protein half-life (squares). As seen from this graph, the transcription rates and protein degradation have constant noise contributions (i.e., $\sim$ constant Fano factors; y-axis) across most of the $\langle p \rangle$ range, while the noise contribution from the protein burst size (i.e., translation) increases as $\langle p \rangle$ increases, indicating that the noise in this system is translationally regulated.

The second and third simulations predicted protein production in nonequilibrium systems and in dividing cells, respectively. When the time-dependent degradation of a protein is taken into consideration (i.e., exponential decay after each burst in Fig. 1B), the Fano factor is $v \sim 2b + 1$, for short times. Thus, noise levels are higher than in the respective steady-state system [3]. During cell division, the Fano factor changes as a function of cell age. It is predicted that $v \sim 1.33b + 1$ just before and $v \sim 0.67b + 1$ just after cell division. Thus, noise levels tend to increase as the cell progresses through the cell cycle [3].

The fourth simulation presented by Thattai and co-workers described an autoregulatory system, which is a common motif in biological systems. In autoregulatory systems, the transcription initiation rate (k_R) decreases as the mean gene product (i.e., protein) concentration $\langle p \rangle$ increases. This system is usually represented by a nonlinear Hill function,

$$k_R = \frac{k_{R,\max}}{1 + \left(\frac{\langle p \rangle}{K_d}\right)^n} \quad (3)$$

where $k_{R,\max}$ is the uninhibited transcription initiation rate, $\langle p \rangle$ is the mean protein concentration, K_d is the dissociation constant between the protein and the interaction site in the transcription machinery, and n is the Hill coefficient. When $\langle p \rangle$ is low, (i.e., autoregulation is occurring), the Fano factor (c.f. Eq. 2) becomes

$$\nu = \left(\frac{1 - k_1/\gamma_p}{1 + bk_1/\gamma_p} \right) \left(\frac{b}{1 + \gamma_p/\gamma_R} \right) + 1 \quad (4)$$

where the burst of protein molecules per mRNA molecule is b (c.f. Fig. 1), the rate of mRNA degradation is γ_R , the rate of protein degradation is γ_p , and k_1 is inversely proportional to K_d . As an example of autoregulation, the authors simulated the CL repressor protein (i.e., gene product) acting on the promoter P_R (i.e., regulator site) of the bacteriophage λ , which has a K_d value of 50 nM. They also represented a similar system in which CL repressor protein acts on the promoter P_{RM} , which has a K_d value of 1000 nM. They convincingly demonstrated that the noise strength decreases as the K_d values decrease, pointing to the importance of autoregulation in noise reduction (c.f. Fig. 3 in [3]).

Regulation of two reporter genes by equivalent promoters

An example of this approach is the work by Elowitz et al. [12]. They used time correlation between the expression of two reporter proteins, cyan fluorescent protein (CFP) and yellow fluorescent protein (YFP), each controlled by separate equivalent promoter sites (i.e., two P_{LacO1} promoter sites located equidistantly in opposite directions from the origin of replication). LacI is a repressor that, when binding to this promoter, increases the production of both CFP and YFP. Therefore, by controlling the expression of LacI (e.g., by changing the concentration of isopropyl β -D-thiogalactopyranoside, which inactivates the expression of LacI), different levels of CFP and YFP expression were attained.

In the absence of intrinsic biological noise, it was expected that CFP and YFP expressions would track each other with time, even in the presence of extrinsic noise that results in temporal variations in protein abundance (Fig. 2A). On the other hand, when intrinsic noise is present, in addition to variations in protein amounts due to extrinsic noise, it was expected that the proportions of CFP and YFP would also vary with time, indicating a loss of time correlation (Fig. 2B). The predicted observations for such systems were that, in the absence and presence of intrinsic noise, bacteria expressing these proteins would all look the same color (right, Fig. 2A) or display color variations (right, Fig. 2B), respectively.

In the work by Elowitz and coworkers [12], expression of both CFP and YFP in *E. coli* cells was monitored by fluorescence microscopy, and computerized image analysis was used to create plots of normalized mean fluorescence intensities. As seen in Fig. 3A, a plot of normalized YFP versus normalized CFP values tends to describe a diagonal with a slope of 1. Deviations perpendicular to the diagonal represent intrinsic noise (i.e., changes in the proportions of CFP and YFP). Dispersion along the diagonal represents extrinsic noise (i.e., changes in the amounts of CFP and YFP). This approach worked well for twelve different examples using four different *E. coli* strains and different variations in promoter regulation (Table 2). In the various scenarios described in this work, the intrinsic noise ($\eta_{in} = CV_{in}$; e.g., red triangles in Fig. 3) decreased as the number of fluorescent molecules produced increased (Fig. 3B). On the other hand, the dominant extrinsic noise reached a maximum at an intermediate value of fluorescence intensity ($\eta_{ex} = CV_{ex}$ in Fig. 3B). It is believed that the observed maxima were dependent on cell-to-cell variations in the LacI concentration resulting from using a plasmid to express LacI [12].

In summary, Elowitz and coworkers found that the production of large and low amounts of CFP and YFP (e.g., in $LacI^-$ and Lac^- strains) had low and high intrinsic noise, respectively [12]. These results are in agreement with previous studies which indicate that intrinsic noise in these systems is a function of transcription rate, regulatory dynamics, and genetic factors [14].

Simultaneous monitoring of a gene regulator and the gene product

Rosenfeld et al. simultaneously monitored the amounts of CL repressor of the P_R promoter that controlled the expression of the CFP in *E. coli* lineages [9]. The amount of CL repressor could be monitored by fluorescence microscopy, because this repressor was expressed as the fusion protein CL-YFP construct. This fusion protein was not typically expressed, but when it was transiently induced, it repressed the P_R promoter and the production of CFP. Thus, following induction, the abundance of CL-YFP was high and CFP abundance was low. The investigators monitored cell division and tracked the abundances of CL-YFP and CFP in the cell lineages as a function of time. They observed that CL-YFP was diluted out as cells divided, which removed the repression of the P_R promoter, which in turn led to increased CFP (c.f. Fig. 1 in [9]).

The data nicely fitted Hill functions of the form $f(R) = \beta / [1 + (R/K_d)^n]$, where $f(R)$ is the protein (CFP) concentration, R is the repressor (CL-YFP) concentration, β is the maximal production rate, K_d is the repressor concentration that gives half-maximal expression, and n is the degree of effective cooperativity in the repression (c.f. Eq. 3). From the deviations from the fit, they determined that the total noise was high (i.e., $CV_{tot} \sim 0.55$) and that the CV_{tot} decreased slightly to ~ 0.4 after correcting for the cell cycle.

In order to further dissect the total noise, the authors carried out a separate experiment in which they regulated two identical genes by the same promoter [12]. This experiment used concepts similar to those outlined in Fig. 2 and made it possible to determine the intrinsic biological noise (i.e., $CV_{in} \sim 0.2$). From the total noise and the intrinsic noise (c.f. Eq. 2), they determined that extrinsic noise predominated in this system (i.e., $CV_{ex} \sim 0.350$). This extrinsic noise was attributed to differences in the micro-environment for each cell [30], and the various sources of noise caused by other cellular processes [12].

High-throughput biological noise analysis

Flow cytometry was used to investigate the architecture of biological noise in >2500 unique yeast proteins, each tagged with GFP to its c-terminus and expressed from its endogenous promoter and its natural chromosomal position [31]. This was the most comprehensive and systematic study of biological noise to date, and included: (i) the dissection of the observed fluctuations into the extrinsic and intrinsic noise of each yeast strain containing a unique protein; (ii) the identification of gene products deviating from the noise-abundance trends in the entire library of yeast strains, and; (iii) the correlation of deviant gene products in the library with Gene Ontology parameters that contribute to noise.

In flow cytometry, the total noise (i.e., CV_{tot}) can be determined from the widths of the fluorescence intensity histograms resulting from the analysis of cells expressing the GFP constructs. As shown in Fig. 4A, two different protein constructs may have similar mean fluorescence intensities but different CV_{tot} values, as shown in the inset. In order to identify the intrinsic and extrinsic noise, the investigators used the concept of gating in 2D plots of forward scattering (FSC), which is a measurement of size, and side scattering (SSC), which is a measurement of cellular granularity. In these plots, they determined the average FSC and SSC, and used these coordinates to define radial regions (gates) in the 2D scattering plots. These gates were then used to select the cells used in the GFP fluorescence intensity distributions (c.f. Fig. 4A). As the gate radius decreased (right to left in Fig. 4B), the CV decreased (black line) until it practically flattened out (indicated by the dashed vertical line). This flat CV value in Fig. 4B suggests that extrinsic noise (blue line) has basically been eliminated, and what remains is intrinsic noise (red line) and instrumental artifacts (gray line), which can be easily corrected. Both the intrinsic and instrumental noise are uncorrelated, and thus can be represented as flat lines in the CV versus gate radius plot. By

subtracting the instrument noise and the intrinsic noise from the total noise over the entire gate radius range, one obtains the extrinsic noise (blue line). The two examples in Fig. 4B demonstrate different levels of intrinsic noise (i.e., CV_{in}) associated with the His4 gene (left) and the Rps25A gene (right).

A plot of $\log_{10}(CV^2)$ versus $\log_{10}(\text{abundance})$ is shown in Fig. 4C for all of the GFP constructs that were investigated in this study [31]. In this figure, CV^2 shows no correlation with abundance for the ungated results. On the other hand, the use of the gate radii, as indicated by a vertical line in Fig. 4B, results in a plot in which CV^2 tends to be inversely proportional to protein abundance. In Fig. 4C, the running median (orange line) of low-abundance proteins (e.g., <1000) have a slope ~ -1 , which corresponds to a Fano factor of ~ 1 (c.f. Eq. 1). This is consistent with ideal Poisson behavior.

One unique feature of the study by Newman et al. is that it provides the intrinsic noise values (i.e., gated CVs) for a large number of gene products [31]. These investigators compared the distance of individual CV values from the mean CV value (DM) when the CVs are plotted as a function of protein abundance. As shown in Fig. 4D, gene products displaying an average trend were mainly associated with stochastic noise, while those displaying deviations from that trend suggest the involvement of other parameters, such as the regulation required to maintain low noise (e.g., transcription factors Flh1, Rap1, and Abf1), or high noise (e.g., factors SAGA, SWI/SNF, Ino80, Isw2, and Swr1).

Furthermore, by correlating intrinsic noise with Gene Ontology terms, Newman and coworkers found that proteins involved in the translation, proteolysis and acidification of vesicles in the secretory pathway are associated with low noise levels, suggesting that these processes are highly regulated [31]. Another particularly interesting observation was that proteins involved in oxidative phosphorylation, heat shock, stress, and amino acid synthesis are associated with high noise. The fact that these proteins are associated with organelles that are distributed to the two cells resulting from cell mitosis suggests that subcellular localization is another contributor to intrinsic biological noise.

Concluding remarks

The roles of biological noise in different biological systems cannot be underscored enough. For example, high biological noise may introduce diversity into populations, thus making their evolution possible under conditions of natural selection. In other instances, such as in embryonic development, low levels of or high tolerance to biological noise may be needed. The characterization of noise in such systems is only possible with suitable single-cell analysis techniques.

Flow cytometry and time lapse microscopy have made it possible to observe individual cell behavior and extract information about the origins of biological noise in yeast and *E. coli*. It is envisioned that other single cell techniques (c.f. Table 2) could use strategies similar to those described in this review to identify the intrinsic noise and extrinsic noise resulting from cellular microenvironments.

Acknowledgments

The author would like to thank Carol Makkyla and Marian Navratil for their insightful comments and suggestions. Support is from NIH grant R01-AG20866 and a Fesler Lampert Chair, University of Minnesota.

References

1. Auffray C, Imbeaud S, Roux-Rouquie M, Hood L. *Philos Trans Royal Soc London Ser A* 2003;361:1125–1139.
2. Blaabjerg O, Petersen PH, Horder M, Thygesen K, Feldtrasmussen U. *J Clin Chem Clin Biochem* 1981;19:617–617.
3. Thattai M, van Oudenaarden A. *Proc Nat Acad Sci USA* 2001;98:8614–8619. [PubMed: 11438714]
4. Arcibal IG, Santillo MF, Ewing AG. *Anal Bioanal Chem* 2007;387:51–57. [PubMed: 16912862]
5. Huang WH, Ai F, Wang ZL, Cheng JK. *J Chromatogr B* 2008;866:104–122.
6. Miyashiro, T.; Goulian, M. Single-cell analysis of gene expression by fluorescence microscopy.. In: Simon, MI.; Crane, BR.; Crane, A., editors. Two-component signaling systems, Pt B (Methods in Enzymology vol 423). Elsevier; Amsterdam: 2007. 2007. p. 458-475.
7. Meyer P, Dworkin J. *Res Microbiol* 2007;158:187–194. [PubMed: 17349779]
8. Wodicka L, Dong HL, Mittmann M, Ho MH, Lockhart DJ. *Nat Biotechnol* 1997;15:1359–1367. [PubMed: 9415887]
9. Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. *Science* 2005;307:1962–1965. [PubMed: 15790856]
10. Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. *Mol Syst Biol* 2007;3:143. [PubMed: 18004276]
11. Struhl K. *Nat Struct Mol Biol* 2007;14:103–105. [PubMed: 17277804]
12. Elowitz MB, Levine AJ, Siggia ED, Swain PS. *Science* 2002;297:1183–1186. [PubMed: 12183631]
13. Casadaban MJ. *J Mol Biol* 1976;104:541–555. [PubMed: 781293]
14. Raser JM, O'Shea EK. *Science* 2004;304:1811–1814. [PubMed: 15166317]
15. Ferrell JE, Machleder EM. *Science* 1998;280:895–898. [PubMed: 9572732]
16. Bucy RP, Panoskaltis-Mortari A, Huang GQ, Li J, Karr L, Ross M, Russell JH, Murphy KM, Weaver CT. *J Exp Med* 1994;180:1251–1262. [PubMed: 7523568]
17. Kuchtey J, Fewtrell C. *J Cell Physiol* 1996;166:643–652. [PubMed: 8600169]
18. Biggar SR, Crabtree GR. *Embo J* 2001;20:3167–3176. [PubMed: 11406593]
19. Lahav G, Rosenfeld N, Sigal A, Geva-Zatorsky N, Levine AJ, Elowitz MB, Alon U. *Nat Genet* 2004;36:147–150. [PubMed: 14730303]
20. Geva-Zatorsky N, Rosenfeld N, Itzkovitz S, Milo R, Sigal A, Dekel E, Yarnitzky T, Liron Y, Polak P, Lahav G, Alon U. *Mol Syst Biol* 2006;2:2006–2033. [PubMed: 16773083]
21. Klevecz RR, Li CM, Marcus I, Frankel PH. *FEBS J* 2008;275:2372–2384. [PubMed: 18410382]
22. Barkai N, Leibler S. *Nature* 2000;403:267–268. [PubMed: 10659837]
23. Zhou T, Chen L, Aihara K. *Phys Rev Lett* 2005;95:178103. [PubMed: 16383875]
24. Blake WJ, Kaern M, Cantor CR, Collins JJ. *Nature* 2003;422:633–637. [PubMed: 12687005]
25. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A. *Nat Genet* 2002;31:69–73. [PubMed: 11967532]
26. Becskei A, Kaufmann BB, van Oudenaarden A. *Nat Genet* 2005;37:937–944. [PubMed: 16086016]
27. Berg OG, Paulsson J, Ehrenberg M. *Biophys J* 2000;79:2944–2953. [PubMed: 11106602]
28. Cook DL, Gerber LN, Tapscott SJ. *Proc Nat Acad Sci USA* 1998;95:15641–15646. [PubMed: 9861023]
29. Ko MSH. *J Theoret Biol* 1991;153:181–194. [PubMed: 1787735]
30. Shapiro JA. *Ann Rev Microbiol* 1998;52:81–104. [PubMed: 9891794]
31. Newman JR, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, Weissman JS. *Nature* 2006;441:840–846. [PubMed: 16699522]
32. Springer M, Paulsson J. *Nature* 2006;439:27–28. [PubMed: 16397487]
33. Kuang Y, Biran I, Walt DR. *Anal Chem* 2004;76:6282–6286. [PubMed: 15516119]

34. Dixon CJ, Bowler WB, Walsh CA, Gallagher JA. *Br J Pharmacol* 1997;120:777–780. [PubMed: 9138681]
35. Huang L, Shen H, Atkinson MA, Kennedy RT. *Proc Natl Acad Sci USA* 1995;92:9608–9612. [PubMed: 7568182]
36. Schuster KC, Urlaub E, Gapes JR. *J Microbiol Methods* 2000;42:29–38. [PubMed: 11000428]
37. Hu S, Michels DA, Fazal MA, Ratisoontorn C, Cunningham ML, Dovichi NJ. *Anal Chem* 2004;76:4044–4049. [PubMed: 15253641]
38. Johnson RD, Navratil M, Poe BG, Xiong G, Olson KJ, Ahmadzadeh H, Andreyev D, Duffy CF, Arriaga EA. *Anal Bioanal Chem* 2007;387:107–118. [PubMed: 16937092]
39. Li L, Romanova EV, Rubakhin SS, Alexeeva V, Weiss KR, Vilim FS, Sweedler JV. *Anal Chem* 2000;72:3867–3874. [PubMed: 10959975]
40. Kamme F, Salunga R, Yu J, Tran DT, Zhu J, Luo L, Bittner A, Guo HQ, Miller N, Wan J, Erlander M. *J Neurosci* 2003;23:3607–3615. [PubMed: 12736331]
41. Peixoto A, Monteiro M, Rocha B, Veiga-Fernandes H. *Genome Res* 2004;14:1938–1947. [PubMed: 15466292]
42. Hoffmann A, Levchenko A, Scott ML, Baltimore D. *Science* 2002;298:1241–1245. [PubMed: 12424381]
43. Santoro MG. *Biochem Pharmacol* 2000;59:55–63. [PubMed: 10605935]
44. Alcedo J, Zou Y, Noll M. *Mol Cell* 2000;6:457–465. [PubMed: 10983991]
45. Guttman-Raviv N, Martin S, Kassir Y. *Mol Cell Biol* 2002;22:2047–2056. [PubMed: 11884593]

Fig. 1A–B.

Numerical simulation of protein production as a function of mRNA translation. **A** Each transcription initiation event (indicated by an *arrow*) causes the production of a burst of proteins of average size b ; the proteins decay between bursts. **B** Fano factor ($\delta p^2 / \langle p \rangle$) as a function of mean protein abundance $\langle p \rangle$ resulting from varying only the protein burst size (*circles*), transcript initiation rate (*triangles*), or the protein half-life (*squares*). When not varied, the mRNA half life is 2 min; the average protein burst size is 20; the transcript initiation rate is 0.01 s^{-1} , and the protein half life is 1 h. Other conditions of the model are described in [3]; copyright National Academy of Sciences, USA

Fig. 2A–B.

Evaluation of biological noise using YFP and CFP expression regulated by equivalent regulatory sequences. The abundances of both proteins are monitored by their respective fluorescence (shown as *red and green traces on the left*, respectively) versus time; the relative abundances of YFP and CFP determine the observed color (*right panels*). **A** In a hypothetical case in which there is no intrinsic noise, the relative abundances of YFP and CFP are constant (i.e., the same color on the *right panel*), while their individual abundances fluctuate as a result of extrinsic noise. **B** When intrinsic noise is present, it causes changes to the relative abundance of the two proteins as a function of time. At a given time, the proportion of the two proteins will cause a change in the observed color (*right panel*). From [12], reprinted with the permission of the AAAS

Fig. 3A–B.

Quantification of noise in a dual gene expression system regulated by equivalent promoters. **A** Plot of normalized mean fluorescence intensities for the YFP reporter (y-axis) versus the CFP reporter (x-axis). Each marker represents either one M22 (*inverted blue triangle*) or one D22 (*green triangle*) *E. coli* cell. Intrinsic noise is calculated from the deviations perpendicular to a diagonal with slope 1, while extrinsic noise is represented by the dispersion along the diagonal. The total noise is calculated using Eq. 2. **B** The total (*blue diamond*), intrinsic (*red triangle*), and extrinsic (*green inverted triangle*) noise levels in D22 *E. coli* as a function of relative fluorescence. The coefficient of variation (CV) is represented by η . As expected, the intrinsic noise decreases as the relative fluorescence increases. Surprisingly, the extrinsic noise, which dominates, has a maximum at 0.3 relative fluorescence. From [12], reprinted with the permission of the AAAS

Fig. 4A–D.

Using flow cytometry to determine the nature of biological noise in yeast. **A** Example of two flow cytometry results measuring two separate genes expressing GFP-fusion products with similar protein abundances (x-axis) but different distribution widths (i.e., CVs). The strains expressing Rp135A-GFP (*red*) and Nop8-GFP (*blue*) have CVs of 11.8 and 38.6, respectively. After sorting and reanalysis, their CVs remain unchanged (see *inset*), suggesting that flow cytometry is a suitable technique for investigating biological noise. **B** Dissecting biological noise based on the use of gated forward and side scattering measurements in flow cytometry. The gate, defined as the circular region in a forward versus side scattering plot centered on the average forward and side scattering measurements, is used to select the cells used to determine the CVs of the fluorescence measurements (see **A**). As the gate is decreased, the CVs of the fluorescence distributions decrease until they plateau (*vertical dashed line*). This CV represents the intrinsic noise (*red trace*) after correction for instrumental noise (*gray trace*). The extrinsic noise (*blue trace*) is then calculated using Eq. 2. The *left and right panels* correspond to His4 and Rps25A genes, respectively. **C** Global noise trends in the expression of all of the yeast genes included in this study. Without gating (*top panel*), the running median (*orange line*) shows that there is little dependence of total noise [represented as $\log_{10}(\text{CV}^2)$] as a function of protein abundance [represented as $\log_{10}(\text{abundance})$]. On the other hand, the intrinsic noise (i.e., when gating is used) is inversely related to protein abundance. In fact, for low-abundance proteins, the slope is ~ -1 , which is in agreement with Poissonian behavior (c.f. Eq. 1). **D** Deviations of CVs from the running median in part **C** are represented as DMs and plotted versus the square root of protein abundance. The plot shows that some biological processes (e.g., ATP synthesis and the TCA cycle) are noisy. From [31], reprinted with the permission of Macmillan Publishers Ltd.

Table 1

Examples of single cell analysis techniques

Technique	Single cell measurement	Cellular system	References
Flow cytometry	Noise in abundances of GFP-fusion proteins	<i>Sacharomyces cerevisiae</i>	[31]
Optical well arrays	Kinetics of <i>recA</i> and <i>lacZ</i> gene expression using GFP reporter	<i>Escherichia coli</i>	[33]
Fluorescence microscopy	Intracellular calcium release to identify subpopulations differing in expression of P2 receptors	Human osteoblasts	[34]
Electrochemical detection	Time profile of bursts of insulin secretion	Rat and human pancreatic beta cells	[35]
Raman microspectroscopy	Changes in Raman spectra indicating coexisting cell types	<i>Clostridium beijerinckii</i> in an acetone–butanol fermentation reactor	[36]
CE-LIF biomolecules	Two-dimensional separation of proteins that are fluorescently labeled on-line	Cultured MC3T3-E1 osteoprogenitor and MCF-7 breast cancer cells	[37]
CE-LFI organelles	Separation and detection of mitochondria labeled with DsRed2	Cultured 143B osteosarcoma cells	[38]
MALDI-MS	Identification of neuropeptides	Neuron cells isolated from <i>Aplysia Californica</i>	[39]
LCM and cDNA microarray analysis	Gene expression profiling pointing to two subpopulations	CA1 neurons from rat hippocampus	[40]
Multiplexed, real-time RT-PCR	Quantification of 20 different mRNAs	Human small intestine cells selected by FACS	[41]

GFP, green fluorescent protein; *CE-LIF*, capillary electrophoresis with laser-induced fluorescence detection; *MALDI-MS*, matrix-assisted laser desorption/ionization mass spectrometry; *LCM*, laser capture microdissection; *RT-PCR*, reverse transcriptase polymerase chain reaction; *FACS*, fluorescence-activated cell sorting

Table 2 Model systems and strategies used to investigate biological noise

Molecules	System features	Biological noise	Reference
LacZ product	Mathematical model, single gene	$v \sim 41^a$	[3]
		$v \sim 81^b$	
LacA product	Mathematical model, single gene	$v \sim 6^a$	[3]
		$v \sim 11^b$	
CL repressor	Autoregulation system of P_R promoter ($K_d=50$ nM)	$v \sim 4^a$	[3]
CL repressor	Autoregulation system of P_{RM} promoter ($K_d=1000$ nM)	$v \sim 6^a$	[3]
Ara product	On–off response	$-c$	[3]
Lac product	On–off response	$-c$	[3]
Promoter P_R , phage λ	Simultaneous monitoring of two gene products under the same regulation in <i>E. coli</i>	$\langle mi \rangle \langle /mi \rangle$ $\langle mi \rangle \langle /mi \rangle$	[9]
P_L lacO1	Simultaneous monitoring of two gene products under the same regulation in <i>E. coli</i>	$\langle mi \rangle \langle /mi \rangle$ $\langle mi \rangle \langle /mi \rangle$	[12]
p53–Mdm2	Simultaneous monitoring of protein and suppressor in MCF7 and H1299 cells	$-c$	[19]
NfκB–IκB	Simultaneous monitoring of protein and suppressor in mice stem cells	$-c$	[42]
HSF1–hsp70 or hsp90	Simultaneous monitoring of protein and suppressor in mammalian cells	$-c$	[43]
Smo–Ptc	Simultaneous monitoring of protein and suppressor in <i>Drosophila</i> embryos	$-c$	[44]
Ime1–Ime2 σ_H –dnaK or dnaJ	Simultaneous monitoring of protein and suppressor in yeast	$-c$	[45]
>2500 proteins	Flow cytometry, selection of yeast cells defining intrinsic noise based on scattering	$CV_{tot} \sim 0.05–0.45$	[31]

^a Intrinsic noise, steady state, $v \sim b + 1$

^b Intrinsic noise, nonequilibrium, $v \sim 2b + 1$

^c Not reported