Predictable trends in protein noise

Kristin Baetz & Mads Kærn

The process of gene expression is inherently stochastic and leads to differences in protein abundance from one cell to another. A new study shows that this protein noise is unexpectedly predictable, providing important new insights into the properties and origins of variability in gene expression.

© 2006 Nature Publishing Group http://www.nature.com/naturegenetics

Lately, there has been a lot of noise about protein noise—the variability among identical cells in the number of protein molecules for a given gene[1–3]. Some studies trace protein noise to sources that are intrinsic to the biochemical process of gene expression, such as turnover ('birth and death') of individual mRNA molecules (translational bursting) and slow promoter kinetics (transcriptional bursting). Others have shown dominant contributions from extrinsic processes, such as fluctuations in upstream signaling and variation in cell-cycle position and cell size. On page 636 of this issue, Bar-Even *et al.*[4] take a broader perspective on a noisy subject and discover that rowdy genes hum a surprisingly similar tune.

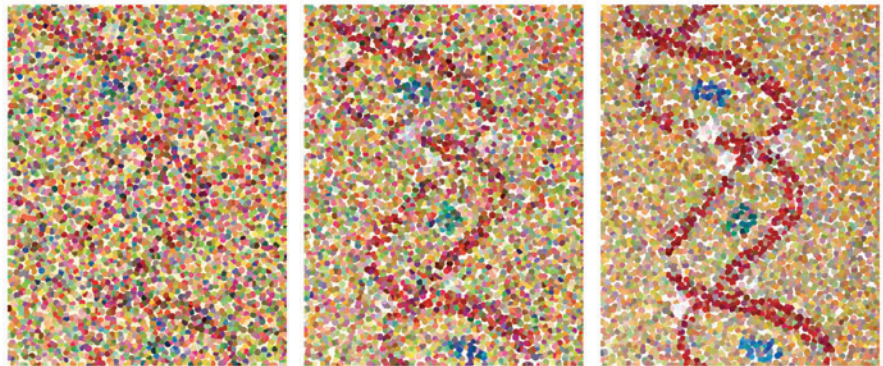

Figure 1 Finding order within disorder.

Scaling behavior

Bar-Even *et al.* determined the protein noise and mean protein abundance (population-averaged number of molecules for a given protein) for 38 different GFP fusion proteins across 11 different conditions. When they plotted protein noise against mean abundance for each fusion protein, they observed a general trend irrespective of the protein fusion, promoter and growth conditions. In most cases, they found that protein noise was remarkably close to inversely proportional to mean protein abundance, with a proportionality factor of ~1,200. This very surprising finding suggests that protein noise correlates in a characteristic manner with mean protein abundance. Protein noise can in this sense be predicted quantitatively. Interestingly, genes associated with stress response deviated from this common trend, with protein noise being significantly higher than that observed for the other genes analyzed.

Analyzing genome-scale mRNA and protein data sets, the authors further determined that the proportionality factor of ~1,200 is consistent with the number of proteins produced per mRNA. This suggests that translational bursting (fluctuation in the level of mRNA) is a major source of protein noise that could explain the observed scaling behavior. Although the study cannot exclude a role for transcriptional bursting (slow promoter kinetics), it clarifies that sources intrinsic to the biochemical process of gene expression are important in generating protein noise and cell-to-cell variability.

Kristin Baetz and Mads Kærn are in the Ottawa Institute of Systems Biology, University of Ottawa, 451 Smyth Road, Ottawa, Ontario K1H 8M5, Canada.
e-mail: kbaetz@uottawa.ca and mkaern@uottawa.ca

Biological implications

Previous studies in yeast have demonstrated that the level of protein noise from reporter genes is sensitive to gene and promoter sequence[2,3], fluctuations in upstream signaling[5,6] and the physical location of the reporter construct within the genome[7] and that it is affected when genes encoding chromatin remodeling complexes are deleted[8]. There is no *a priori* reason to believe that these different sources of protein noise should have the same effect across the genome and thus give rise to the common trend between protein noise and mean protein abundance observed by Bar-Even *et al.* This raises two fundamental questions. First, is the common trend 'scale-free' and applicable to the entire range of mean protein abundance? Second, how does the common trend arise? Although our current knowledge is too limited to provide definitive answers, below we try to shed some light on the issues involved.

Is protein noise scale-free? There are several indications that different sources of protein noise dominate at different ranges of protein abundance. Bar-Even *et al.* observe that the common trend is not extendable to high protein abundance, where it gradually becomes overshadowed by population effects such as variations in cell size and cell cycle position, consistent with previous studies[9]. For high protein abundance, the protein noise is observed to be minimal and at a level that is largely independent of the mean abundance. In other words, a second common trend—namely, constant low protein noise—is observed for proteins of high abundance.

It is questionable if the observed common trend in protein noise extends to low protein abundances. Extrapolating the trend from the range of intermediate abundance to a mean protein abundance of 200 molecules yields a standard deviation of about 500 molecules per cell. This standard deviation appears to be quite high and would seem to imply that cells should have a high chance of having few or no copies of essential proteins known to be present at very low abundance, such as cell cycle regulators, transcription factors and proteins involved in chromosome segregation. It is possible that genes encoding low-abundance proteins have properties leading to lower protein noise than genes encoding intermediate-abundance

proteins. Alternatively, cells may have evolved as-yet-uncharacterized mechanisms that allow reliable function despite high noise in key regulatory proteins. One could speculate that such a mechanism might involve the assembly of protein complexes that provide proteolytic protection, coupled with active partitioning of these complexes at cell division. In this context, the investigation by Bar-Even *et al.* highlights that we have much to learn about how cellular processes involving low-abundance molecules can achieve the functional robustness seen in nature.

Why a common trend? Before the study by Bar-Even *et al.*, there was no reason to believe that protein noise should be the same for proteins of similar abundance. This would require that the number of proteins synthesized per mRNA be approximately the same for different genes (that is, ~1,200 proteins per mRNA), if contributions from other sources of protein noise were minimal. It is possible that protein noise has been minimized for the majority of genes used in the study, as they encode proteins with essential housekeeping functions in proteasome complexes, ergosterol biosynthesis and rRNA processing. The biological reasoning behind this noise reduction hypothesis is as follows: high noise in essential proteins increases the chances that cells will lose the protein and hence mimic a lethal null mutant. Hence, high noise in essential proteins should be subject to evolutionary pressure[10]. The Bar-Even study confirms that this is likely the case. The authors observed that dispensable stress genes are associated with higher protein noise than housekeeping genes. This noise reduction was especially apparent for the essential components of the proteasome.

Even though protein noise is reduced for the proteasome, ergosterol and rRNA processing proteins, it is still unclear why they have a similar proportionality constant of ~1,200. If translational bursting is the dominant source of protein noise, it should be possible to reduce protein noise by decreasing the number of proteins synthesized per mRNA below the estimated value of ~1,200 proteins per mRNA. A twofold increase in mRNA abundance and a corresponding decrease in translational efficiency should yield a further twofold decrease in protein noise without changing protein abundance. Although such adjustments could, in principle, be used to virtually eliminate protein noise originating from mRNA fluctuations, the energetic cost would be prohibitive. It is possible that the observed proportionality constant arises from an evolutionary cost-benefit optimization of protein noise to cellular energy. However, as we are only beginning to understand the biological roles of protein noise, further studies of beneficial and detrimental noise effects on cell function will be needed to resolve this question.

1. Rao C.V., Wolf D.M. & Arkin A.P. *Nature* **420**, 231–237 (2002).
2. Kaern M., Elston T.C., Blake W.J. & Collins J.J. *Nat. Rev. Genet.* **6**, 451–464 (2005).
3. Raser, J.M. & O'Shea, E.K. *Science* **309**, 2010–2013 (2005).
4. Bar-Even, A. *et al. Nat. Genet.* **38**, 636–643 (2006).
5. Blake, W.J., Kaern, M., Cantor, C.R. & Collins, J.J. *Nature* **422**, 633–637 (2003).
6. Colman-Lerner, A. *et al. Nature* **437**, 699–706 (2005).
7. Becskei, A., Kaufmann, B.B. & van Oudenaarden, A. *Nat. Genet.* **37**, 937–944 (2005).
8. Raser, J.M. & O'Shea, E.K. *Science* **304**, 1811–1814 (2004).
9. Volfson, D. *et al. Nature* **439**, 861–864 (2006).
10. Fraser, H.B., Hirsh, A.E., Giaever, G., Kumm, J. & Eisen, M.B. *PLoS Biol.* **2**, e137 (2004).

© 2006 Nature Publishing Group http://www.nature.com/naturegenetics