When Noisy Neighbors Are a Blessing: Analysis of Gene Expression Noise Identifies Coregulated Genes

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
When Noisy Neighbors Are a Blessing: Analysis of Gene Expression Noise Identifies Coregulated Genes

Jan Philipp Junker¹ and Alexander van Oudenaarden¹,²,*
¹Departments of Physics and Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
²Hubrecht Institute–KNAW (Royal Netherlands Academy of Arts and Sciences) and University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, Netherlands
*Correspondence: avano@mit.edu
DOI 10.1016/j.molcel.2012.02.002

In this issue of Molecular Cell, Stewart-Ornstein et al. (2012) use systematic pair-wise correlation analysis of expression noise in a large number of yeast genes to identify clusters of functionally related genes and signaling pathways responsible for elevated noise.

Gene regulatory networks are pervasive entities in biology, playing a central role in cellular information processing and acting on a variety of levels, from the sensory networks that control a yeast cell’s reaction to osmotic shock to the developmental networks that determine the formation of an embryo from a fertilized egg (Aion, 2007). In many cases, large numbers of genes are placed under the control of the same upstream regulator (Davidson, 2010). But how can we systematically find out which genes are part of such a “regulon”? The usual approach consists of perturbing the upstream regulator and observing the effects on downstream gene expression. However, two major barriers can limit the efficacy of this method: (1) a priori knowledge of the upstream regulator is required, and (2) it can be very challenging to control the strength of the perturbation, particularly in the case of genetic manipulation. Perturbations should be strong enough to reliably yield observable effects, yet weak enough to avoid nonspecific side effects on gene expression. Wouldn’t it be faster and simpler to exploit the inherent noisiness of biological systems (Balázsi et al., 2011) rather than applying an exogenous perturbation to identify coregulated genes? Stewart-Ornstein et al. (2012) report a new method to identify clusters of genes that share a common upstream regulator based on single-cell correlations between pairs of genes in unperturbed yeast cells.

The underlying idea is simple: If the expression of an upstream regulator is noisy, its variability should be transmitted to all downstream targets (Figure 1A). If the regulator signal is high in a given cell, all downstream genes should be highly expressed. Conversely, if the regulator signal is low in another cell, target genes will also be lowly expressed. Hence, all genes within a regulon should be correlated among each other, but not with genes outside the regulatory network (Figure 1B). By flow-cytometry analysis of pair-wise correlations of GFP and mCherry fluorescently tagged proteins in unperturbed yeast cells, Stewart-Ornstein et al. were able to demonstrate that this is the case.

In the first part of their work, the authors decompose gene expression noise into intrinsic (gene specific, uncorrelated) and extrinsic (multiple genes, correlated) components, by comparing noise levels in diploid yeast cells expressing either one or two copies of the same fluorescently tagged protein (Elowitz et al., 2002; Becskei et al., 2005; Volfson et al., 2006). While intrinsic noise depends on the mean expression level in agreement with previous work, they find that extrinsic noise varies significantly between genes. Most genes exhibit a basal level of extrinsic noise, but a subset of about 20% of the genes studied display elevated extrinsic noise. Among those, the authors find enrichment of genes harboring specific transcription factor-binding sites, with binding sites of the stress responsive transcription factors Msn2/4 being most highly enriched. This observation suggests that the high extrinsic noise measured might originate from common upstream factors. To corroborate this hypothesis, Stewart-Ornstein et al. then examine the correlations between the Msn2/4 target Pgm2 and a large set of proteins. Indeed, they find that their query protein Pgm2 is most highly correlated to other Msn2/4 targets. In other words, the Msn2/4 targets fluctuate coherently, probably driven by variability in the Msn2/4 signal. Importantly, the authors also demonstrate that for a given gene the magnitude of the correlation to Pgm2 is predictive of its response to exogenous perturbations, establishing a direct link between their method and the traditional approach.

In the next step, the authors increase the number of pair-wise correlations in order to find additional noisy regulons. Correlating 182 genes to 44 query proteins (in addition to Pgm2), they find four major clusters, three of which correspond to well-defined functional categories (stress-response, mitochondrial, and amino acid biosynthesis proteins). Again, these groups of genes are enriched for specific transcription factor-binding sites. Next, the authors use genetic manipulation to trace the source of noise in the newly identified noise regulons. The PKA pathway is known to activate Msn2/4 (Görner et al., 1998) and is therefore a candidate noise source in the stress-response regulon. Indeed, heterozygous mutations in members of this pathway lead to changes in the correlation between two Msn2/4 target genes, indicating that PKA regulates noise in...
the system. Using a larger screen in which correlations between genes belonging to the mitochondrial and amino acid biosynthesis pathways are analyzed in strains heterozygous for several signaling-related factors, the authors further identify the TOR pathway as a likely source of noise for those genes.

These findings clearly demonstrate that expression noise in gene regulatory networks can be transmitted from a master upstream regulator to its downstream targets and that this process can be exploited to identify regulatory networks. But what is the biological advantage of having such “communities” of noisy neighbors? Stochastic gene expression in unicellular organisms has been shown to enhance the fitness of a population in rapidly changing environmental conditions (Blake et al., 2006). Maintaining diversity in a population through stochastic gene expression can be a successful “bet-hedging” strategy, anticipating changes in the environment. It might be beneficial for functionally related genes to fluctuate coherently to optimize the efficacy of the bet-hedging scheme. In agreement with this hypothesis, it is interesting to note that the authors find the stress-response Msn2/4 pathway to be particularly noisy and that increased basal levels of Msn2 indeed increase survival upon heat shock.

What will be the next breakthrough in single-cell correlation analysis look like? Application of this approach to cells and tissues of higher eukaryotes, where many yet undiscovered regulons are likely to exist, would be a significant step forward. Measuring RNA by single-cell RNA-Seq, qPCR, or single-molecule in situ hybridization instead of fluorescently tagged proteins would restrict the analysis to transcriptional interactions but circumvent the need for time-consuming transgenesis. Moreover, correlation analysis should in principle also be possible in gene networks exhibiting lower levels of noise. An equally important advance will be to establish the hierarchy of genes in a cluster (Figure 1B) by observing how correlations and absolute levels change over time or upon a given genetic perturbation. In the next years, we will certainly witness a growing number of applications of single-cell gene correlation analysis in the fascinating field of gene regulatory networks.

REFERENCES


Figure 1. Networks and Correlations
(A) Gene expression noise is transmitted from an upstream regulator to a downstream target. Expression levels of the target in individual cells will depend on the intensity of the regulator signal in the respective cell.
(B) In a population of cells, genes belonging to the same regulatory network will be correlated among each other (green arrows) but not to genes controlled by other pathways (gray arrow).