

NEWS AND VIEWS

Translation's coming of age

Christine Vogel*

Center for Genomics and Systems Biology, New York University, New York, NY, USA

* Corresponding author. Center for Genomics and Systems Biology, New York University, 12 Waverly Place, New York, NY 10003, USA. Tel.: +1 212 998 3976; Fax: +1 212 995 4015; E-mail: cvogel@nyu.edu

Molecular Systems Biology 7: 498; published online 24 May 2011; doi:10.1038/msb.2011.33

This is an open-access article distributed under the terms of the Creative Commons Attribution Noncommercial Share Alike 3.0 Unported License, which allows readers to alter, transform, or build upon the article and then distribute the resulting work under the same or similar license to this one. The work must be attributed back to the original author and commercial use is not permitted without specific permission.

Proteins are the key players in our cells, and their biological concentrations are regulated through an intricate interaction of transcription, translation, as well as mRNA and protein degradation. In a work recently published in *Nature*, Schwanhausser *et al* (2011) report one of the first and most comprehensive characterizations of the complete sequence of events producing proteins in mammalian cells. The authors measured concentrations and degradation rates for >5000 mRNAs and proteins, respectively, and tied these data into a mathematical model that, in unprecedented completeness, describes the cellular dynamics that govern protein production. Their analysis confirms, extends, and generalizes what has been suspected before: 'transcription is only half the story' (Plotkin, 2010), and translational control is at least as important in determining the final concentrations of proteins (Figure 1) (reviewed in de Sousa Abreu *et al*, 2009; Vogel *et al*, 2010).

When wondering about the impact of a perturbation on gene expression, analysis of mRNA expression levels is often the method of choice. Microarrays present a simple and standardized method to assess mRNA concentrations and are complemented by recent next-generation sequencing technology. For many years and out of necessity, mRNA concentrations were taken as a proxy of protein concentrations, assuming reasonable concordance between mRNA and protein expression regulation. Meanwhile, mass spectrometry-based proteomics methods have matured enough to present a workable solution for direct measurements of protein concentrations. The technology can still not capture the entire mammalian proteome, i.e. >20 000 genes, and reliable, quantitative measurements of several thousands of protein concentrations require weeks of labor, but instrumentation and software are now available to routinely incorporate proteomics analysis into any gene-expression study.

This technology has opened the scope of traditional proteomics experiments, which for many years comprised the comparison of a perturbed condition with a control. While such comparisons deliver invaluable information on the qualitative nature of the protein response, they have also restricted the types of experiments that were designed.

Now, much faster and more sensitive mass spectrometers are available and enable systematic measurements of the cellular *dynamics*, i.e. the speed at which the proteome reacts. As a consequence, kinetic models of cell-scale responses can be built, allowing for the design of time-dependent experiments, and the estimation of genome-wide rates of transcription, translation, and molecule degradation.

Schwanhausser *et al*'s (2011) work is a landmark among such efforts to break into these new frontiers. The authors use next-generation sequencing, large-scale proteomics based on mass spectrometry, and mathematical modeling to define the dynamics of protein production for 5000 genes in mouse fibroblasts. The work represents the largest of such studies available today, covering close to a quarter of the mouse genome. The authors focused on unperturbed, exponentially growing cells. mRNA concentrations and concentration changes were measured by next-generation sequencing; mRNA degradation rates were estimated through labeling of newly synthesized mRNA with a nucleoside analog (Friedel *et al*, 2009). Conceptually similar, the authors use isotopically labeled proteins (Doherty *et al*, 2009) to estimate protein degradation over 13.5 h. While comparison of heavy and light peptides delivers information on changes in protein concentrations, the authors also estimate absolute protein concentrations from the collected mass spectra.

In a simple mathematical model, changes in mRNA concentrations over time can be described as a function of rates of transcription and mRNA degradation. Similarly, protein concentrations are determined by rates of translation and protein degradation. Given that concentrations and degradation rates had been experimentally measured, Schwanhausser *et al* used this model to estimate the remaining, unknown rates, i.e. those of transcription and translation. The authors also incorporated rates of cell division into the model and estimated the relative contribution of these processes to cellular energy usage.

Do not underestimate translation, as its role may even be bigger than that of transcription—says the model by Schwanhausser *et al* (Figure 1). The correlation between mRNA and protein concentrations in mouse fibroblasts is 0.37–0.41, which is largely consistent with previous studies (Vogel *et al*,

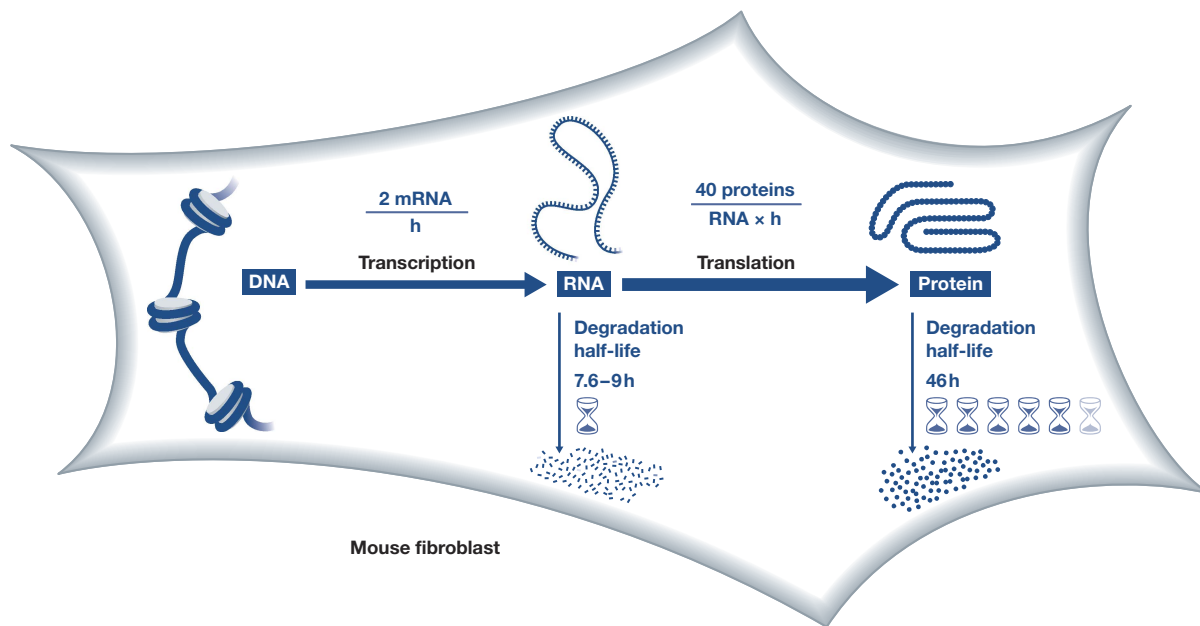


Figure 1 Francis Crick's central dogma of biology revolves around the transcription of mRNA from DNA, the translation of proteins from mRNA, and the degradation of the respective molecules. A new study by Schwanhausser *et al* (2011) measures concentrations of mRNAs and proteins as well as the corresponding degradation rates for >5000 genes in mouse cells, and with the help of a mathematical model they are able to deliver a complete description of the dynamics governing mammalian protein production. The thickness of the arrows indicates the relative importance of the processes in defining protein concentrations. The values indicate median rates or half-lives for those genes observed by Schwanhausser *et al*.

2010). This correlation implies that ~40% of the variation in protein concentrations is defined by mRNA concentrations, i.e. the results of transcription and mRNA degradation. Using their model, Schwanhausser *et al* show that another 41–54% of the variation in concentration across proteins can be attributed to differences in translation rates—while the rates of degradation have surprisingly small roles. The model holds up even when analyzing an entirely different cell line. In sum, the 'second half' of the central dogma of biology has a role much bigger than that has been recognized to date.

How surprising are these results? The mammalian genome encodes not only ~2000 transcription factors but also >600 RNA-binding proteins (de Lima Morais *et al*, 2011), presenting an enormous pool of potential post-transcriptional regulators. While discrepancies between mRNA and protein expression are known from many anecdotal examples, large-scale measurements of mRNA and protein concentrations usually also report low correlations between the two measures (de Sousa Abreu *et al*, 2009). Recent computational analyses have indicated that the correlation doubles when incorporating post-transcriptional regulation (Vogel *et al*, 2010). Further, several studies, including that by Schwanhausser *et al* (2011), report that protein concentrations cover a larger dynamic range (at least six orders of magnitude) than mRNA concentrations (about four orders of magnitude)—again suggesting substantial regulation following transcription. Finally, recent evolutionary analyses examining the concentrations of orthologs across different organisms reported that protein concentrations appear to be more conserved than mRNA concentrations (Schrimpf *et al*, 2009; Laurent *et al*, 2010). Together, a picture has formed that emphasizes the significance of post-transcriptional gene expression regulation—and

the study by Schwanhausser *et al* very timely demonstrates this for a mammalian system.

Such work raises entirely new questions. First of all, our understanding of the mechanisms underlying posttranscriptional regulation is still poor. It also remains to be seen how the contribution of translation regulation varies across cell types, and how it affects individual genes versus the entire proteome. We do not know how the differential contributions of transcription, translation, and degradation react to perturbations of the system—and whether the comparatively simple mathematical model will still hold true. Further, the measurements in Schwanhausser *et al*'s study are averages across an unsynchronized population of cells—single-molecule methods will have to complement the data. Nevertheless, Schwanhausser *et al*'s work highlights a shift in paradigm, a shift toward the routine incorporation of protein measurements, and a shift toward a holistic characterization of all cellular processes leading to functional proteins.

Conflict of interest

The author declares that he has no conflict of interest.

References

- de Lima Morais DA, Fang H, Rackham OJ, Wilson D, Pethica R, Chothia C, Gough J (2011) SUPERFAMILY 1.75 including a domain-centric gene ontology method. *Nucleic Acids Res* **39**: D427–D434
- de Sousa Abreu R, Penalva LO, Marcotte E, Vogel C (2009) Global signatures of protein and mRNA expression levels. *Mol Biosyst* **5**: 1512

- Doherty MK, Hammond DE, Clague MJ, Gaskell SJ, Beynon RJ (2009) Turnover of the human proteome: determination of protein intracellular stability by dynamic SILAC. *J Proteome Res* **8**: 104–112
- Friedel CC, Dolken L, Ruzsics Z, Koszinowski UH, Zimmer R (2009) Conserved principles of mammalian transcriptional regulation revealed by RNA half-life. *Nucleic Acids Res* **37**: e115
- Laurent J, Vogel C, Kwon T, Craig S, Boutz DR, Huse H, Nozue K, Walia H, Whiteley M, Ronald P, Marcotte EM (2010) Protein abundances are more conserved than mRNA abundances across diverse taxa. *Proteomics* **10**: 4209–4212
- Plotkin JB (2010) Transcriptional regulation is only half the story. *Mol Syst Biol* **6**: 406
- Schrimpf SP, Weiss M, Reiter L, Ahrens CH, Jovanovic M, Malmstrom J, Brunner E, Mohanty S, Lercher MJ, Hunziker PE, Aebersold R, von Mering C, Hengartner MO (2009) Comparative functional analysis of the *Caenorhabditis elegans* and *Drosophila melanogaster* proteomes. *PLoS Biol* **7**: e48
- Schwanhauser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M (2011) Global quantification of mammalian gene expression control. *Nature* **473**: 337–342
- Vogel C, de Sousa Abreu R, Ko D, Le SY, Shapiro BA, Burns SC, Sandhu D, Boutz DR, Marcotte EM, Penalva LO (2010) Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. *Mol Syst Biol* **6**: 400



Molecular Systems Biology is an open-access journal published by *European Molecular Biology Organization* and *Nature Publishing Group*. This work is licensed under a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License.