Molecular Cell
Technology Preview

Central Dogma Goes Digital

Yihan Lin[1] and Michael B. Elowitz[1,2,*]
[1]Division of Biology and Biological Engineering and Department of Applied Physics
[2]Howard Hughes Medical Institute
California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA
*Correspondence: melowitz@caltech.edu
http://dx.doi.org/10.1016/j.molcel.2016.03.005

In this issue of *Molecular Cell*, Tay and colleagues (Albayrak et al., 2016) describe a new technique to digitally quantify the numbers of protein and mRNA in the same mammalian cell, providing a new way to look at the central dogma of molecular biology.

The two gene expression processes involved in the central dogma of molecular biology—transcription and translation—are both influenced by fluctuations (i.e., "noise"). Thus, a clonal population of cells can be highly heterogeneous, with different cells having very different copy numbers of mRNA and protein. Such variability in gene expression can lead to diverse phenotypes and affect a wide range of cellular functions (Eldar and Elowitz, 2010), including development, homeostasis, and disease progression. Understanding the causes and consequences of variability in gene expression requires accurate quantification of mRNA and protein at the single-cell level. In this issue, Tay and colleagues (Albayrak et al., 2016) present an elegant and practical technique to digitally and simultaneously quantify the copy numbers of mRNA and protein in individual mammalian cells. This technique does not require genetic manipulations and can be easily adopted by other researchers. Albayrak et al. integrated this technique with mathematical modeling to explain the lack of correlation between the copy numbers of mRNA and protein for a single gene (Albayrak et al., 2016). In this case, they focused on CD147, a multifunctional transmembrane protein involved in various processes, including tumor invasion (Muramatsu and Miyauchi, 2003).

Counting mRNA molecules in single cells is increasingly commonly performed in laboratories. In contrast, counting protein molecules in single cells is much more challenging. Existing techniques require genetic manipulations or sophisticated devices and often only yield relative levels of protein abundance (Albayrak et al., 2016; Hughes et al., 2014; Shi et al., 2012; Taniguchi et al., 2010). To address these limitations, Albayrak et al. describe an elegant approach by combining proximity ligation assay (PLA) with droplet digital PCR (ddPCR) to count the absolute copy numbers of proteins in the lysate of single mammalian cells (Figure 1A).

As originally described by Landegren and colleagues (Fredriksson et al., 2002), PLA enables highly sensitive and specific protein detection by combining two antibodies (or nucleic acid aptamers) attached with DNA probes. Binding of both antibodies to the same protein brings the two DNA probes into close proximity. In the presence of a connector oligonucleotide, they can then be ligated together. The ligated DNA molecules can subsequently be quantified by techniques such as PCR. To achieve absolute quantification of protein abundance, Albayrak et al. analyzed the resulting DNA molecules with ddPCR on a commercially available platform (Albayrak et al., 2016) (Figure 1A). The resulting DNA counts can then be converted into protein counts based on a calibration curve, which was generated with a recombinant protein. The limit of detection of this new approach is about 10,000 proteins per cell, enabling quantification of medium- or high-abundance proteins.

To analyze the two gene expression processes involved in the central dogma—transcription and translation—Albayrak et al. performed digital quantification of both mRNA and protein in the same single cells (Albayrak et al., 2016) (Figure 1A). In this assay, the lysate of a single mammalian cell was split into two fractions, which were subjected to protein quantification and mRNA quantification, respectively. For mRNA, absolute quantification was accomplished by reverse transcription followed by ddPCR (with the limit of detection of about 52 mRNAs per cell).

After quantifying copy numbers of CD147 mRNA and protein in 131 human HEK293T cells, Albayrak et al. observed only very weak correlation between the copy numbers of mRNA and protein (Albayrak et al., 2016) (Figure 1B). This result suggests that the amount of CD147 mRNA is a poor proxy for the level of CD147 protein at the same time in a HEK293T cell.

In fact, low instantaneous correlations between mRNA and protein copy numbers at the single-cell level have been observed in different organisms. They were reported previously in *E. coli* (Taniguchi et al., 2010) and more recently in mammalian cells (Darmanis et al., 2016). Mechanistically, as suggested by Taniguchi et al. (Taniguchi et al., 2010), such low correlations may originate from at least two sources: first, disparities between the half-lives of mRNA and protein can cause protein concentrations to effectively lag their mRNAs. Second, extrinsic noise in protein translation (such as variability in levels of translational machinery) can cause the amount of protein produced per mRNA molecule to vary from cell to cell. In *E. coli*, where typical mRNA and protein half-lives are ~5 min and ~180 min, respectively, the former source was suggested to explain a large fraction of the low correlation (Taniguchi et al., 2010).

To understand the causes of the low correlations they observed, Albayrak et al. constructed a two-state model to

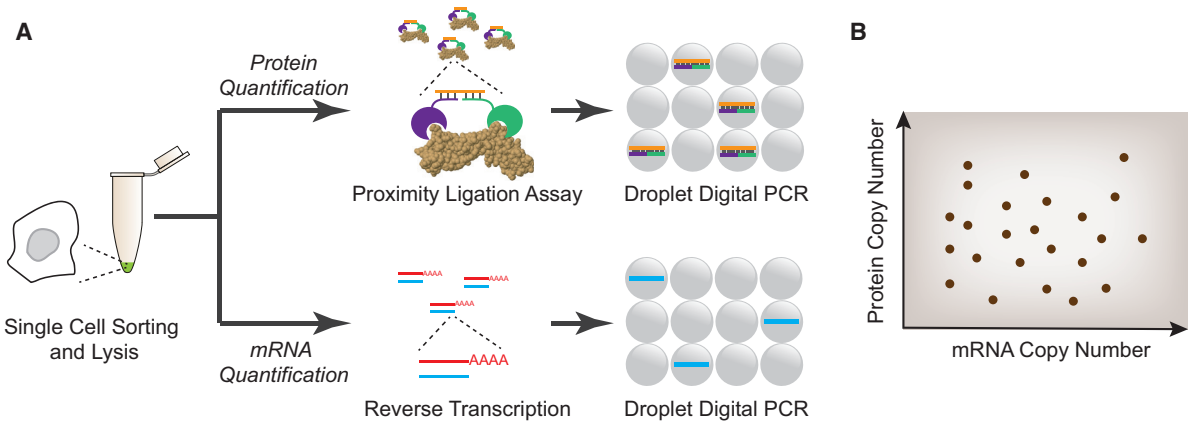

Figure 1. Digital Quantification of Both mRNA and Protein in the Same Cell

(A) A method for quantifying mRNA and protein molecules in individual cells. Single mammalian cells are sorted and lysed. The lysate is split into two fractions for protein and mRNA quantification, respectively. Protein copy number is quantified by proximity ligation assay followed by droplet digital PCR (top). mRNA copy number is quantified by reverse transcription followed by droplet digital PCR (bottom).

(B) The correlation between mRNA and protein copy numbers (schematic) can provide insights into the mechanisms underlying gene expression.

describe the gene expression of CD147 (Albayrak et al., 2016). In this model, they assume that (1) the promoter switches randomly between on and off states and (2) the extrinsic noise comes from protein translation. The model, which was parameterized with experimentally determined means and variances of mRNA and protein copy numbers, recapitulated the low mRNA-protein correlation (Albayrak et al., 2016), suggesting that extrinsic posttranscriptional noise may play a key role in disrupting CD147 mRNA and protein copy number correlation in single mammalian cells.

Further insights into mammalian gene expression can be gained by addressing two remaining questions: What are the exact sources of extrinsic noise? And is the lack of correlation between mRNA and protein copy numbers a general feature for mammalian genes at the single-cell level? It should be noted that at the population level, the correlation between mean levels of mRNA and protein across different genes is noticeable, presumably due to the absence of stochastic effects (Taniguchi et al., 2010).

The technique developed by Tay and colleagues joins a growing list of tools for simultaneous quantification of mRNA and protein in single cells (Albayrak et al., 2016; Darmanis et al., 2016; Frei et al., 2016; Taniguchi et al., 2010). The pervasiveness of dynamic gene regulation in various cellular regulatory systems (Levine et al., 2013) highlights the increasing importance of these single-cell analysis tools. This technique's unique ability to provide absolute protein and mRNA counts, together with its simple workflow, will allow researchers to dissect the fundamental principles of gene regulation by taking "digital snapshots" of gene expression in single cells. Further developments of the technique may provide multiplexing capability and enable us to digitally quantify the expression of a network of genes in the same cell. At the same time, it also highlights the need for fully dynamic measurements, e.g., through time-lapse microscopy, in which temporally delayed relationships between variables in the same cell, such as an mRNA and its protein product, can be detected and analyzed.

REFERENCES

Albayrak, C., Jordi, C.A., Zechner, C., Lin, J., Bichsel, C.A., Khammash, M., and Tay, S. (2016). Mol. Cell *61*, this issue, 914–924.

Darmanis, S., Gallant, C.J., Marinescu, V.D., Niklasson, M., Segerman, A., Flamourakis, G., Fredriksson, S., Assarsson, E., Lundberg, M., Nelander, S., et al. (2016). Cell Rep. *14*, 380–389.

Eldar, A., and Elowitz, M.B. (2010). Nature *467*, 167–173.

Fredriksson, S., Gullberg, M., Jarvius, J., Olsson, C., Pietras, K., Gústafsdóttir, S.M., Ostman, A., and Landegren, U. (2002). Nat. Biotechnol. *20*, 473–477.

Frei, A.P., Bava, F.A., Zunder, E.R., Hsieh, E.W., Chen, S.Y., Nolan, G.P., and Gherardini, P.F. (2016). Nat. Methods *13*, 269–275.

Hughes, A.J., Spelke, D.P., Xu, Z., Kang, C.C., Schaffer, D.V., and Herr, A.E. (2014). Nat. Methods *11*, 749–755.

Levine, J.H., Lin, Y., and Elowitz, M.B. (2013). Science *342*, 1193–1200.

Muramatsu, T., and Miyauchi, T. (2003). Histol. Histopathol. *18*, 981–987.

Shi, Q., Qin, L., Wei, W., Geng, F., Fan, R., Shin, Y.S., Guo, D., Hood, L., Mischel, P.S., and Heath, J.R. (2012). Proc. Natl. Acad. Sci. USA *109*, 419–424.

Taniguchi, Y., Choi, P.J., Li, G.W., Chen, H., Babu, M., Hearn, J., Emili, A., and Xie, X.S. (2010). Science *329*, 533–538.