TOWARDS PARAMETERIZING THE CENTRAL DOGMA OF MOLECULAR BIOLOGY ON A GENOME-WIDE SCALE

Meera Gupta*,[1,2] Hsin-Jung (Sophia) Li,[1] Ned Wingreen,[1] Zemer Gitai,[1] **Martin Wühr[1,2 #]**

1 Department of Molecular Biology & the Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA
2 DOE Center for Advanced Bioenergy and Bioproducts Innovation
*presenting author: meerag@princeton.edu
correspondence: wuhr@princeton.edu, 617-2307625

https://cabbi.bio/

Project Goals: Flux through any biochemical network depends on enzyme abundance. We aim to develop experimental techniques and statistical framework [1] to quantify all parameters of the central dogma of molecular biology that determine protein expression levels on a genome wide scale. To this end we combine RNA seq, ribosomal profiling, and mass-spectrometry based quantitative proteomics. We aim to determine how the expression level of proteins is set and how cells adapt their proteome to different environments. This will help us understand the economies of enzyme production. This insight will likely help engineer cell's enzyme levels to produce desired metabolic products. We have started the development of these techniques in *E. coli* since its simplicity speeds up the technology development. We anticipate the generated insights and technology to be easily transferrable to yeast and other microbes relevant for biotechnology.

Abstract

Proteins are the direct determinants of cellular functions. Being able to understand how each protein is formed is critical to tuning their expression levels to obtain the desired product. Modulation of protein levels can be achieved by controlling transcription, translation, or protein degradation. While, we understand this regulation for few of the individual genes, a quantitative and comprehensive understanding of regulatory processes is still lacking. To address this shortcoming, we need to parametrize the central dogma of molecular biology for the entire proteome. Recent advances in high throughput genomics have enabled us to globally quantify transcription and translation rates. However, till date, the measurement of degradation contribution at a global scale remains a challenge. This is mainly due to the lack of accurate and precise measurements of absolute and relative protein expression levels.

As a preliminary result, we have global measurements of differential degradation rates in Nitrogen-limited vs the Phosphorous-limited growth conditions for *E. coli*. These measurements were obtained by comparing translation rates, measured through ribosomal profiling, with the relative protein expression levels in the steady state growth conditions. Remarkably, the results indicated that using the recently published TMTc+ method [2] for protein measurements removes the systematic biases in relative proteomics measurements. However, this approach provides only insight into proteins degradation rates that differ between conditions. To confirm our measurements, and to obtain absolute degradation rates we measured the decline of protein

abundances upon ribosomal translation inhibition. This approach has been previously used on a proteome-wide scale with TAP-tagged libraries [3] . However, we find strong evidence that this approach is prone to artifacts and the half-life of many proteins is drastically altered upon drug addition. To overcome this limitation, we have started to measure protein turn-over rates by a pulse-chase experiment with heavy isotope labeled amino-acids. Our preliminary shallow experiments indicate that only a small subset of proteins in *E. coli* (<10%) shows protein half-lives which are significantly shorter than doubling time. Once these techniques have been established, we intent to apply them to yeast species that are relevant for metabolic engineering and determine the parameters that set metabolic enzyme levels under various growth conditions.

References

1. Peshkin, L., et al., *Bayesian Confidence Intervals for Multiplexed Proteomics Integrate Ion-Statistics with Peptide Quantification Concordance.* bioRxiv, 2019: p. 210476.
2. Sonnett, M.Y., E. ; Wühr, M, *Accurate, Sensitive, and Precise Multiplexed Proteomics using the Complement Reporter Ion Cluster.* 2017, Princeton University.
3. Belle, A., et al., *Quantification of protein half-lives in the budding yeast proteome.* Proc Natl Acad Sci U S A, 2006. **103**(35): p. 13004-9.

This work was supported by the U.S. Department of Energy (DOE) under Grant # DE-SC0018420.