PROTEIN EXPRESSION

Finding abundance regulators

A new pooled screening method in yeast allows scientists to probe how protein levels are regulated by mutating thousands of genes at once.

OLIVIA OZGUC AND SIBYLLE VONESCH

> **Related research article** Schubert OT, Bloom JS, Sadhu MJ, Kruglyak L. 2022. Genome-wide base editor screen identifies regulators of protein abundance in yeast. *eLife* **11**:e79525. doi: 10.7554/eLife.79525

Each cell in our body contains over 10,000 different proteins that perform the activities the cell needs to live. But how do cells know how much of each protein to make? The information needed to make each protein is encoded in our genes, which are used as templates to make messenger RNAs (mRNAs) in a process called transcription. The mRNAs are in turn translated into proteins by ribosomes. Each of these steps is executed and tightly controlled by several proteins, which together form a regulatory network that determines the amounts of a protein that get synthesized. For most proteins, we only know certain parts of this network, making it difficult to predict the factors that affect protein levels in the cell, which can play a role in disease.

Geneticists can identify which proteins help to regulate protein levels in the cell by inactivating individual genes to see whether stopping the production of the protein coded for by that gene affects the levels of another protein. Unfortunately, doing this for each regulator protein individually is expensive and time-consuming. So-called 'pooled screens' can survey many inactivating mutations at once in a single tube full of cells, but it is challenging to read out the effects of these perturbations on protein levels. Advances in single-cell RNA sequencing have allowed scientists to link specific mutations to changes in mRNA levels (*Datlinger et al., 2017*). However, mRNA levels are not always good predictors for protein abundance, because several cellular mechanisms that regulate protein abundance act only after transcription (*Vogel and Marcotte, 2012*). Some of these mechanisms can even buffer the effects of genetic mutations, such that a change in mRNA level due to a mutation does not lead to a change in protein level (*Buccitelli and Selbach, 2020*).

Now, in eLife, Olga Schubert, Leonid Kruglyak and colleagues at the University of California, Los Angeles report on a new pooled screening approach to directly study how protein abundance is regulated in the yeast *Saccharomyces cerevisiae* (*Schubert et al., 2022*). To generate the necessary inactivating mutations in the yeast cells, the team used a DNA base editor – a molecular tool that can be directed to convert one type of DNA base to another at specific sites determined by a targeting sequence (a so-called gRNA). The editor was originally developed for human cells (*Komor et al., 2016*), so Schubert et al. first had to show that it could efficiently make these mutations in yeast. They did this by using the editor to inactivate genes necessary for survival in yeast, which led to the cells dying.

Next, Schubert et al. developed an approach to read out the effect of inactivating mutations on the levels of 11 proteins of interest (*Figure 1*). First, they generated 11 strains of yeast each expressing a protein of interest fused to a fluorescent protein. These strains allowed Schubert et al. to have a straightforward readout of protein abundance: for proteins engineered this way, fluorescence can be used as a proxy for protein abundance. Next, the base editor was used to introduce inactivating mutations for all yeast genes in these strains, with each cell only
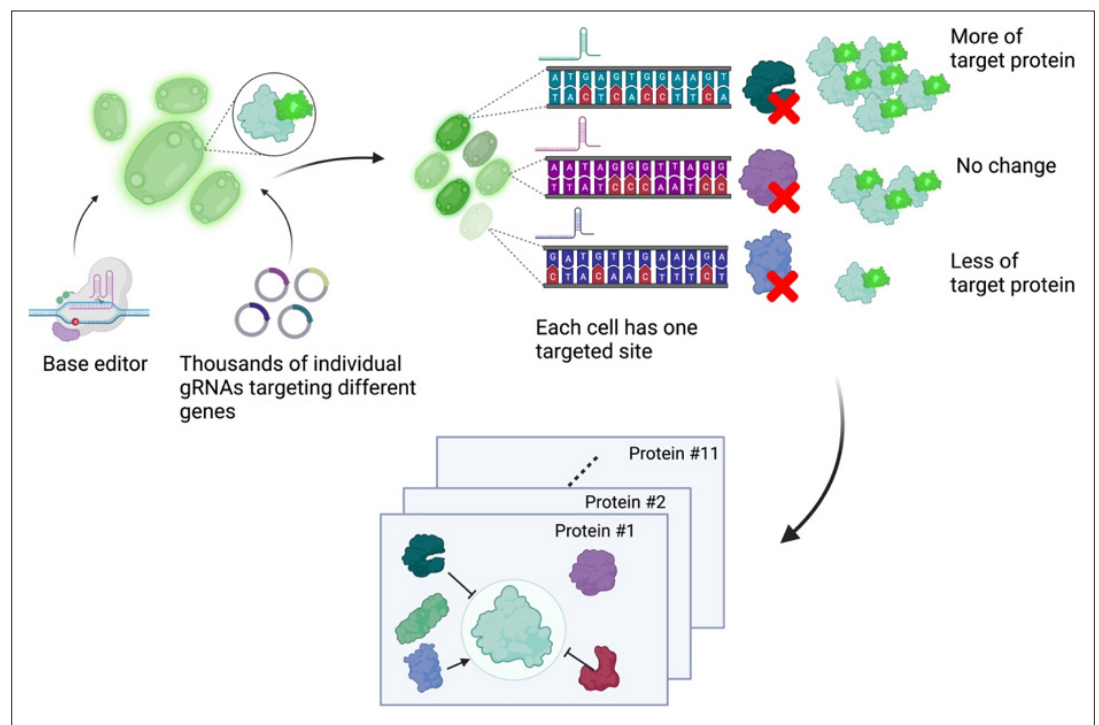
Figure 1. A base editor screen to investigate regulatory relationships affecting protein levels. Left: a library of thousands of guide RNA plasmids (gRNA, circles with sections in different colours) targeting genes across the genome are introduced, along with the base editor, into an engineered yeast strain containing the genetic code for a protein of interest fused to the fluorescent protein GFP. Centre: each cell has a different target gene disrupted by the base editor (shown as DNA sequences in different colours), resulting in a specific protein (the green, purple and blue shapes next to the sequences) not being produced in each cell. Right: after editing, each cell will have different levels of fluorescence depending on which gene was disrupted in it. The top cell has higher fluorescence than an unedited cell because the disrupted gene in it codes for a protein that normally lowers the levels of the protein of interest. The centre cell has the same fluorescence levels as an unedited cell because the disrupted gene in it codes for a protein that does not regulate the levels of the protein of interest. The bottom cell has low levels of fluorescence because the disrupted gene in it codes for a protein that normally increases the levels of the protein of interest. The cells are then sorted depending on their fluorescence levels, and the identity of the gene mutated in each cell is determined by sequencing the gRNA. Bottom: this pooled screening approach allows scientists to define regulatory networks that determine the abundance of each target protein and reveals general properties of networks governing protein expression.

Figure created using BioRender.com

receiving mutations at one target site encoded by the gRNA. After editing, cells with high or low fluorescence were isolated using a method called fluorescence-activated cell sorting, and the identity of the perturbed genes determined by sequencing of the gRNA.

The results showed that mutations that only affected one or two proteins of interest were often found in genes coding for transcriptional regulators. On the other hand, mutating genes that make key components for the translation machinery reduced abundance of most of the proteins under study. This implies that genes involved in translation are more likely to be hubs in protein regulatory networks. Schubert et al. also found that mutations that affected more proteins of interest were also more likely to reduce protein abundance and decrease cellular fitness, since they were often found in genes needed for survival.

Besides confirming known regulatory relationships, Schubert et al. also found many new ones. For example, they propose a new regulatory pathway that links the cell's ability to sense extracellular amino acids to the activation of the stress-responsive protein Yhb1, increasing stress resistance. The experiments also identified distinct regulatory proteins for different forms of GAPDH, a multifunctional enzyme known for its role in glucose metabolism. This

suggests that each form of the enzyme may be required in different metabolic states. GAPDH is an important human enzyme, and it can interact with proteins implicated in neurodegenerative diseases, so understanding how it works and how it is regulated has broad implications beyond yeast (*Butterfield et al., 2010*).

The findings of Schubert et al. highlight how identifying new regulators, even for well-studied proteins, is important for basic science. In the future, coupling this method with tools that introduce mutations that do not fully inactivate the target genes (*Roy et al., 2018*; *Sharon et al., 2018*) may allow a more nuanced view of how protein abundance is encoded in the genome. This, combined with improvements in the speed and sensitivity of protein quantification (*Messner et al., 2021*; *Brunner et al., 2022*), may soon allow scientists to map regulators for thousands of proteins at once. Integrating Schubert et al.'s approach with methods like high-throughput imaging or phospho-proteomics (*Humphrey et al., 2018*) will also increase our understanding of proteins that are mainly regulated through phosphorylation or by their localization in the cell. Overall, the ability to assess more proteins in diverse environments and organisms will deepen our understanding of the regulatory principles that govern protein expression.

Olivia Ozguc is in the Center for Microbiology, VIB-KU Leuven, Leuven, Belgium and in the Centre of Microbial and Plant Genetics, Department of Microbial and Molecular Systems, KU Leuven, Leuven, Belgium
olivia.ozguc@kuleuven.be

Sibylle Vonesch is in the Center for Microbiology, VIB-KU Leuven, Leuven, Belgium and in the Centre of Microbial and Plant Genetics, Department of Microbial and Molecular Systems, KU Leuven, Leuven, Belgium
sibylle.vonesch@kuleuven.be
http://orcid.org/0000-0003-2485-1048

Competing interests: The authors declare that no competing interests exist.

Published 03 November 2022

References

Brunner AD, Thielert M, Vasilopoulou C, Ammar C, Coscia F, Mund A, Hoerning OB, Bache N, Apalategui A, Lubeck M, Richter S, Fischer DS, Raether O, Park MA, Meier F, Theis FJ, Mann M. 2022. Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *Molecular Systems Biology* **18**:e10798. DOI: https://doi.org/10.15252/msb.202110798, PMID: 35226415

Buccitelli C, Selbach M. 2020. mRNAs, proteins and the emerging principles of gene expression control. *Nature Reviews Genetics* **21**:630–644. DOI: https://doi.org/10.1038/s41576-020-0258-4, PMID: 32709985

Butterfield DA, Hardas SS, Lange MLB. 2010. Oxidatively modified glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and Alzheimer's disease: many pathways to neurodegeneration. *Journal of Alzheimer's Disease* **20**:369–393. DOI: https://doi.org/10.3233/JAD-2010-1375, PMID: 20164570

Datlinger P, Rendeiro AF, Schmidl C, Krausgruber T, Traxler P, Klughammer J, Schuster LC, Kuchler A, Alpar D, Bock C. 2017. Pooled CRISPR screening with single-cell transcriptome readout. *Nature Methods* **14**:297–301. DOI: https://doi.org/10.1038/nmeth.4177, PMID: 28099430

Humphrey SJ, Karayel O, James DE, Mann M. 2018. High-throughput and high-sensitivity phosphoproteomics with the EasyPhos platform. *Nature Protocols* **13**:1897–1916. DOI: https://doi.org/10.1038/s41596-018-0014-9, PMID: 30190555

Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. 2016. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**:420–424. DOI: https://doi.org/10.1038/nature17946, PMID: 27096365

Messner CB, Demichev V, Bloomfield N, Yu JSL, White M, Kreidl M, Egger AS, Freiwald A, Ivosev G, Wasim F, Zelezniak A, Jürgens L, Suttorp N, Sander LE, Kurth F, Lilley KS, Mülleder M, Tate S, Ralser M. 2021. Ultra-fast proteomics with scanning SWATH. *Nature Biotechnology* **39**:846–854. DOI: https://doi.org/10.1038/s41587-021-00860-4, PMID: 33767396

Roy KR, Smith JD, Vonesch SC, Lin G, Tu CS, Lederer AR, Chu A, Suresh S, Nguyen M, Horecka J, Tripathi A, Burnett WT, Morgan MA, Schulz J, Orsley KM, Wei W, Aiyar RS, Davis RW, Bankaitis VA, Haber JE, et al. 2018. Multiplexed precision genome editing with trackable genomic barcodes in yeast. *Nature Biotechnology* **36**:512–520. DOI: https://doi.org/10.1038/nbt.4137, PMID: 29734294

Schubert OT, Bloom JS, Sadhu MJ, Kruglyak L. 2022. Genome-wide base editor screen identifies regulators of protein abundance in yeast. *eLife* **11**:e79525. DOI: https://doi.org/10.7554/eLife.79525

Sharon E, Chen S-AA, Khosla NM, Smith JD, Pritchard JK, Fraser HB. 2018. Functional genetic variants revealed by massively parallel precise genome editing. *Cell* **175**:544–557.. DOI: https://doi.org/10.1016/j.cell.2018.08.057, PMID: 30245013

Vogel C, Marcotte EM. 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics* **13**:227–232. DOI: https://doi.org/10.1038/nrg3185, PMID: 22411467