HIGHLIGHTS

URLs

green fluorescent protein
<http://us.expasy.org/cgi-bin/nice-prot.pl?P42212>

PROTEOMICS

How to see the big picture

Scientific research can often focus on the smallest problems, so it's important to keep the big picture in mind. For example, how does global protein expression change under different conditions, and what are the subcellular localizations of the proteins expressed? In *Nature*, the groups of O'Shea and Weissman now provide a way to answer these questions, because, in their two papers, they describe methods for the global analysis of protein expression and protein localization in *Saccharomyces cerevisiae*.

The availability of high-throughput technologies and the complete genome sequence of *S. cerevisiae*, which contains 6,234 annotated open reading frames (ORFs), has made it feasible to analyse this eukaryotic proteome systematically. One problem, though, is how to tag the ORFs. Tagging can disrupt protein localization signals or lead to overexpression, which can subsequently result in protein mislocalization.

To overcome this problem, the O'Shea and Weissman groups created *S. cerevisiae* fusion libraries by using homologous recombination to integrate a specific tag in-frame immediately after the stop codon of each ORF in its natural chromosome location. To study global protein expression, they made a tandem affinity purification (TAP)-tagged fusion library, whereas to study global protein localization, they made a **green**

fluorescent protein (GFP)-tagged fusion library. As these tags don't disrupt the upstream regulatory or promoter sequences of each gene, the fusion proteins are likely to behave in the same way as normal proteins.

Creating a TAP-tagged library allowed the authors to show that ~80% of the yeast proteome is expressed under normal growth conditions, and also enabled them to identify misannotated genes. Approximately 1,500 of the annotated ORFs did not produce gene products, and, whereas ~1,000 of these ORFs probably encode proteins that have functions under specialized conditions that are not characteristic of authentic genes.

The use of the TAP-tagged library also allowed the authors to show that protein abundance can range from less than 50 to more than 1,000,000 molecules per cell. So, many proteins — including essential proteins and most transcription factors — are present at levels that cannot be detected by other proteomics techniques. Furthermore, by comparing their data with studies of messenger RNA levels, the authors showed that the protein:mRNA ratio is, on average, 4,800 — a ratio that is surprisingly constant over a range of mRNA abundances. This indicates that mRNA and protein abundance are closely correlated.

By creating a GFP-tagged library, the O'Shea and Weissman groups were able to classify 4,156 proteins

(75% of the yeast proteome) into 22 distinct subcellular localizations. The resolution limits of fluorescence microscopy were overcome by visualizing the localization of the GFP-tagged proteins alone and their co-localization with red fluorescent protein-tagged reference proteins. Importantly, this work provided localization data for previously unlocalized yeast proteins that represent ~30% of the yeast proteome.

This global analysis of protein localization, when analysed in the context of genetic, transcriptional and protein–protein interaction data, has helped to “...reveal the logic of transcriptional co-regulation, and provides a comprehensive view of interactions within and between organelles in eukaryotic cells”. And, together, the two papers from the O'Shea and Weissman groups have given us our most complete and physiologically relevant view so far of global protein expression and localization in a eukaryotic cell.

Rachel Smallridge

References and links

ORIGINAL RESEARCH PAPERS Huh, W.-K. *et al.* Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003) | Ghaemmaghami, S. *et al.* Global analysis of protein expression in yeast. *Nature* **425**, 737–741 (2003)

FURTHER READING Wohlschlegel, J. A. & Yates, J. R. Where's Waldo in yeast? *Nature* **425**, 671–672 (2003)

WEB SITES

Erin O'Shea's laboratory:
<http://www.ucsf.edu/ekolab/>

Jonathan Weissman's laboratory:
<http://www.ucsf.edu/jswlab/>

Yeast GFP fusion localization database:
<http://yeastgfp.ucsf.edu>