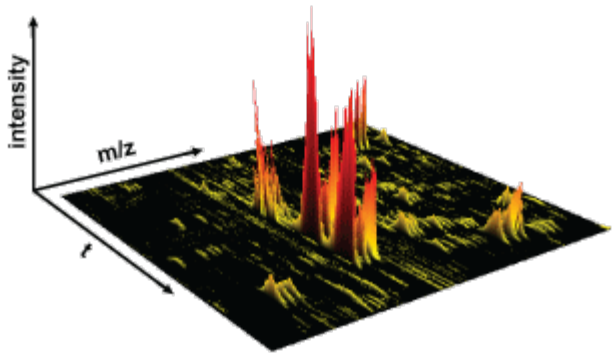
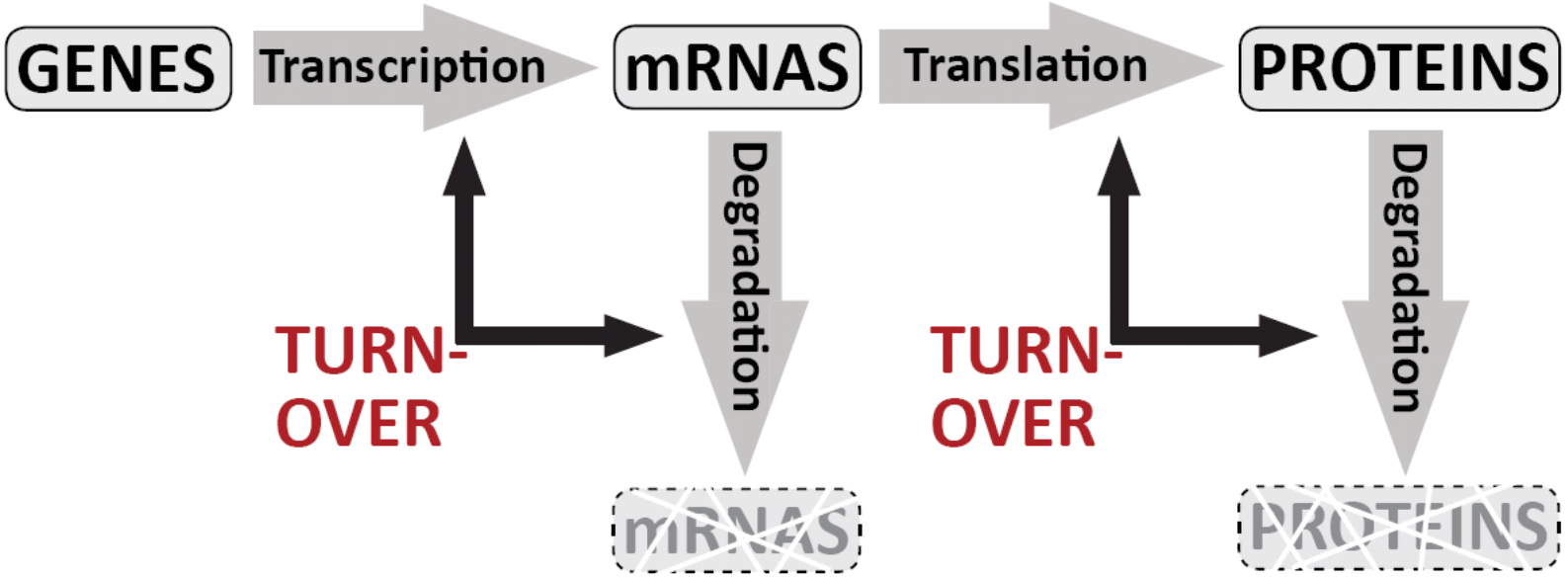
Genome-wide analysis of protein and mRNA half-lives reveals dynamic properties of mammalian gene expression

Matthias Selbach



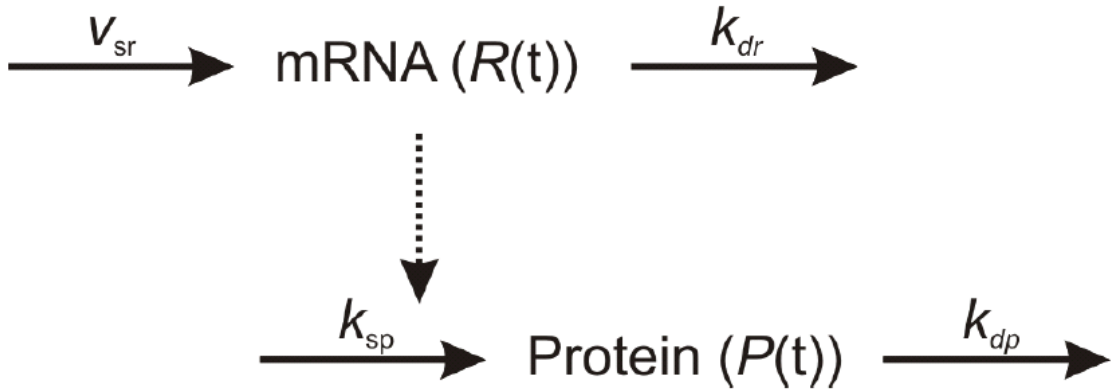
Cell Signaling and Mass Spectrometry
Max Delbrück Center for Molecular Medicine
Berlin, Germany
www.mdc-berlin.de/selbach

Dynamic regulation of gene expression



➢Dynamic properties of gene expression depend on protein and mRNA levels and half-lives

Mathematical modeling of gene expression

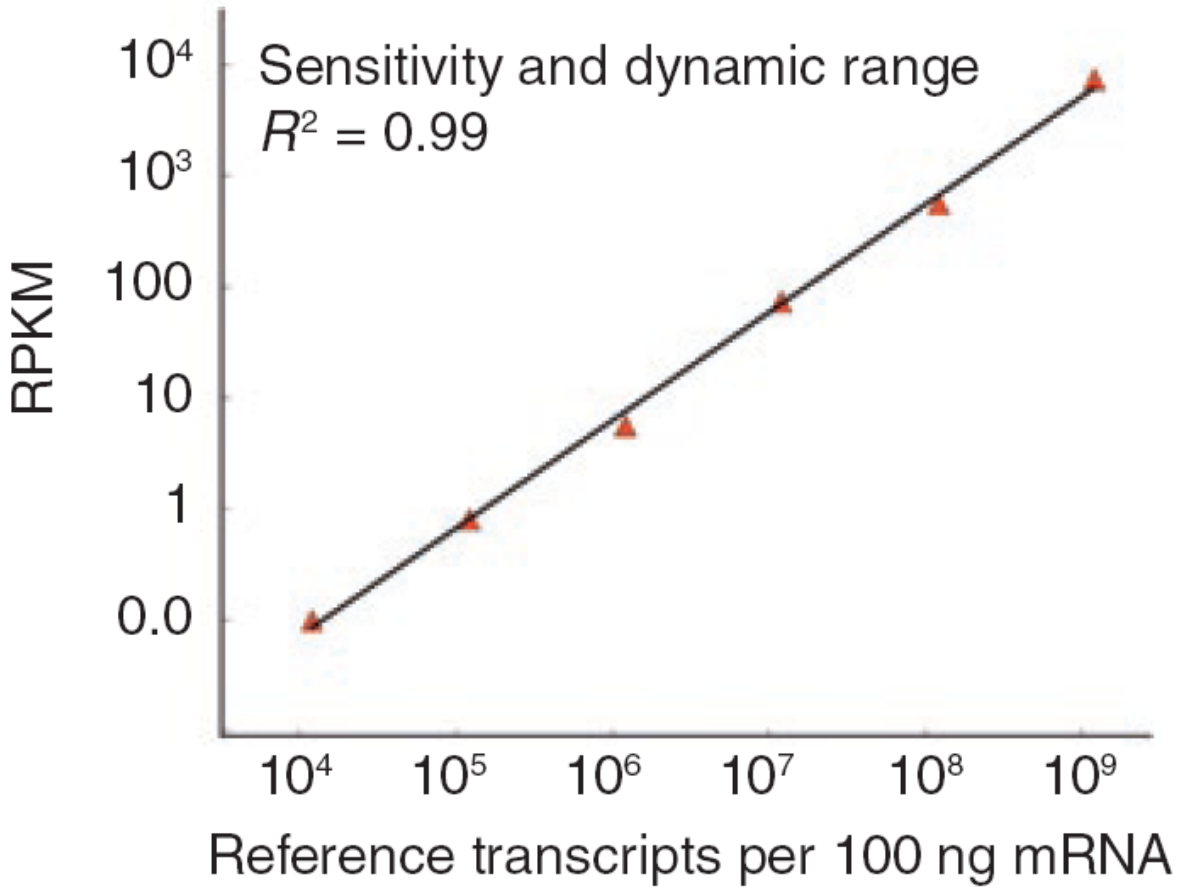


$$\frac{dR}{dt} = v_{sr} - k_{dr}R$$

$$\frac{dP}{dt} = k_{sp}R - k_{dp}P$$

Absolute mRNA levels: next generation sequencing

reads per kb of exon model per million mapped reads (RPKM)



Mortazavi et al., N. Meth.,2008

Methods for absolute protein quantifaction

Stable isotope-based methods:

• AQUA peptides	Gerber et al., PNAS, 2003
• iTRAQ	Ross et al., MCP, 2004
• QconCAT	Beynon et al., N. Meth., 2005
• absolute SILAC	Hanke et al., JPR, 2008

Label free / computational methods:

• protein abundance index (PAI/emPAI)	Ishihama et al, MCP, 2005
• normalized spectral abundance factor (NSAF)	Mosley et al., J.Prot., 2009
• absolute protein expression measurements (APEX)	Lu et al., Nat. Biotech, 2007
• top3 peptide intensities (top3)	Silva et al., MCP, 2006
	Malstrom et al., Nature, 2009

Methods for absolute protein quantifaction

Stable isotope-based methods:

• AQUA peptides	Gerber et al., PNAS, 2003
• iTRAQ	Ross et al., MCP, 2004
• QconCAT	Beynon et al., N. Meth., 2005
• absolute SILAC	Hanke et al., JPR, 2008

Label free / computational methods:

• protein abundance index (PAI/emPAI)	Ishihama et al, MCP, 2005
• normalized spectral abundance factor (NSAF)	Mosley et al., J.Prot., 2009
• absolute protein expression measurements (APEX)	Lu et al., Nat. Biotech, 2007
• top3 peptide intensities (top3)	Silva et al., MCP, 2006
	Malstrom et al., Nature, 2009
• intensity-based absolute quantification (iBAQ)	Schwanhaeusser et al.,unpublished

Benchmarking different methods

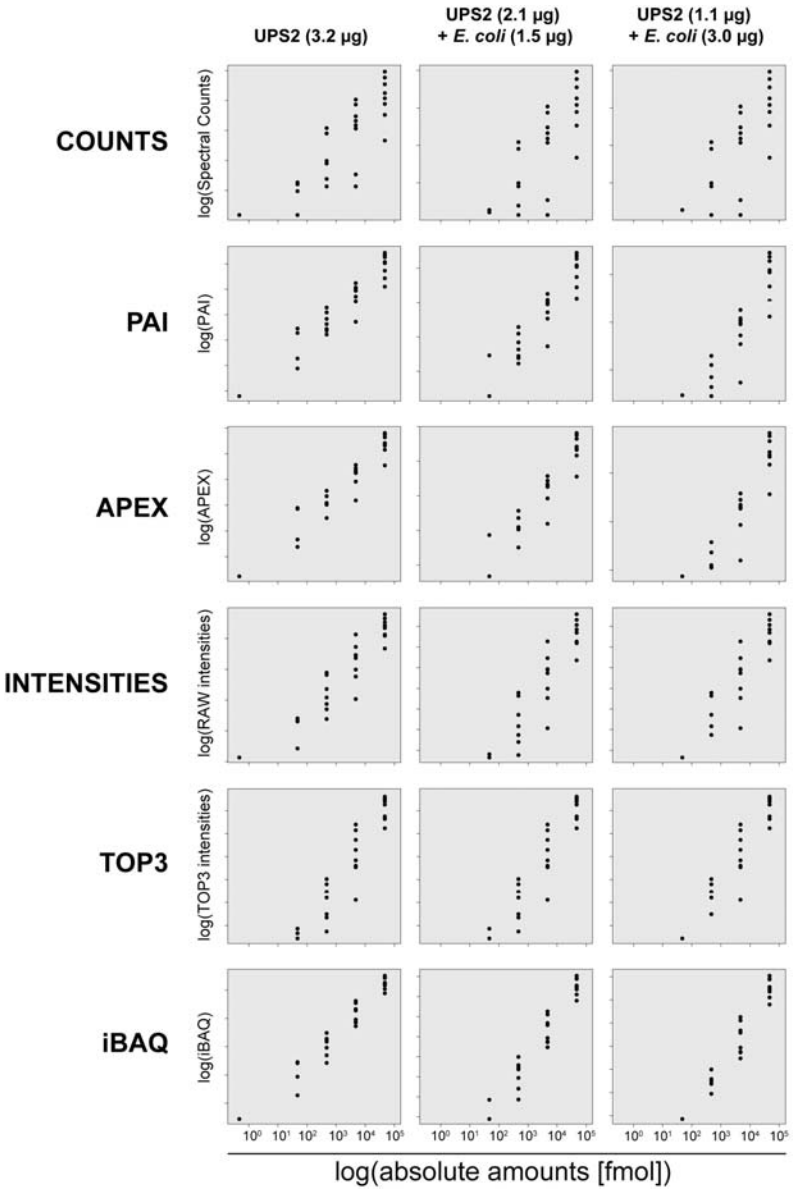
- should be independent of protein sequence
- be accurate over several orders of magnitude
- should not be affected by overall sample complexity

48 recombinant human proteins over six orders of magnitude (UPS2, Sigma)

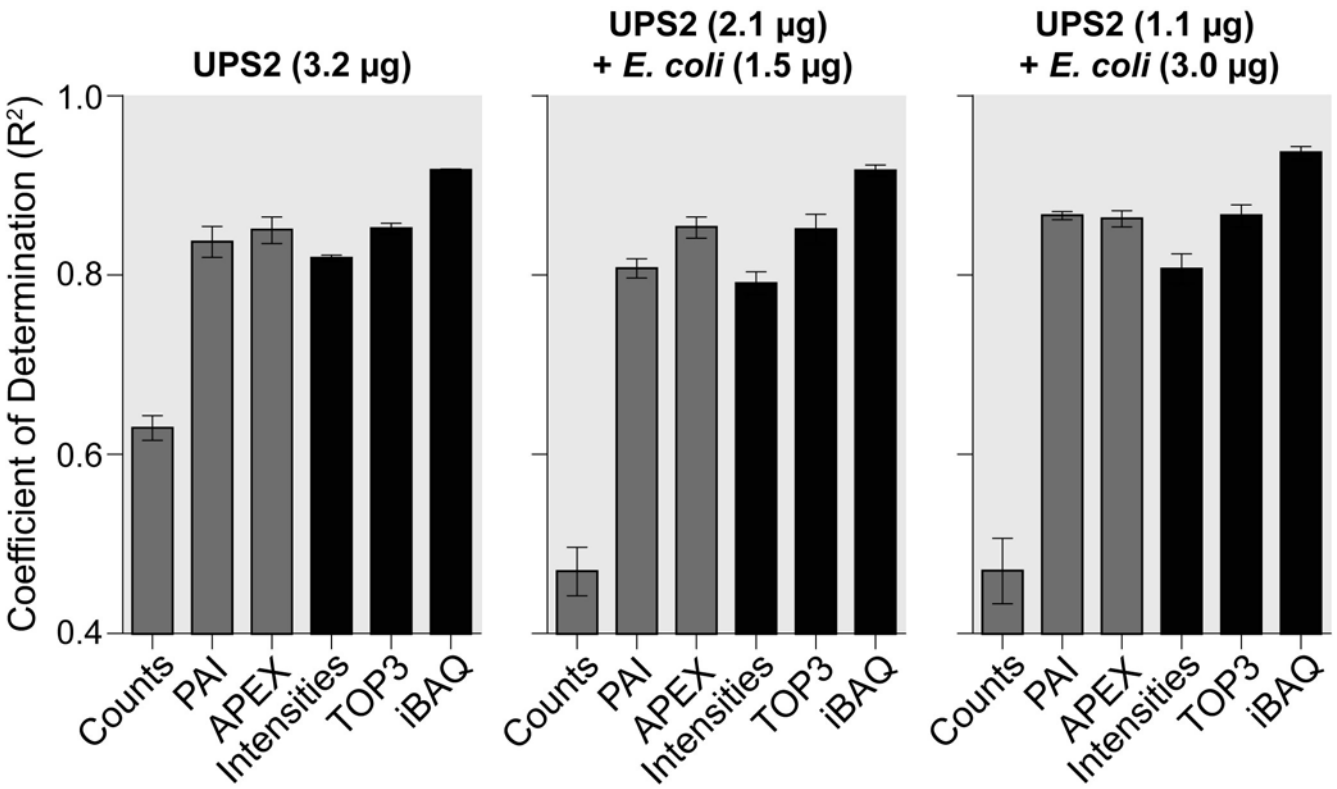
+

E. coli lysate (different amounts)

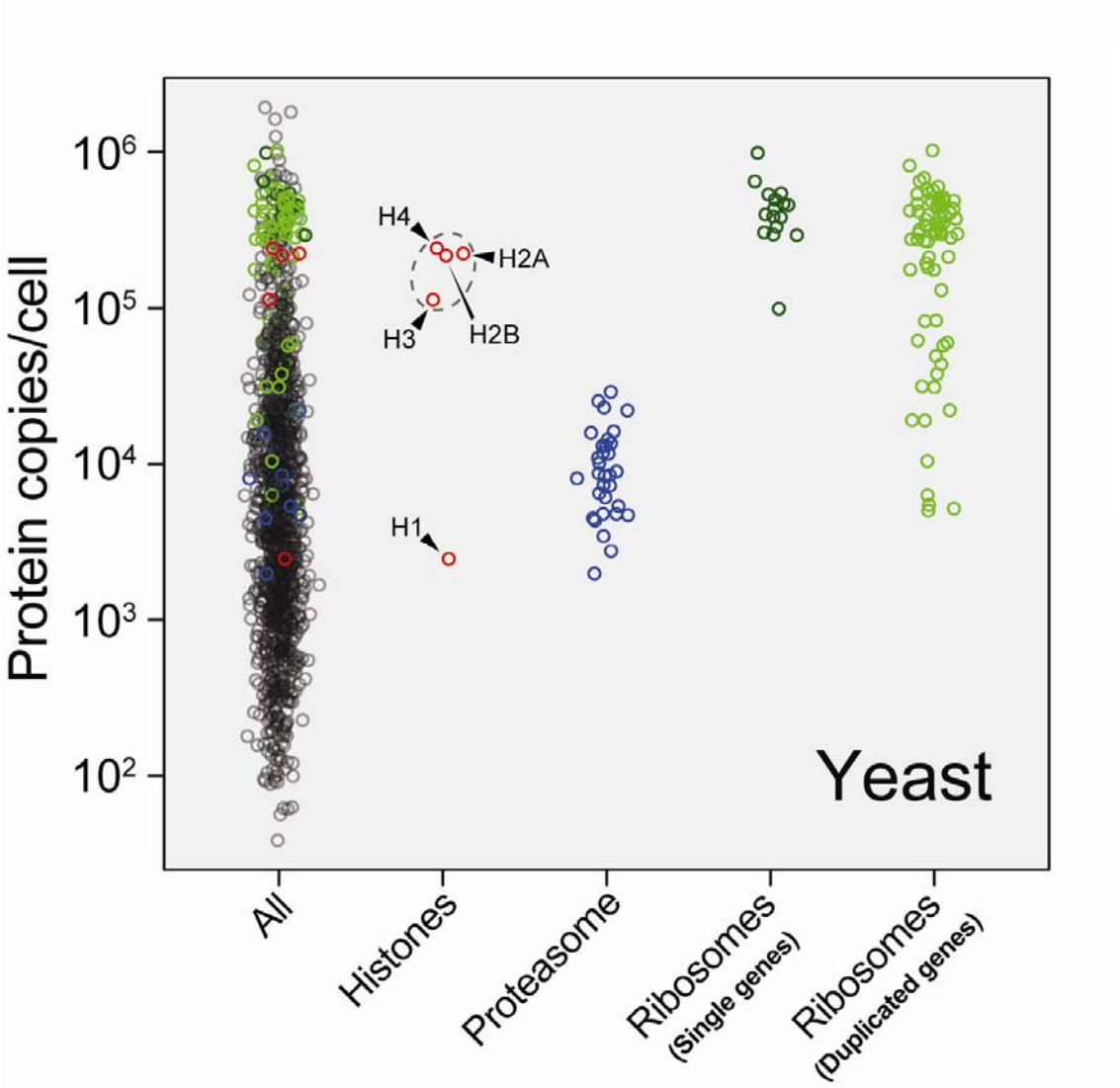
Benchmarking different methods

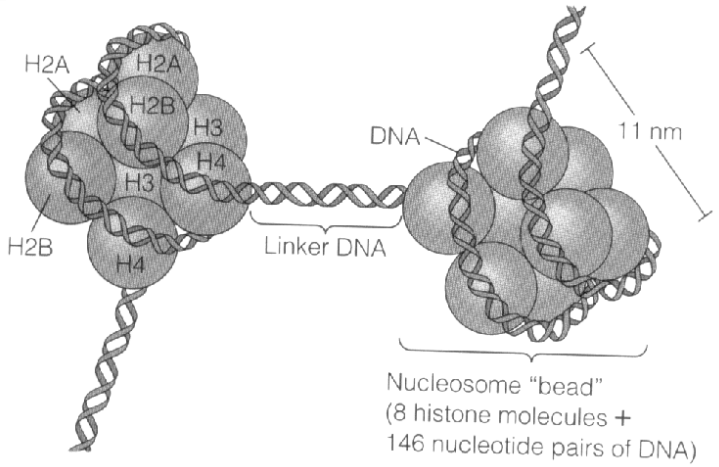


Benchmarking different methods



Cellular copy numbers of yeast and mouse (NIH3T3) proteins

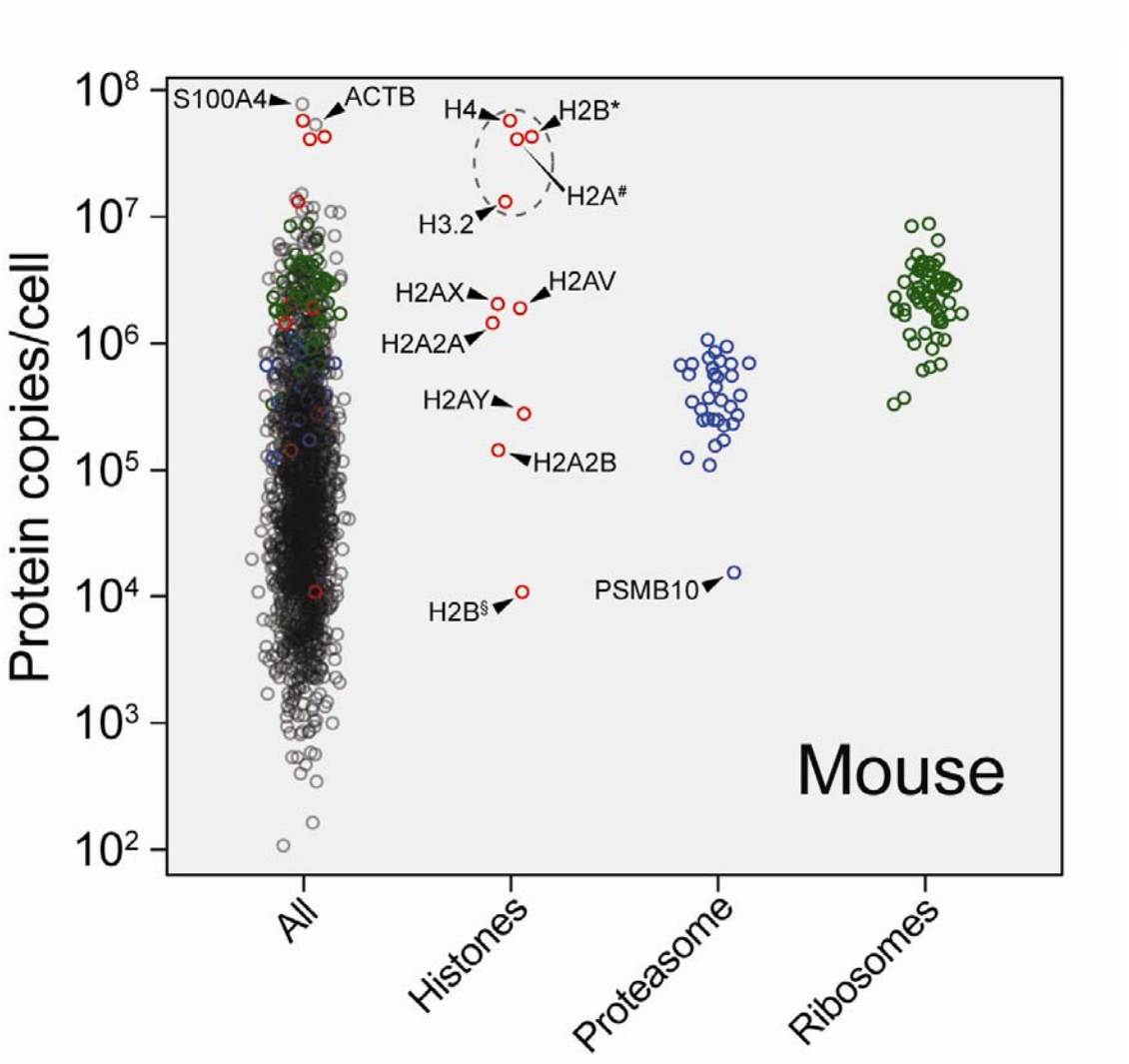


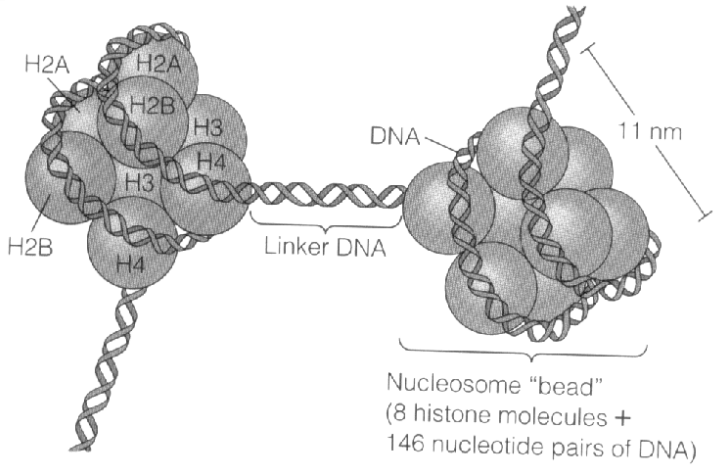


Yeast nucleosomes:
12 Mb genome / 150 bp
= 80,000 nucleosomes

iBAQ: ~100,000
nucleosomes

Cellular copy numbers of yeast and mouse (NIH3T3) proteins





mouse nucleosomes:
2.6 Gb haploid genome size x 2 / 150b = 3.5 x 10^7 nucleosomes

iBAQ: ~ 3 x 10^7 nucleosomes

Summary I: intensity-based absolute quantification (iBAQ)

- most accurate estimate of absolute protein abundance
- all available information is used (intensities of all peptides)
- spiked-in standard proteins minimize variability in sample handling
- very simple approach:
  - universal standards can be used (AQUA, absolute SILAC)
  - no complex machine learning procedures required (APEX)
  - no need for three different MS methods (top3)

Large scale analysis of mRNA and protein turnover

Classical approach:

- inhibition of transcription / translation using drugs
- track mRNA / protein decay over time

Problem:

- drug treatment has severe side effects and ultimately kills the cells under study

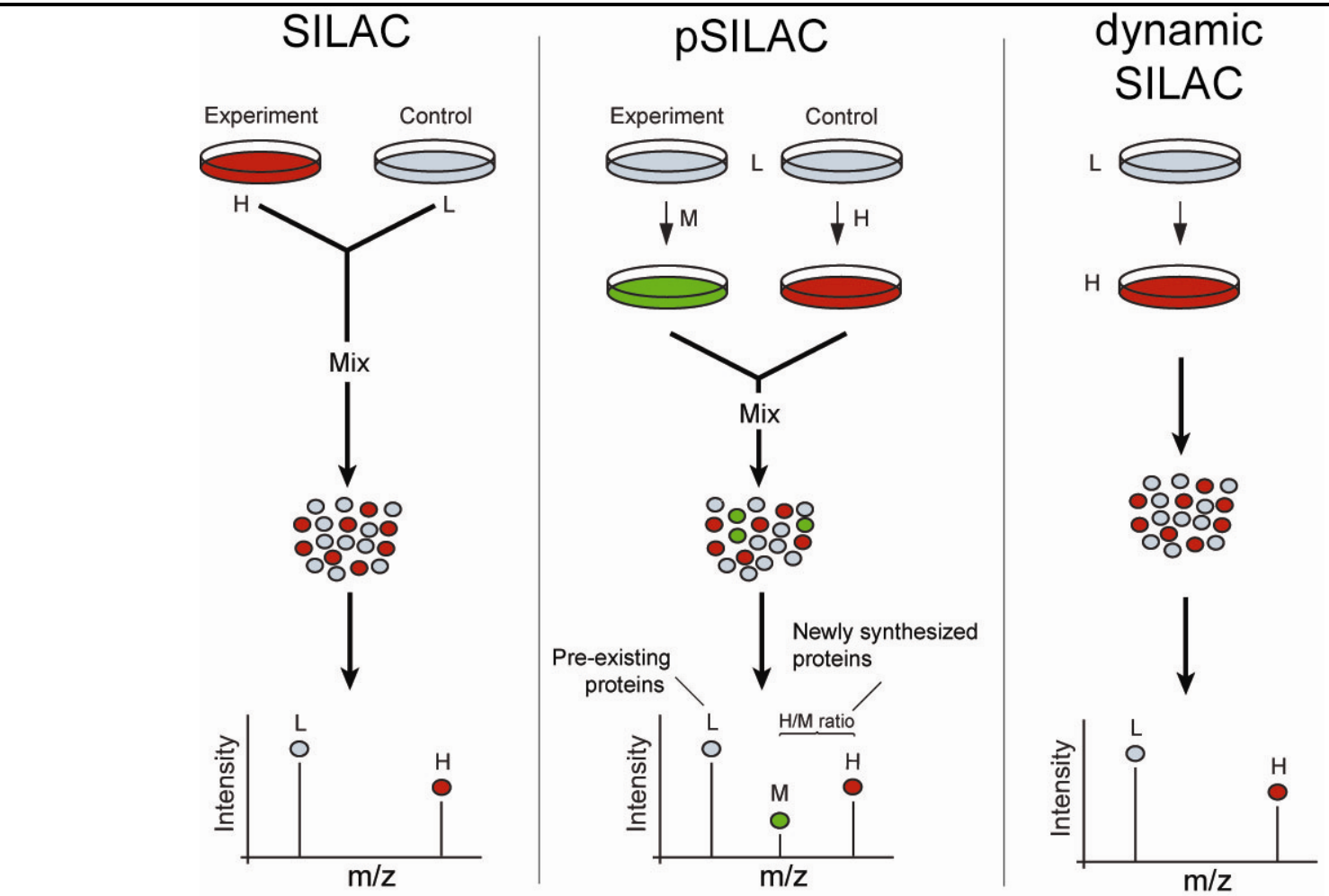
Labeling with isotopes: Rudolf Schoenheimer

Rudolf Schoenheimer, 1898-1941

Deuterium as an Indicator in the Study of Intermediary Metabolism, Schoenheimer and Rittenberg, J. Biol. Chem., 1935

„In order successfully to label a physiological substance, it is essential that the chemical and physical properties of the labeled substance be so similar to the unlabeled one that the animal organism will not be able to differentiate between them. The chemist, on the other hand, must be able to distinguish and to estimate them in small quantities and at high dilutions. A possibility for such a label is the use of an isotope."

Measuring protein dynamics using SILAC

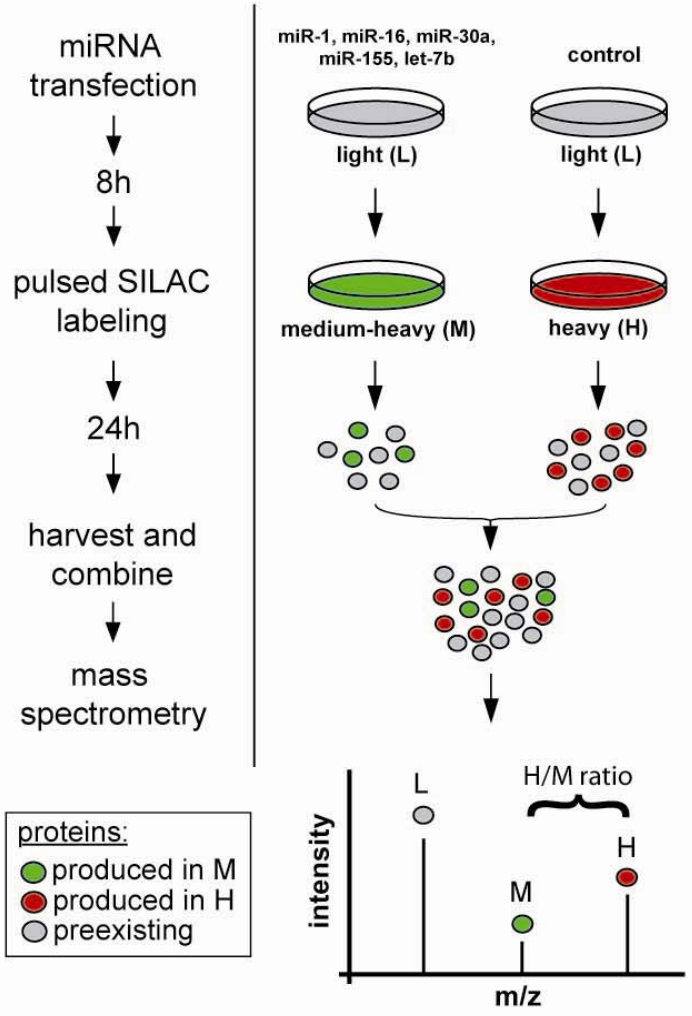


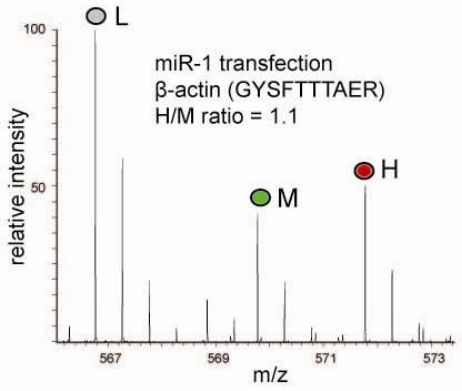
Relative changes in protein levels (Ong et al., 2002)

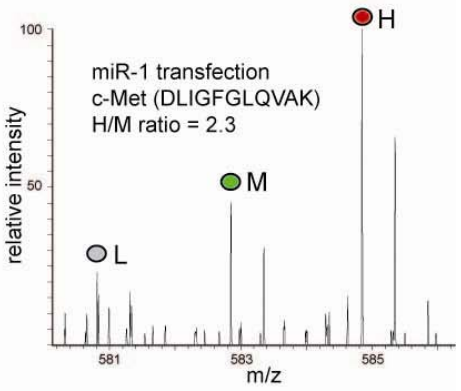
Relative changes in protein production (Schwanhaeusser et al, 2009)

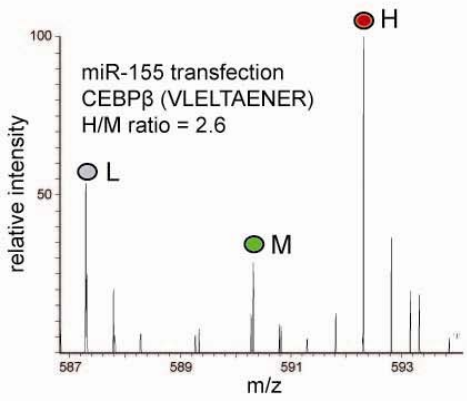
protein turnover (Doherty et al., 2009)

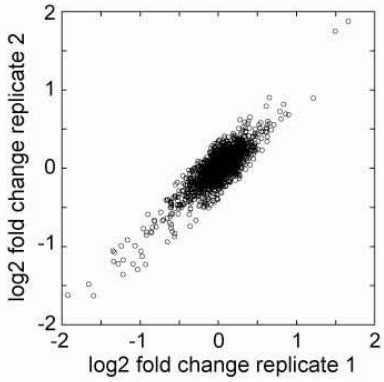
pSILAC identification of miRNA targets











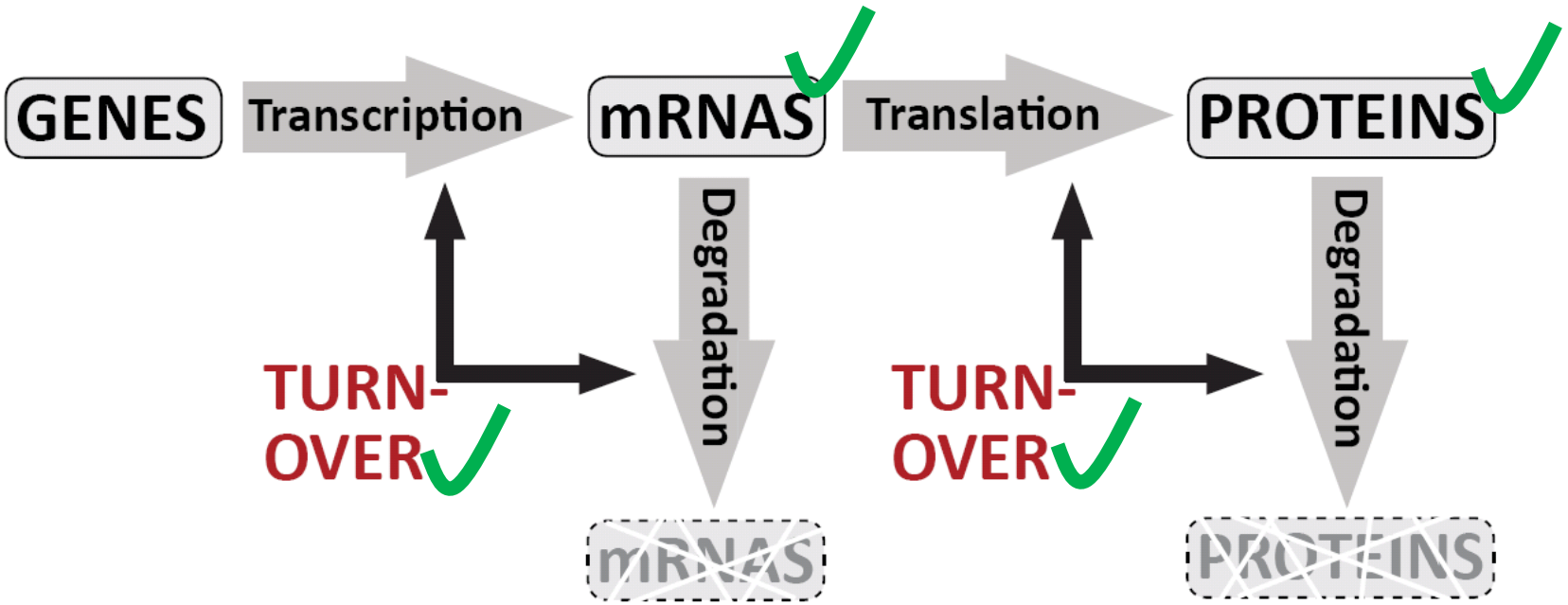
Nature, 2008

Metabolic pulse labeling of proteins and mRNAs

High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. RNA. 2008 Sep;14(9):1959-72.

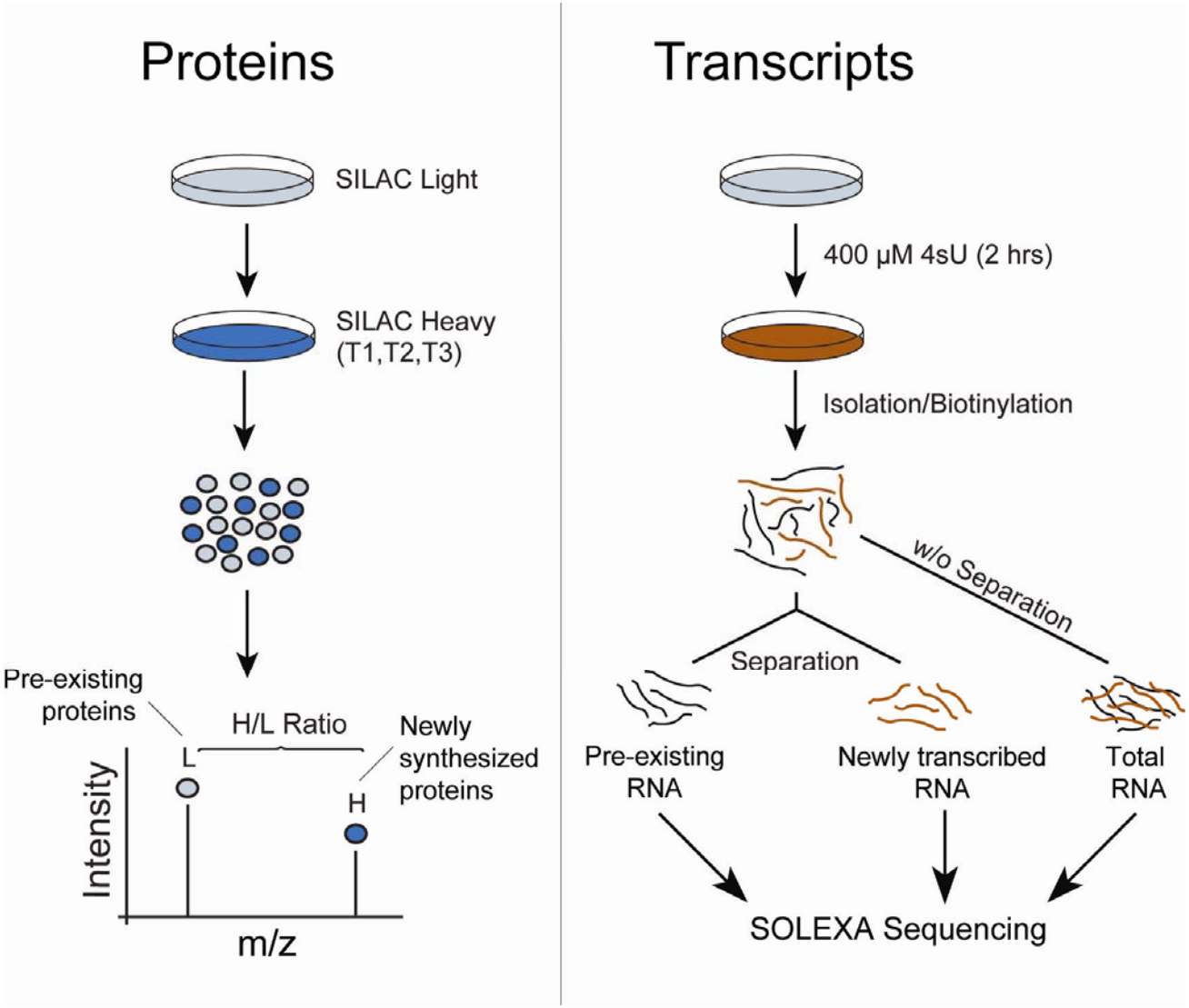
Dölken L, Ruzsics Z, Rädle B, Friedel CC, Zimmer R, Mages J, Hoffmann R, Dickinson P, Forster T, Ghazal P, Koszinowski UH. Max von Pettenkofer-Institute, Ludwig Maximilians-University Munich, Munich 80337, Germany. doelken@mvp.uni-muenchen.de

Dynamic regulation of gene expression

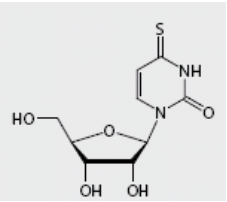


➢Dynamic properties of gene expression depend on protein and mRNA levels and half-lives

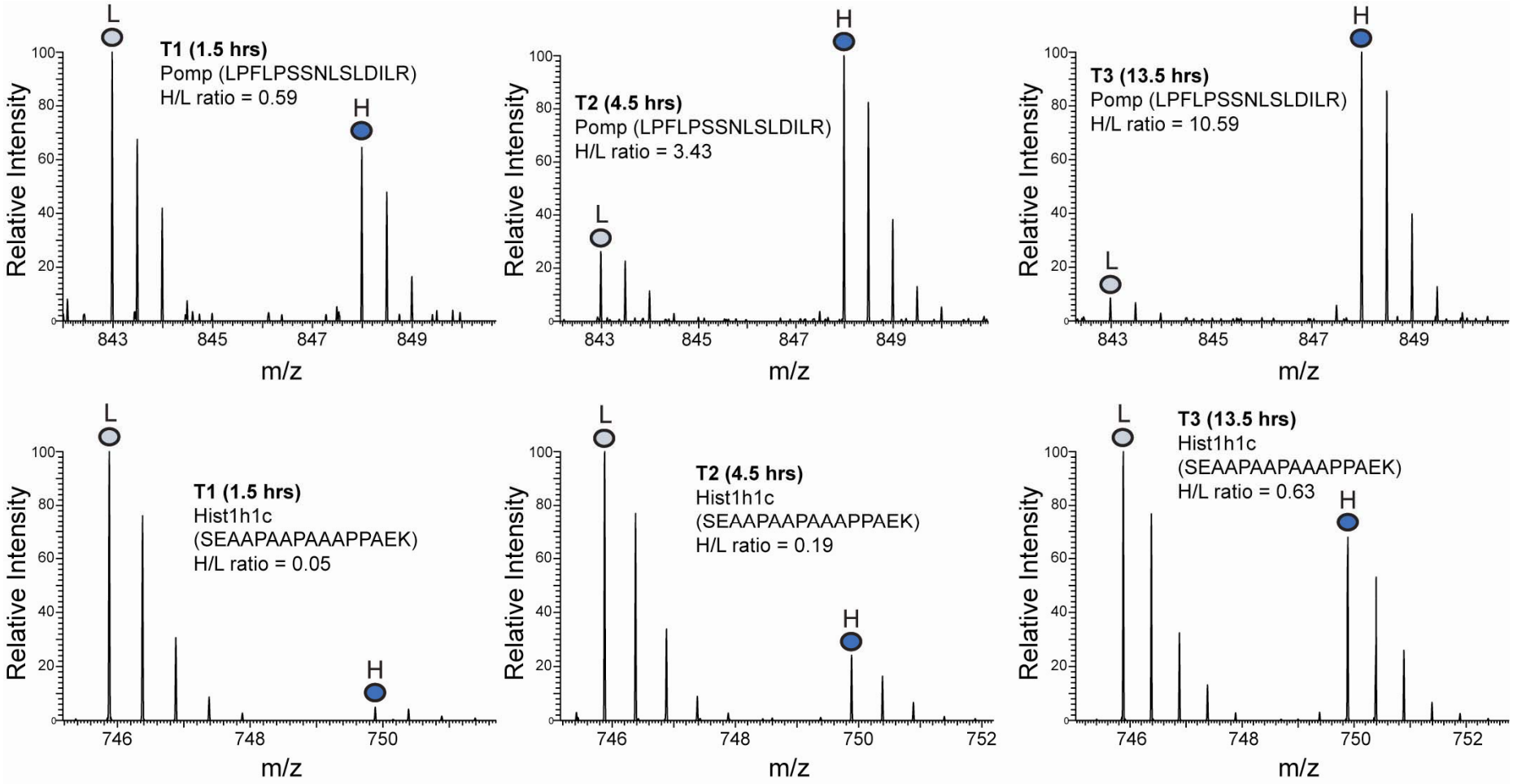
Dynamic properties of mammalian gene expression: mRNA and protein levels and half-lives



4-Thiouridine



Metabolic pulse labeling of proteins



Summary

- parallel metabolic pulse labeling with SILAC and 4SU can be used to assess protein and mRNA turnover at the same time (unperturbed system, no inhibitors)

- deep sequencing and mass spectrometry can quantify protein and mRNA half-lives and absolute cellular protein and mRNA copy numbers (iBAQ and RPKM)

- while mRNA and protein levels are correlated, half-lives are very different

- important consequences for dynamic properties of gene expression
    - energy constraints
    - posttranscriptional regulation (like RNA-binding proteins)
    - dynamic behavior

Thank you!

Dorothea Busse
Jana Wolf

Wei Chen
Na Li

Björn Schwanhäusser

www.mdc-berlin.de/selbach