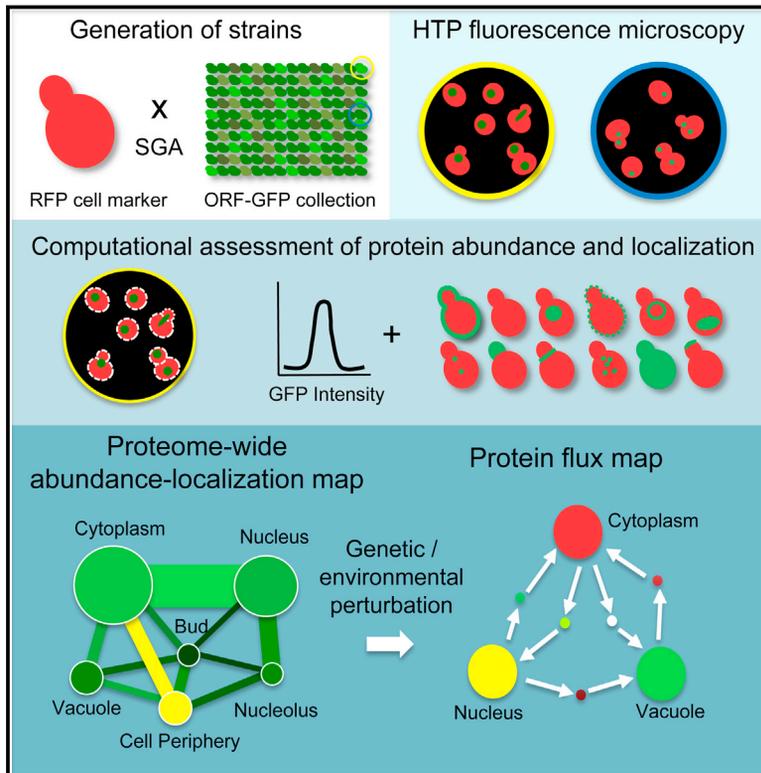


Yeast Proteome Dynamics from Single Cell Imaging and Automated Analysis

Graphical Abstract



Authors

Yolanda T. Chong, Judice L.Y. Koh, ..., Charles Boone, Brenda J. Andrews

Correspondence

brenda.andrews@utoronto.ca

In Brief

High-content, high-throughput imaging and automated analysis enables quantification of the abundance and localization of the GFP-tagged ORFeome in yeast, including how protein dynamics shift over time and respond to different perturbations.

Highlights

- Combined SGA technology, high throughput imaging, and automated analysis using SVM
- Quantified in vivo abundance and localization of $\sim 4,100$ GFP fusion proteins
- Identified proteome changes over time in rapamycin, hydroxyurea, and *rpd3 Δ* strain
- Images and abundance and localization data for all screens in CYCLOPs database



Yeast Proteome Dynamics from Single Cell Imaging and Automated Analysis

Yolanda T. Chong,^{1,4,6} Judice L.Y. Koh,^{1,4,7} Helena Friesen,¹ Kaluarachchi Duffy,^{1,2} Michael J. Cox,^{1,2} Alan Moses,³ Jason Moffat,^{1,2,5} Charles Boone,^{1,2,5} and Brenda J. Andrews^{1,2,5,*}

¹The Donnelly Centre

²Department of Molecular Genetics

³Department of Cell and Systems Biology

University of Toronto, Toronto, ON M5S3E1, Canada

⁴Co-first author

⁵Co-senior author

⁶Present address: Cellular Pharmacology, Discovery Sciences, Janssen Pharmaceutical Companies, Johnson & Johnson, 30 Turnhoutseweg, Beerse 2340, Belgium

⁷Present address: Cancer Therapeutics and Stratified Oncology, Genome Institute of Singapore, Agency for Science, Technology and Research (A*STAR), 60 Biopolis Street, #02-01 Genome, Singapore 138672, Singapore

*Correspondence: brenda.andrews@utoronto.ca

<http://dx.doi.org/10.1016/j.cell.2015.04.051>

SUMMARY

Proteomics has proved invaluable in generating large-scale quantitative data; however, the development of systems approaches for examining the proteome in vivo has lagged behind. To evaluate protein abundance and localization on a proteome scale, we exploited the yeast GFP-fusion collection in a pipeline combining automated genetics, high-throughput microscopy, and computational feature analysis. We developed an ensemble of binary classifiers to generate localization data from single-cell measurements and constructed maps of ~3,000 proteins connected to 16 localization classes. To survey proteome dynamics in response to different chemical and genetic stimuli, we measure proteome-wide abundance and localization and identified changes over time. We analyzed >20 million cells to identify dynamic proteins that redistribute among multiple localizations in hydroxyurea, rapamycin, and in an *rpc3Δ* background. Because our localization and abundance data are quantitative, they provide the opportunity for many types of comparative studies, single cell analyses, modeling, and prediction.

INTRODUCTION

The regulation of most biological processes involves changes in protein abundance and localization. However, because of an absence of quantitative data describing global protein localization and abundance, our understanding of proteome responses to perturbations remains rudimentary. Methods for rapid acquisition of in vivo quantitative data about the dynamic proteome are needed to enable the identification of members of protein complexes and coregulated proteins, give global information

about post-translational modifications, protein stability and kinetics, and suggest dependency relationships.

One of the most comprehensive reagent sets available for proteome-wide surveys of cell biological phenotypes such as protein localization is the budding yeast open reading frame (ORF)-GFP collection (Huh et al., 2003). Each strain in the ORF-GFP collection carries a unique fusion gene construct in which an ORF is fused to the GFP gene, generating a full-length protein with a COOH-terminus GFP fusion, whose expression is driven by the endogenous ORF promoter. The ORF-GFP collection contains ~4,100 strains (of a possible 5,797; *Saccharomyces* Genome Database <http://www.yeastgenome.org/>) that gave a GFP signal above background in standard growth conditions and for the majority of proteins, the GFP moiety appears to have little effect on protein function (Huh et al., 2003).

The ORF-GFP fusion collection has been analyzed using wide-field microscopy and manual inspection to assign ~70% of the proteome to subcellular locations (Huh et al., 2003). Additional studies have examined the collection in various conditions and several have quantified abundance of the GFP-tagged proteins (Benanti et al., 2007; Breker et al., 2013; Mazumder et al., 2013), but all have relied on manual assessment of localization, which is non-quantitative and suboptimal since assignments made by different individuals can be subject to biases and often show poor agreement (for review, see Chong et al., 2012).

A few recent studies have used automated approaches to analyze protein localization in yeast. Unsupervised clustering of subcellular localization patterns identified colocalized proteins (Handfield et al., 2013; Loo et al., 2014) and a classification approach assessed a mixture of spatial patterns to identify changes in protein localization in a microchemostat array (Dénervaud et al., 2013). However, while these approaches are information-rich, they have not attempted to computationally assign a specific localization for each protein and thus are of limited use to biologists. Here, we describe a high-content screening and machine learning approach to measure protein abundance and localization changes in a systematic and quantitative fashion on a genome scale.

RESULTS AND DISCUSSION

Automated Assessment of Protein Abundance and Localization

Our goal was to create an automated platform to quantify the abundance and localization of the ~4,100 visible fusion proteins in the ORF-GFP collection. To do this efficiently, we used yeast synthetic genetic array (SGA) technology (Tong et al., 2001) to introduce a cytosolic red fluorescent protein (RFP), a marker of cell boundaries, into the ORF-GFP array (a list of all strains assayed is shown in Table S1). SGA technology allows introduction of any mutation of interest into the array, thus enabling us to combine high throughput genetics and imaging. Our experimental pipeline also includes high-throughput microscopy, automated image analysis, and pattern classification through machine learning. To develop and validate our screening protocol, we imaged our wild-type collections during early log-phase growth, capturing data for >200 cells on average for each individual strain (~3 million cells total, $n = 3$). Cell boundaries were defined in the red channel and features from both the red and the green channel were extracted for each cell using CellProfiler (Carpenter et al., 2006) (Experimental Procedures; Supplemental Experimental Procedures; CellProfiler pipeline available at <http://cyclops.ccb.utoronto.ca/DOWNLOAD/Download.html>).

As an initial benchmark of the quality of our data, we extrapolated protein abundance from the mean GFP intensity (I_g) for each protein in the collection (Table S1A–S1D). Our measurements agreed with other studies that estimated protein abundance in strains from the original ORF-GFP collection using flow cytometry ($r \sim 0.9$ for 2,172 proteins; Figure S1A) (Newman et al., 2006), western blot analysis of TAP-tagged strains ($r \sim 0.64$ for 2,595 proteins; Figure S1B) (Ghaemmaghami et al., 2003), and mass spectrometry ($r \sim 0.66$ for 2,684 proteins; Figure S1C) (Kulak et al., 2014), and GFP intensities across all of our biological replicates were highly correlated ($r > 0.98$; Figure S1D). These findings validate our image acquisition and feature extraction pipelines for studying global protein abundance at the single cell level.

To extract information about protein localization patterns from our images, we adopted a machine learning approach. Individual cells were segmented based on their cytosolic RFP and texture measurements from both the RFP and GFP channels were extracted. These measurements, from a representative subset of strains known to localize to single subcellular compartments (Huh et al., 2003), were used to generate training sets for automated classification of remaining proteins. To discriminate among multiple compartments, we found that we needed an ensemble of 60 binary classifiers, which we developed from training sets that comprised >70,000 cell instances hand-picked from images from one of our wild-type screens (Data S1; Supplemental Experimental Procedures). We were able to computationally distinguish among 16 subcellular compartments (Supplemental Experimental Procedures); examples of cells with an ORF-GFP corresponding to each localization are shown in Figure 1A. The classifiers were used to generate a set of 16 quantitative localization scores (LOC scores) for each protein that reflects the proportion of cells assigned to each of the 16 defined subcellular compartments under a given condition

(Experimental Procedures; Table S2A). We assessed the performance of our classifiers in several ways. First, we compared our computationally derived pattern assignments to visually assigned localization annotations (Huh et al., 2003) and found 94% agreement among the set of 1,097 proteins assigned to a single compartment by both methods (Figure S1E, upper). Second, we assessed localization of proteins assigned to one or more of seven major localization classes covering 91% of the proteome by comparison to the manual assessment of the yeast ORF-GFP collection (Huh et al., 2003) and found that our computational approach achieved >70% in recall and precision (Figures S1E, lower, and S1F).

We explored instances in which our classifier ensemble incorrectly assigned a localization. We found that misclassified proteins tended to be lower abundance (median I_g for correctly classified proteins = 1.1×10^{-3} , median I_g for incorrectly classified proteins = 6.6×10^{-4} ; $p < 2.2 \times 10^{-16}$, Wilcoxon rank sum test). Most false positives also reflected mis-assignment of localizations to compartments that look highly similar. For example, the vacuole and nucleus are both large round compartments and our computational approach misclassified a fraction of vacuolar proteins as localizing to the nucleus or to both the nucleus and the vacuole (~35% of the proteins assigned to the vacuole in Huh et al., 2003) (Figure S1E, lower). In contrast, our automated imaging approach had high precision for nucleus (~70%) with only 0.7% of our nuclear proteins assigned to the vacuole in Huh et al., 2003). Likewise, small punctate compartments are difficult to distinguish computationally and we observed occasional proteins misclassified as cortical patches, spindle pole, peroxisome, or nucleolus. The original visual assignments of the ORF-GFP collection were validated using secondary assays for 700 proteins whose localization was difficult to determine (Huh et al., 2003); our computational approach achieved a relatively high level of precision and recall based on a single experiment using automated analysis (Figures S1E and S1F).

To display the results from our genome-wide screens in a way that would facilitate the visualization of functional information about the proteome, LOC scores, and protein abundance information were extrapolated from GFP signals to generate an “abundance localization map” or ALM (Figure 1B; ALMs 1–18 in Data S2). The wild-type ALM consists of 2,834 proteins (nodes) connected to one or more of 16 localization classes (hubs). Because of auto-fluorescence of yeast cells, our approach could not distinguish between cells with protein of very low abundance and a complete absence of the protein (see Experimental Procedures); 646 proteins with abundance below our threshold are not included in the ALM (Table S1) as well as 664 proteins that were below the threshold for significant LOC scores. In addition to localization information, absolute protein abundance measurements are represented in the ALM by node color (Figure 1B). A large cohort of proteins (1,029) localized to both the cytoplasm and the nucleus; in the ALM, these are divided into two groups: proteins whose localization is predominantly cytosolic and proteins whose localization is predominantly nuclear.

By representing these data as an ALM, we are able to visualize proteins with a common localization. The expanded views of

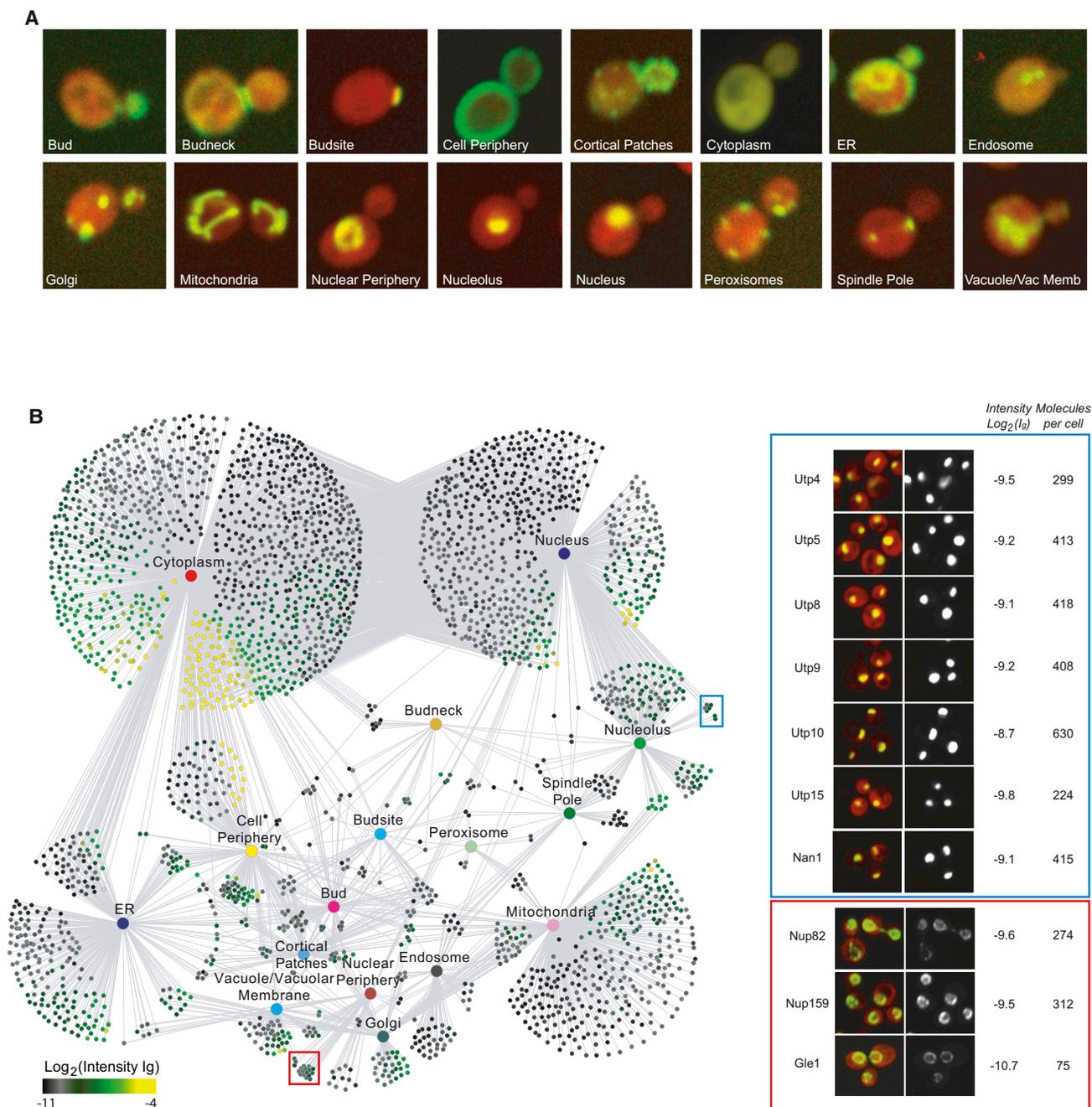


Figure 1. Abundance Localization Map of a Wild-Type Strain

(A) Images of single cells, showing typical proteins assigned by our classifiers to each localization. Each strain produces a cytosolic RFP to mark cell boundaries and a different GFP-tagged protein.

(B) Left: abundance Localization Map (ALM) illustrating protein abundance and localization information for 2,834 proteins (nodes) connected to one or more of 16 colored hubs that represent distinct cellular locations. Absolute protein abundance measurements are represented on a gray-green-yellow scale for each protein. Right: expanded views of components of the network are shown in blue and red boxes. Micrographs of cells expressing proteins are shown, whose localization is typical of the assigned compartment (nucleolus [blue box] and nuclear periphery [red box]) with protein abundance (in molecules per cell), estimated from the GFP intensity values.

See also [Figure S1](#), [Tables S1](#), [S2](#), and [S3](#), and [Data S2](#).

components of the network in [Figure 1B](#) illustrate members of known protein complexes that colocalize, with representative micrographs. For example, our automated classifiers correctly

assigned all seven components of the t-UTP complex to the nucleolus ([Figure 1B](#), blue box) and 25 of the 27 components of the Nup82 nuclear pore complex available in the GFP

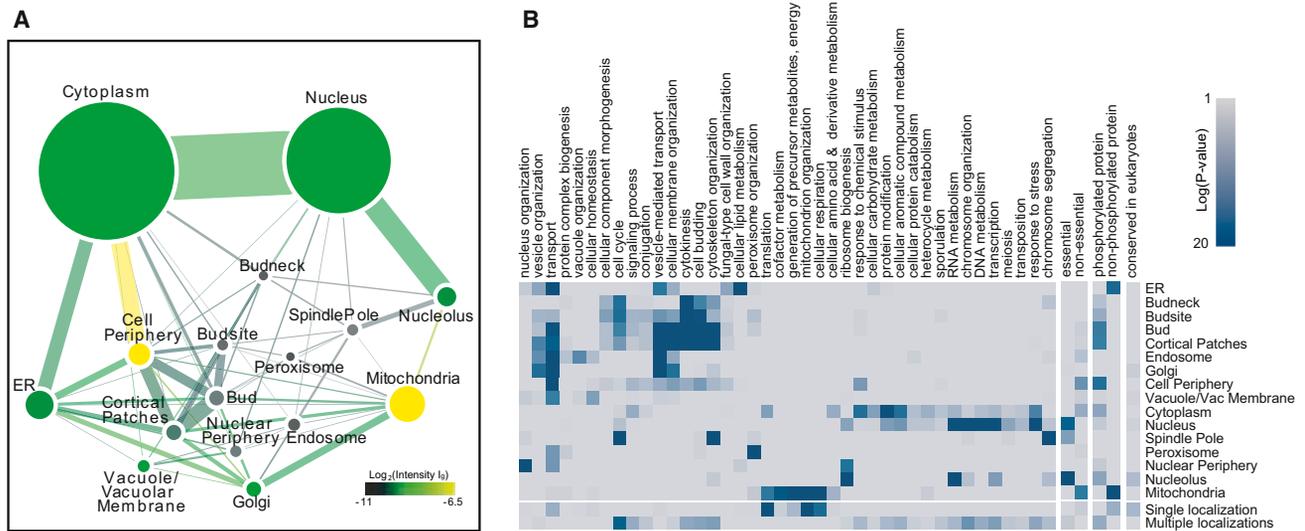


Figure 2. Thematic Diagram and Functional Enrichment Analysis of Shared Protein Localization

(A) A summary network of protein localization based on the data illustrated in Figure 1B is shown.

(B) Functional enrichment analysis of proteins assigned to various cell localizations using our automated classifiers. Log (p values) indicating the significance of the functional enrichments are color-coded (key at top of graph).

See also Figure S2.

collection (Aitchison and Rout, 2012) to the nuclear periphery (some of which are shown in Figure 1B, red box). Two components of the Nup82 complex, Sec13, and Nup116, appeared to behave differently because they were not assigned to the nuclear periphery; our computational localization assignments for these proteins correspond with published manual annotations (Huh et al., 2003). These examples illustrate the utility of ALMs for providing *in vivo* validation of protein complexes identified through biochemical or other means (see below for more discussion).

For the purposes of visualizing broader trends in our data, we summarized the ALM into a network where each compartment is represented by a node whose size indicates the number of proteins with that localization (Figure 2A). The edges connecting nodes represent dual occupancy of proteins to more than one subcellular compartment and edge thickness indicates the number of proteins that are shared between compartments (Figure 2A). Node and edge colors reflect mean protein abundance (Figure 2A; Data S2). By depicting our ALM data in this manner, we are able to easily visualize the proportion of shared localization patterns between various subcellular compartments, as well as the average abundances of shared proteins. For example, the highest fraction of proteins with multiple localizations were those shared between the nucleus and the cytoplasm (67% of 1,538 colocalized proteins), consistent with the well-characterized shuttling of many proteins involved in transcription, RNA metabolism, and DNA repair (Aitchison and Rout, 2012). We saw strong colocalization of proteins to cortical patches, bud, budsite, and cell periphery (38% of the 264 localization assignments to these four compartments), largely due to the localization of cell polarity proteins associated with the actin cytoskeleton to each of these compartments at different times in the cell cycle.

We saw enrichment of proteins with annotated roles associated with specific organelles in cognate cellular compartments, e.g., proteins with annotated roles in cytoskeletal organization, budding, and cytokinesis were highly enriched in the bud-associated compartments and cortical patches (Figure 2B). Essential proteins were enriched in nucleus, nucleolus, and spindle pole and were more frequently associated with multiple localization classes (Figure 2B) and physical interactions among proteins were enriched significantly within each organelle except for the cytoplasm (Figure S2A). Finally, in agreement with previous studies (Huh et al., 2003), the majority of the localizations were to the cytoplasm (37%) and nucleus (28%; Figures 1B, 2A, and S2B), with most of these proteins assigned to both classes (Figures 2B and S2C). However, even outside of the shared nuclear and cytoplasmic localizations, 18% of the proteome (511 proteins) showed multiple localizations in up to five different compartments (Figures S2C and S2D), emphasizing that the proteome is far more complex than a mutually exclusive organization of subcellular units. Proteins that localize to multiple compartments were enriched in roles that involve regulation, such as cell cycle, signaling, cytokinesis, and budding and were more likely to be phosphoproteins, consistent with a role for phosphorylation in regulating subcellular movement (Figure 2B).

Analysis of total GFP fluorescence also provided useful proteome-wide information about protein abundance. For example, the most abundant proteins in the cell were enriched in the cytoplasm (some of which colocalized at the cell periphery) and within the ER, Golgi, vacuole/vacuolar membrane, and nucleolus, which function in part to maintain the integrity of the intracellular environment (Figures 2A and S2E). Indeed, highly abundant proteins were enriched for roles in cell homeostasis (vacuole/vacuolar membrane, $p < 5.9 \times 10^{-5}$)

and transport (cell periphery, $p < 4.5 \times 10^{-8}$; ER, $p < 2.0 \times 10^{-6}$; Golgi, $p < 2.7 \times 10^{-5}$), as well as translation (cytoplasm, $p < 3.4 \times 10^{-69}$) and metabolism (cytoplasm, $p < 1.5 \times 10^{-6}$) (Figures 2A and S2F). Our automated analysis allowed us to distinguish proteins that varied in concentration over three orders of magnitude, from an estimated lower limit of ~ 40 molecules per cell for Are1, an ER protein that was undetectable by western blot as a TAP-tagged protein, to an upper limit of $\sim 19,000$ molecules per cell for Gln1, glutamine synthase, one of the highest abundance proteins in the TAP data set (Ghaemmaghami et al., 2003) (Figure 1B; Table S1).

Our data set can also be mined using more detailed single cell analyses. For example, we compared single cells containing two GFP-tagged proteins that are known to relocate between nucleus and cytoplasm in a cell cycle-dependent manner, Mcm2 (Yan et al., 1993) and Whi5 (Costanzo et al., 2004), and two proteins that are localized to the nucleus but are degraded in a cell cycle-specific manner, Sic1 (Visintin et al., 1998) and Far1 (Blondel et al., 2000). For each protein, we plotted the fraction of single cells assigned to cytoplasm, nucleus, or whose protein was below our threshold of detection for abundance (Supplemental Experimental Procedures). The Mcm2 and Whi5 proteins in all of the single cells were classified as localizing to either nucleus or cytoplasm (Figure S2G) with a larger fraction of cells containing nuclear Mcm2 than nuclear Whi5, consistent with their cell-cycle biology. In contrast, for Sic1 and Far1, most of the cells that had detectable protein were classified as nuclear, while the majority of cells had undetectable protein levels (Figure S2G), suggesting that the protein had been degraded. Thus, processing single cell data offers the potential to reveal complex regulatory mechanisms governing protein localization and abundance.

The increased sensitivity afforded by our approach allowed us to reliably assign a quantitative localization for 52 of the 156 proteins previously annotated as “ambiguous” (Table S3). For example, we confidently assigned Itr2, a myo-inositol transporter, Yeh1, a sterol ester hydrolase, and Bor1, a boron efflux carrier (<http://www.yeastgenome.org/>), to the cell periphery. Orthogonal functional genomic data sets supported our novel assignments of several proteins to subcellular compartments. We assigned the uncharacterized protein Ymr010w to the Golgi, and *YMR010W* shows strong synthetic lethal interactions with genes involved in transport/retention of proteins in the ER and Golgi (Costanzo et al., 2010; Koh et al., 2010). We also assigned Yir003w (Aim21) to cortical patches, and *YIR003W* has many genetic interactions with known actin patch genes (Costanzo et al., 2010). Together, these data suggest that our classifiers efficiently capture localization patterns and abundance information for most proteins and that these localizations can reveal important biology and contribute to predictions of protein function.

Quantitative Analysis of Proteome Abundance Changes in Response to Environmental and Genetic Perturbations

We exploited the ALMs to explore the response of the proteome to either environmental or genetic perturbation in an unbiased and quantitative fashion. We chose two chemical perturbations—rapamycin and hydroxyurea (HU) treatment—both of which compromise yeast proliferation and have been used

extensively as reagents to study nutrient deprivation and DNA replication stress, respectively. Rapamycin is an inhibitor of the TORC1 kinase complex, which controls cell growth by affecting protein synthesis and metabolism (Loewith and Hall, 2011), while HU inhibits DNA replication through deoxyribonucleotide depletion (Alvino et al., 2007).

As a proof-of-principle for analysis of a genetic perturbation, we also examined proteome abundance and dynamics in a mutant lacking the catalytic subunit of a lysine deacetylase (KDAC), Rpd3 (Yang and Seto, 2008), which we had previously analyzed in systematic genetic screens (Kaluvarachchi Duffy et al., 2012). Rpd3-mediated lysine deacetylation is a dynamic posttranslational modification with a well-defined role in negatively regulating gene expression through modification of histones. Rpd3 also deacetylates many non-histone proteins, regulating properties such as protein stability and protein-protein interactions (Kaluvarachchi Duffy et al., 2012; Henriksen et al., 2012). Our quantitative approach is ideal for identifying changes in proteins that are not related to transcriptional changes.

Because of the automated nature of our analysis, we were able to examine 15 different screening conditions (three times for HU treatment; nine times for rapamycin treatment; three *rdp3Δ* screens) and extract features from a total population of >20 million cells. With our classifiers, we assigned localization patterns for the GFP-fusion proteins expressed in $\sim 95\%$ of cells imaged. Our compendium of high-resolution images as well as quantitative information on protein abundance and localization in both wild-type and perturbed cells is available as a searchable web-accessible interface called Collection of Yeast Cells and Localization Patterns (CYCLOPs) (<http://cyclops.ccb.utoronto.ca>). Protein abundance information revealed 674 unique changes across the 15 conditional screens (Table S4; proteins with abundance changes using a stringent 2-fold cutoff (δ PL); δ PL < -1 (red) or δ PL > 1 (green) shown in Figure 3 and Table S5). Importantly, we detected protein abundance changes consistent with known biology and with other studies. For example, deletion of *RPD3* largely resulted in increases in protein abundance (Figure 3), consistent with the well-characterized role of Rpd3 as a repressor of gene expression. We also observed 22 proteins, including several ribosome biogenesis and ribosomal proteins, that were downregulated in rapamycin and upregulated in *rdp3Δ* (Figure 3, center right). These findings reflect that inhibition of TORC1 by rapamycin leads to repression of Ribi (ribosome biogenesis) and RP (ribosomal protein) coding genes in a manner dependent on Rpd3 (for review, see Broach, 2012). Inhibition of TORC1 by rapamycin also causes repression of energy-consuming processes (Dechant and Peter, 2008)—our imaging assay showed a time-dependent decrease in abundance of regulators of protein translation by rapamycin (Figures 3, bottom right, and S3A) while levels of proteins involved in vacuolar protein degradation (e.g., Prc1 peptidase) and energy regeneration (e.g., Cox4 subunit of cytochrome c oxidase) were increased (Figures 3, top right, and S3A). Interestingly, there was very little overlap between our HU and rapamycin screens with respect to protein abundance changes (Figure 3). This difference largely appears to reflect activation of the Environmental Stress Response (ESR) (Gasch et al., 2000) following rapamycin treatment and growth inhibition (209/355

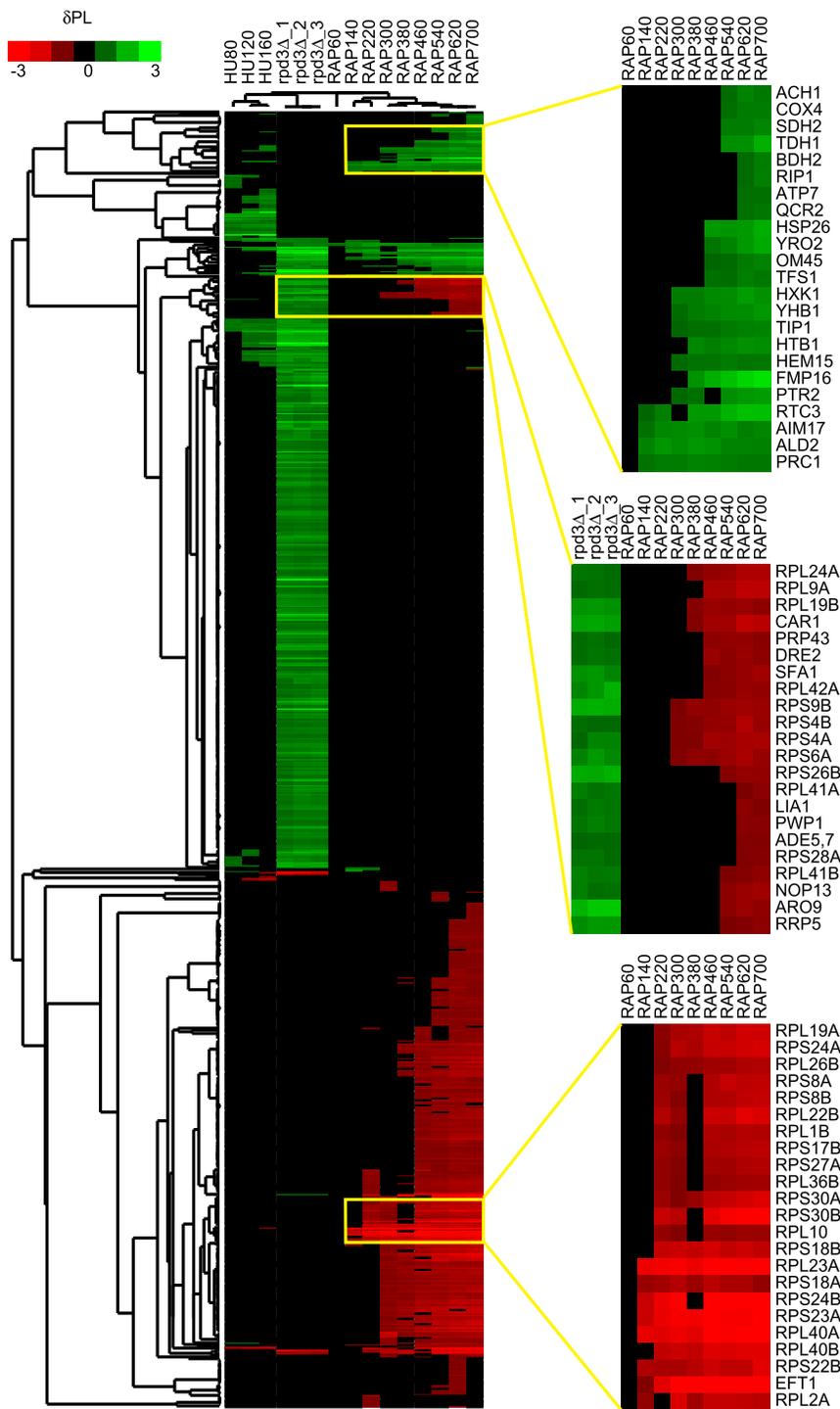


Figure 3. Protein Abundance Changes following Perturbations Detected Using Automated Imaging

Hierarchical clustergram of GFP proteins showing a >2-fold change in protein level (GFP intensity). δ PL values were calculated from treated over untreated data sets and clustered using average linkage with uncentered correlation. Parts of the clustergram showing distinct trends in the data are boxed in yellow. To the right are shown expanded views of the boxed regions, illustrating groups of proteins that have increased abundance in rapamycin (top), increased abundance in *rpd3* Δ together with decreased abundance in rapamycin (middle), and decreased abundance in rapamycin (bottom). See also Figure S3 and Tables S4 and S5.

rapamycin causes inhibition of growth (Loewith and Hall, 2011), which is associated with the ESR (Brauer et al., 2008).

The changes in protein abundance in cells treated with rapamycin correlated closely with published protein abundance changes assessed by mass spectrometry (Fournier et al., 2010) (Figure S3B, upper). To ask if protein abundance changes correlated with changes in transcript levels, we compared the protein abundance values from our rapamycin experiment with published data sets measuring mRNA levels. We saw that rapamycin caused a delayed relationship between mRNA and protein levels with maximal correlation between mRNA expression at 120 min and protein levels at 620 min of rapamycin treatment ($R = 0.7$), similar to the findings by mass spectrometry (Fournier et al., 2010) (Figure S3B, lower).

To ask if transcript and protein abundance were also correlated following a genetic perturbation, we considered any protein whose abundance changed detectably in the *rpd3* Δ strain (no stringent cut-off applied). Of all the proteins that increased in abundance, 79% were associated with genes whose transcript levels increased in the absence of *RPD3* ($p < 3.3 \times 10^{-12}$) (Figure S3C) (Fazzio et al., 2001), and this set of proteins was enriched for components of the cellular

of total changes occurred in proteins encoded by ESR genes; Table S5) but not following HU treatment (only 28/91 total changes occurred in proteins encoded by ESR genes; Table S5) (Dubacq et al., 2006). This difference may be explained by the effect of HU versus rapamycin on growth of the cells: HU treatment activates the intra-S phase checkpoint, causing cycling cells to accumulate in S/G2 (Alvino et al., 2007), whereas

stress response (Figure S3A), consistent with Rpd3 biology (Alejandro-Osorio et al., 2009). While we see generally good agreement between our protein levels and published mRNA levels, proteins that change in abundance in the absence of changes in the corresponding transcript are of particular interest as candidates for post-translational regulation. For example, of the seven proteins that changed in abundance >5-fold with no

corresponding change in published mRNA levels (Table S4) (Fazio et al., 2001), three have been found to be hyperacetylated in an *rpm3Δ* strain by mass spectrometry (Henriksen et al., 2012): Pbi2, Tsl1, and Hyr1. Acetylation of these proteins may be regulating their abundance, consistent with known roles for acetylation in negatively regulating protein stability in yeast (Robert et al., 2011). Because our data are quantitative, we can identify unexpected changes in the proteome, even with a high background of protein abundance changes resulting from Rpd3 repressing transcription.

Flux Networks Map Localization and Abundance Changes in Environmental and Genetic Perturbations

As described above, ALMs provide a useful means of displaying proteome-scale information about protein abundance and localization in living cells. We reasoned that comparative analysis of ALMs derived from perturbed cell populations would enable us to examine protein dynamics in an unbiased and quantitative manner. To do this, we used LOC scores derived from wild-type and perturbed ALMs to define a “z-LOC score.” The z-LOC score quantifies changes in protein localization between sub-cellular compartments; a negative z-LOC value denotes the exit of a protein from a particular compartment, while a positive z-LOC value represents movement toward a specific location. For example, in the *rpm3Δ* strain, Pct1, an enzyme involved in phospholipid biosynthesis (Howe et al., 2002), showed a negative z-LOC score for the nuclear periphery and a positive z-LOC score for the nucleus, indicating displacement from the nuclear periphery to the nucleus (Figures 4A–4D).

To display z-LOC score data in a biologically useful way, we generated “flux networks,” which are directional, quantitative vectors that define the relationship between two ALMs and represent the dynamic localization and abundance changes following a perturbation. In the flux networks, large nodes represent subcellular locations while small nodes are moving proteins, whose abundance relative to unperturbed cells is color-coded. The vectors illustrate the direction of movement of the protein and the thickness of the vector corresponds to the proportion of cells in the population with the indicated change in localization. The flux networks for our *rpm3Δ*, HU, and rapamycin experiments are shown in Figures 4C, S4A, and 5A. Some proteins, such as Dig2, a repressor of mating-specific gene expression (Figures 4C and 4D), are computationally scored with a specific localization pattern only in the perturbed situation (Dig2 retains some cytoplasmic localization but is enriched in the nucleus in the *rpm3Δ* strain). In this case, the vector on the *rpm3Δ* flux network shows movement into a compartment with no specific departure localization (Figure 4C). Interestingly, both Pct1 (see above) and Dig2 are acetylated proteins in vivo and both are hyperacetylated in an *rpm3Δ* strain (Henriksen et al., 2012), suggesting that Rpd3-mediated deacetylation may regulate their localization. In general, flux networks provide a quantitative snapshot of protein changes within the cell following environmental or genetic perturbation.

A recent screen visually identified localization changes in the GFP collection treated with HU (Tkach et al., 2012), allowing us an opportunity to benchmark our computational method for detecting protein localization changes. Our computational analysis

identified 54 of the 97 proteins for which we could visually identify the changes reported by Tkach et al. in our images (Figure S4A; Table S6A; Supplemental Experimental Procedures). For another 36 of the 97 proteins, our computational approach identified the relevant localization in wild-type and HU but the change was not statistically significant; generally for these proteins a small fraction of the cells showed the localization change. Our computational approach also identified a novel set of 40 significant localization changes that had not been scored in the previous manual screen (Figure S4A; Table S6B). The majority of these (34) were proteins that had a mixed localization pattern in wild-type (for example nucleus and cytoplasm), which translates to more subtle quantitative localization changes in HU that would be challenging to score manually. Thus our computational approach is able to efficiently and quantitatively detect changes in protein localization with high sensitivity, capturing dynamic information that cannot be reasonably scored by assessment of thousands of images by eye.

Integrating protein interaction information into flux networks highlights co-translocation of protein complexes or functionally related proteins. For instance, HU treatment resulted in relocalization of several proteins that form P-bodies or cytoplasmic granules, which sequester untranslated mRNAs under certain stresses like glucose deprivation (Teixeira and Parker, 2007; Tkach et al., 2012). We observe nine P-body proteins recruited from a diffuse cytoplasmic pattern to foci (Figures S4A and S4B), which expands upon similar observations (Tkach et al., 2012; Dénervaud et al., 2013).

Proteins Change in Localization or Abundance but Usually Not Both

Flux networks can also be used to explore the general relationship between protein abundance and localization changes. Following 80 min of HU treatment, 28 localization and 40 abundance changes ($\delta PL > 1$) were observed in the yeast proteome (Tables S2 and S4). Remarkably, only one protein, Smf3, an iron transporter, changed both localization and abundance using our criteria. Treatment with HU induces activation of the Aft regulon, which controls iron mobilization (Dubacq et al., 2006), leading to an increase in *SMF3* transcript abundance; however, relocalization of Smf3 has not been previously observed. Similar trends were seen after prolonged exposure to HU and in cells perturbed by rapamycin treatment or deletion of *RPD3*, suggesting that the response of the proteome to genetic or chemical insults involves a change in protein abundance or localization but rarely both (Figure S4C).

Our Rapamycin flux networks (Figure 5) also uncover intriguing time-dependent changes in the proteome that recapitulate known and meaningful biology. For example, our flux network displaying changes after 140 min of rapamycin treatment (RAP140) (Figure 5A) revealed translocation of transcription factors Sfp1, which promotes RP gene transcription, and Stp1, which is involved in amino acid sensing, from the nucleus to the cytoplasm. In contrast, Stb3, a repressor of RP gene expression, moved from the cytoplasm to the nucleus, reflecting the need to inhibit protein translation in starvation conditions, which are mimicked by rapamycin treatment (Figure 5A) (Broach, 2012) We also detected subcellular movement of two other

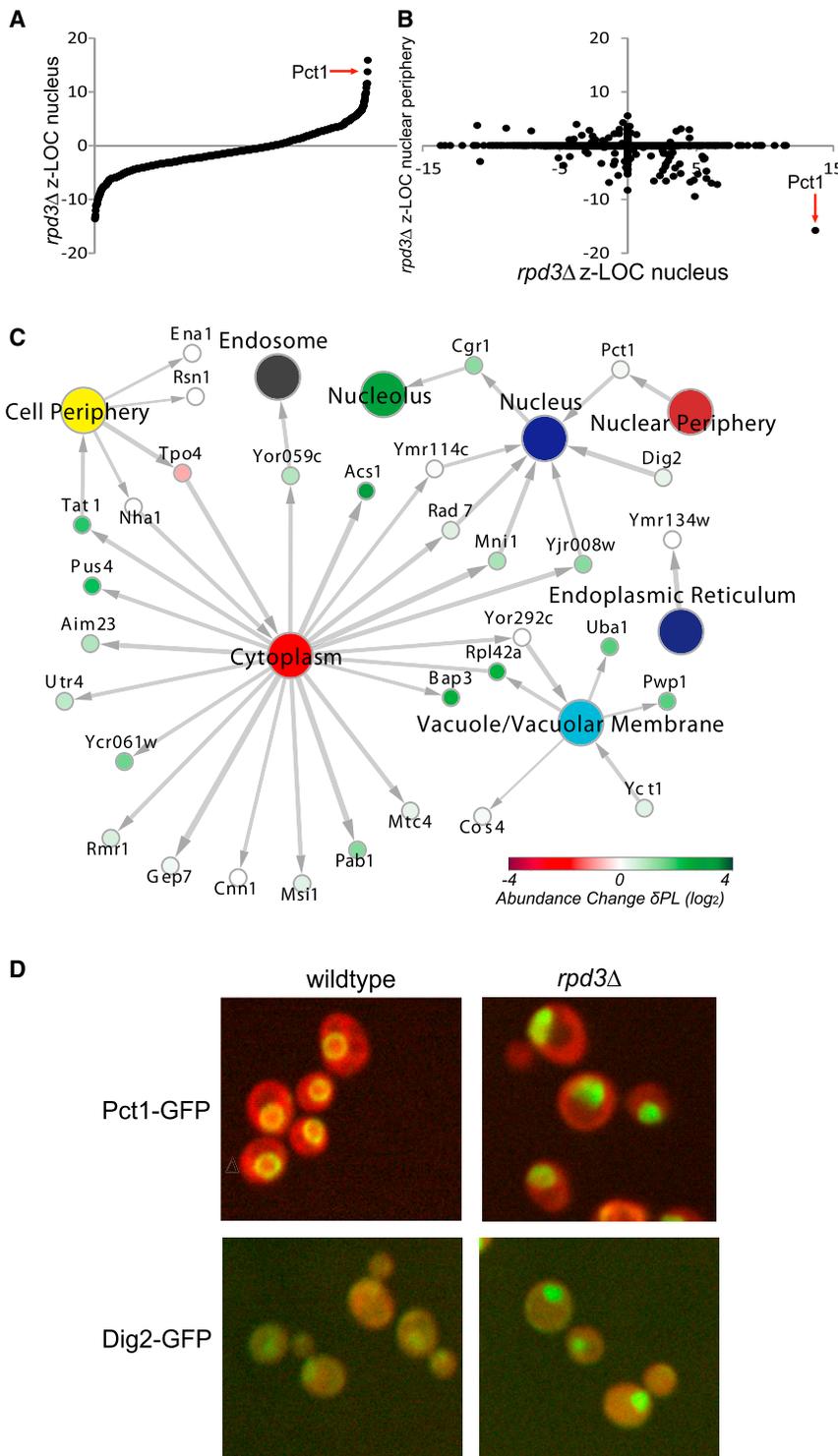


Figure 4. Automated Assessment of the Proteome in an RPD3 Deletion Mutant

(A) Changes in protein localization to the nucleus in an *rpd3Δ* strain. The graph illustrates the distribution of z-LOC scores of all GFP-fusion proteins scored using our nuclear classifier in an *rpd3Δ* strain relative to wild-type. One outlier, Pct1, with increased nuclear localization is highlighted by the arrow.

(B) Pairwise comparison between z-LOC scores for the nucleus versus those for nuclear periphery localization in *rpd3Δ* relative to wild-type. Pct1 is highlighted as an example of a protein showing obvious relocalization from the nuclear periphery to a more general nuclear localization in the *rpd3Δ* strain.

(C) Flux network illustrating abundance and localization changes in the proteome in an *rpd3Δ* strain. The wild-type and *rpd3Δ* ALMs were used to generate the flux network where large nodes represent cell locations while the small nodes are moving proteins.

(D) Representative micrographs showing Pct1-GFP and Dig2-GFP localization in wild-type and *rpd3Δ* cells.

See also Figure S4.

partments. A striking example is Rts3, a PP2A and Sit4 phosphatase-associated protein of unknown function (Gavin et al., 2006; Breitreutz et al., 2010). Rts3 redistributed from the nucleus to the cytoplasm at early time points (Figure 5B), with a re-concentration in the nucleus after prolonged rapamycin treatment, which excludes it from the flux network in Figure 5A. This type of movement is often associated with signaling molecules that respond to a change in a stimulus and then adapt and return to their steady state localization, e.g., the Hog1 MAP kinase in response to high osmolarity (Ferrigno et al., 1998). Our flux network also revealed an increase in Rts3-GFP abundance, which we confirmed using a TAP-tagged version of Rts3 (Figure 5B). The transient relocalization and the abundance change suggest that Rts3 may play a key role in response to TORC1 inhibition. Indeed, we and others have found that Rts3 is required for cell survival in rapamycin (Figure 5B) (Parsons et al., 2004; Xie et al., 2005; Hood-DeGrenier, 2011). Thus, flux network-based analysis

transcription factors, Gln3 and Gat1, whose translocation to the nucleus is a hallmark of TORC1 inhibition (Figure 5A) (Broach, 2012).

Observing cells over time following rapamycin treatment allowed us to track proteins that move transiently between com-

of the proteome over time can highlight unusual proteome dynamics in response to environmental or genetic stress and facilitate attribution of protein function.

Several proteins that translocated out of the nucleolus following exposure of cells to rapamycin for 140 min also shared physical

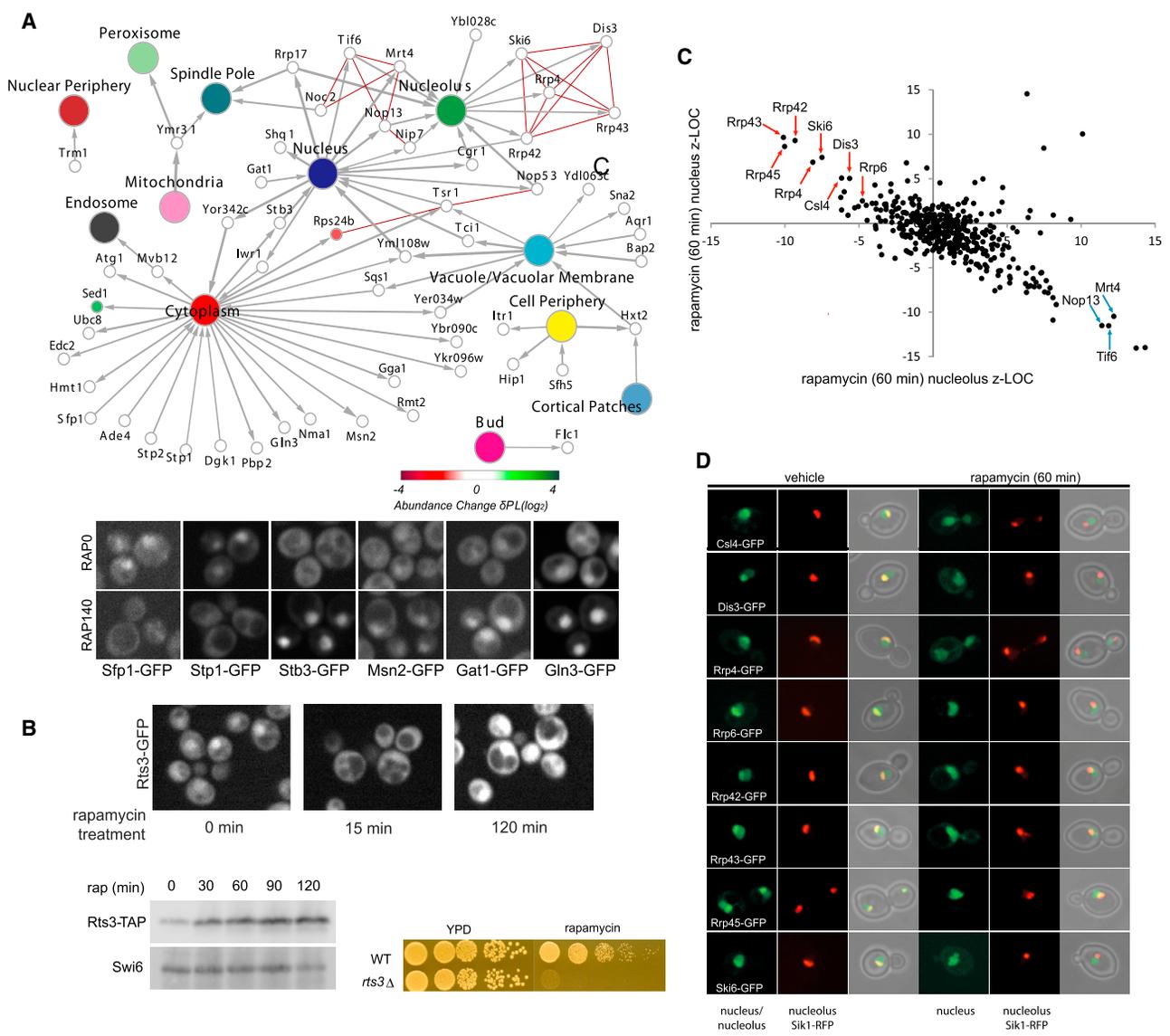


Figure 5. Proteome Dynamics in Rapamycin-Treated Cells over Time

(A) Flux network illustrating protein localization and abundance changes after 140 min of rapamycin treatment. Proteins annotated as having physical interactions (BioGrid) are indicated with red edges. Representative micrographs of cells expressing GFP-fusions of transcription factors whose localization is influenced by rapamycin treatment are shown below: untreated cells (RAP0) and cells treated with rapamycin for 140 min (RAP140).

(B) Micrographs showing transient relocation of Rts3 from largely nuclear in untreated cells to diffusely cytoplasmic within 15 min of rapamycin treatment, with a re-concentration in the nucleus after prolonged rapamycin treatment (120 min). At bottom left is a western blot showing Rts3-TAP, following treatment with rapamycin for times indicated, with Swi6 loading control. Bottom right shows spot dilutions revealing sensitivity of an *rts3Δ* strain to rapamycin.

(C) Pairwise comparison between z-LOC scores for the nucleus versus those for the nucleolus in cells treated with rapamycin for 60 min. Components of the exosome are highlighted with red arrows while proteins involved in ribosome biogenesis are highlighted with blue arrows. Rapamycin treatment induces exosome complex redistribution from being concentrated in the nucleolus to becoming dispersed throughout the nucleus as early as 60 min.

(D) Representative micrographs of untreated and rapamycin-treated cells expressing GFP-fusions to exosome components (left; green images), a Sik1-RFP nucleolar marker (middle; red images) and an overlay of the two images (right).

interactions (Figure 5A, protein interactions indicated by red edges). Rrp43, Rrp42, Rrp4, Ski6, and Dis3 are members of the exosome, a complex involved in processing of rRNA, small nucleolar RNAs (snoRNAs), small nuclear RNAs (snRNAs), tRNAs, and cryptic unstable transcripts (CUTs) (Lykke-Andersen et al., 2009). A re-distribution from the nucleolus to the nucleus was

observed for eight out of ten components of the exosome after a 60 min exposure to rapamycin (Figure 5C). We confirmed this movement of the exosome using colocalization of GFP-tagged exosome components in cells expressing the nucleolar marker Sik1-RFP (Figure 5D). Subcellular movement of the exosome during rapamycin treatment has not been observed previously,

and suggests a significant nuclear role for the complex and perhaps the need for specific RNA regulation or processing events during the cellular response to rapamycin. Thus, integrating flux networks with protein interaction data can provide new insight into protein dynamics and point to new compartment-specific functions for protein complexes.

PERSPECTIVE

Although incredibly powerful, functional genomic approaches that explore gene expression (Kemmeren et al., 2014), protein-protein interactions (Rees et al., 2011) and genetic interactions (Costanzo et al., 2010) fail to yield a spatio-temporal resolution that will be required to understand biological processes as complex dynamic systems. Here, we describe a high-content screening system that allows us to rapidly survey proteome dynamics in living cells. This approach is designed to define proteome flux using a computational image-based method and provide a proof-of-principle for both a technical and a conceptual platform that ought to be easily adapted to other systems.

This study is our first attempt at a quantitative exploration of abundance and localization of the proteome. We used a high quality, available resource, the ORF-GFP collection, which includes 71% of the full proteome, with 1,492 strains excluded that failed to yield a GFP signal above background in the original manual assessment of the collection under standard growth conditions (Huh et al. [2003] and *Saccharomyces* Genome Database <http://www.yeastgenome.org/>). Our experimental approach can be readily modified to include technical improvements that should enable more comprehensive analysis of the proteome. A number of the “missing” proteins are detectable by western blot when C-terminally tagged with the TAP moiety (344 proteins) (Ghaemmaghami et al., 2003) and peptides corresponding to another 543 proteins have been identified in at least one mass spectrometry study (Nagaraj et al., 2012; Kulak et al., 2014); these proteins are clearly produced under standard conditions. Some proteins may be undetectable due to destabilization caused by the C-terminal tag; for these, insertion of the tag at the N terminus of the protein may produce useful alleles. In general, use of yeast-optimized versions of GFP that outperform the original GFP(S65T) for brightness, photostability, and function as fusion proteins (Lee et al., 2013) is likely to allow visualization of more low abundance proteins. In addition, low abundance proteins may be detectable using a recently developed protein tag known as the SunTag, which can recruit multiple copies of an antibody-GFP fusion protein and permits imaging of single molecules (Tanenbaum et al., 2014). The remaining proteins, which are likely not expressed under standard conditions, will have to be characterized on a case-by-case basis; growth in other conditions, chosen from transcription data or that affect growth of the corresponding mutant, may enable image-based analysis of these proteins.

Automated image analysis of the proteome offers biological insight at several different levels. First, our compendium of images offers a qualitative view of the abundance and localization of each protein. In contrast to some other imaging data sets, we provide data for an average of 80 cells per image, allowing

assessment of general trends for each protein, including cell-cycle distribution and variability within a population. Second, these data reveal previously unidentified regulatory events and offer many opportunities for generation of particular hypotheses. Third, although our data set contains many interesting individual findings, we predict that this type of data will be of particular value for comparative analyses using quantitative methodologies. Comparison of image compendia from different mutant strains or conditions with relevant proteomic data sets promises to provide insight into regulated proteolysis, post-translational modifications, and the regulation of protein-protein interactions. We also demonstrate that we can use single cell measurements to identify and distinguish two types of changes in cell-cycle-regulated proteins and many other types of single-cell analysis should be possible with our data.

Finally, quantitative data sets are particularly important for the development of models for pathway activity and cellular function. Just as transcriptome data have revolutionized our understanding of gene regulation, quantitative proteomic data, with the dimension of localization added to more straightforward abundance measures, will change the way researchers understand the cellular consequences of genetic and environmental stress. We suggest that integrating condition-specific quantitative abundance and localization data for the proteome will allow even greater predictive power and represent an essential component underlying modeling approaches.

EXPERIMENTAL PROCEDURES

Generation of Yeast Strains and Growth Conditions

S. cerevisiae strains used in this experiment are listed in Table S1. Using a modified SGA protocol (Tong et al., 2001) *RPL39pr-tdTomato* and marked alleles were introduced into the arrayed yeast GFP collection (Huh et al., 2003). Haploid *MATa* strains from SGA were grown and imaged in low fluorescence synthetic medium (Sheff and Thorn, 2004) supplemented with methionine, NAT, and 2% glucose.

Image Acquisition, Analysis, and Pattern Classification

Large-scale acquisition of fluorescent micrographs was conducted using a high-throughput spinning-disc confocal microscope (Opera, PerkinElmer). For co-localization analysis, images were acquired using a spinning-disc confocal (WaveFX, Quorum Technologies) connected to a DMI 600B fluorescence microscope (Leica Microsystems) and Imagem charge-coupled device camera (Hamamatsu C9100-13, Hamamatsu Photonics).

Image analysis, segmentation, and acquisition of measurements (Figure S5A) were conducted using CellProfiler (Carpenter et al., 2006). Mean GFP intensity was used to represent the protein abundance in the cell and changes > 2-fold compared to wild-type were defined as significant. To generate the training sets, objects from the WT1 screen were visually inspected using CellProfiler Analyst (Jones et al., 2008), to select objects suitable for training different patterns (see Data S1) representing distinct locations and stages of the cell cycle. The resulting handpicked “designer” training sets consisted of over 70,000 objects. Using a supervised machine learning approach, an ensemble of 60 binary classifiers was generated to identify cells in 3 stages of the cell cycle, 16 localization patterns, ghost objects, and dead cells (Figures S5B–S5D). Classifiers were generated using SVM (Platt, 1998) (sequential optimization, PolyKernel with 1 as exponent) and improvement in accuracy was achieved through bootstrap aggregation (Breiman, 1996). An automated classification system was developed, using WEKA (Hall et al., 2009) and in-house statistical programs to generate training models from WT1 and to apply classifiers to all single cell instances across 18 screens.

Defining Changes in Localization

For a protein i , the LOC score reflects the proportion of cells that are assigned to a specific localization class j under condition k :

$$LOC_{ijk} = \frac{n_{ijk}}{\sum_j n_{ijk}}$$

For each localization a cutoff was determined such that each strain was assigned to that localization if the fraction of cells assigned to that localization passed the cutoff (Table S7).

Data Visualization

The abundance localization maps (ALMs) and flux networks were generated using Cytoscape (Cline et al., 2007). Two types of node were defined in the ALMs: 16 localization classes as “hub” nodes, and individual proteins as “protein” nodes. For the ALMs in Data S2, node positions were determined by edge-weighted spring-embedded layout, and the visual distance of the edge between the hub and protein node depicts the quantitative localization membership of the protein to the specific organelle (i.e., shorter distance for higher LOC score). For the WT1 ALM shown in Figure 1B, the positions of the localization clusters were determined by edge-weighted spring-embedded layout and protein nodes were manually shifted to more clearly show abundance groupings. The node colors in gray-green-yellow-scale are proportional to the I_0 value (e.g., darker node for more abundant protein).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, seven tables, and two data sets and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2015.04.051>.

AUTHOR CONTRIBUTIONS

B.J.A., C.B., and J.M. designed and supervised the project. Y.T.C., M.J.C., and S.K.D. carried out and analyzed experiments. J.L.Y.K., Y.T.C., H.F., and A.M. performed large-scale analysis and interpretation. B.J.A., C.B., J.M., J.L.Y.K., and Y.T.C. created figures. B.J.A., C.B., J.M., Y.T.C., J.L.Y.K., and H.F. wrote the manuscript.

ACKNOWLEDGMENTS

We thank Jeff Liu, David Wade-Farley, Quaid Morris, Lee Zamparo, Harm Van Bakel, Kyle Tsui, and Corey Nislow for technical assistance and advice. We thank Michael Costanzo for helpful discussions and comments on the manuscript. This work was supported by grant MOP-97939 from the Canadian Institutes for Health Research to B.A. and C.B., and with funds from the Ontario Institute for Cancer Research to J.M. Infrastructure for high-content imaging was acquired with funds from the Canadian Foundation for Innovation and the Ministry of Research and Innovation (Ontario) (grant 21745 to B.A. and C.B.). B.A., J.M., and C.B. are Senior Fellows in the Genetic Networks program of the Canadian Institute for Advanced Research.

Received: November 7, 2014

Revised: February 5, 2015

Accepted: April 9, 2015

Published: June 4, 2015

REFERENCES

Aitchison, J.D., and Rout, M.P. (2012). The yeast nuclear pore complex and transport through it. *Genetics* 190, 855–883.

Alejandro-Osorio, A.L., Huebert, D.J., Porcaro, D.T., Sonntag, M.E., Nillasithanukroh, S., Will, J.L., and Gasch, A.P. (2009). The histone deacetylase Rpd3p is required for transient changes in genomic expression in response to stress. *Genome Biol.* 10, R57.

Alvino, G.M., Collingwood, D., Murphy, J.M., Delrow, J., Brewer, B.J., and Raghuraman, M.K. (2007). Replication in hydroxyurea: it's a matter of time. *Mol. Cell. Biol.* 27, 6396–6406.

Benanti, J.A., Cheung, S.K., Brady, M.C., and Toczyski, D.P. (2007). A proteomic screen reveals SCFGrr1 targets that regulate the glycolytic-gluconeogenic switch. *Nat. Cell Biol.* 9, 1184–1191.

Blondel, M., Galan, J.M., Chi, Y., Lafourcade, C., Longaretti, C., Deshaies, R.J., and Peter, M. (2000). Nuclear-specific degradation of Far1 is controlled by the localization of the F-box protein Cdc4. *EMBO J.* 19, 6085–6097.

Brauer, M.J., Huttenhower, C., Airoidi, E.M., Rosenstein, R., Matese, J.C., Gresham, D., Boer, V.M., Troyanskaya, O.G., and Botstein, D. (2008). Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. *Mol. Biol. Cell* 19, 352–367.

Breiman, L. (1996). Bagging predictors. *Mach. Learn.* 2, 123–140.

Breitkreutz, A., Choi, H., Sharom, J.R., Boucher, L., Neduva, V., Larsen, B., Lin, Z.Y., Breitkreutz, B.J., Stark, C., Liu, G., et al. (2010). A global protein kinase and phosphatase interaction network in yeast. *Science* 328, 1043–1046.

Breker, M., Gymrek, M., and Schuldiner, M. (2013). A novel single-cell screening platform reveals proteome plasticity during yeast stress responses. *J. Cell Biol.* 200, 839–850.

Broach, J.R. (2012). Nutritional control of growth and development in yeast. *Genetics* 192, 73–105.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J., et al. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* 7, R100.

Chong, Y.T., Cox, M.J., and Andrews, B. (2012). Proteome-wide screens in *Saccharomyces cerevisiae* using the yeast GFP collection. *Adv. Exp. Med. Biol.* 736, 169–178.

Cline, M.S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., Christmas, R., Avila-Campilo, I., Creech, M., Gross, B., et al. (2007). Integration of biological networks and gene expression data using Cytoscape. *Nat. Protoc.* 2, 2366–2382.

Costanzo, M., Nishikawa, J.L., Tang, X., Millman, J.S., Schub, O., Breitkreuz, K., Dewar, D., Rupes, I., Andrews, B., and Tyers, M. (2004). CDK activity antagonizes Whi5, an inhibitor of G1/S transcription in yeast. *Cell* 117, 899–913.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L., Toufighi, K., Mostafavi, S., et al. (2010). The genetic landscape of a cell. *Science* 327, 425–431.

Dechant, R., and Peter, M. (2008). Nutrient signals driving cell growth. *Curr. Opin. Cell Biol.* 20, 678–687.

Dénervaud, N., Becker, J., Delgado-Gonzalo, R., Damay, P., Rajkumar, A.S., Unser, M., Shore, D., Naef, F., and Maerkl, S.J. (2013). A chemostat array enables the spatio-temporal analysis of the yeast proteome. *Proc. Natl. Acad. Sci. USA* 110, 15842–15847.

Dubacq, C., Chevalier, A., Courbeyrette, R., Petat, C., Gidrol, X., and Mann, C. (2006). Role of the iron mobilization and oxidative stress regulons in the genomic response of yeast to hydroxyurea. *Mol. Genet. Genomics* 275, 114–124.

Fazio, T.G., Kooperberg, C., Goldmark, J.P., Neal, C., Basom, R., Delrow, J., and Tsukiyama, T. (2001). Widespread collaboration of Isw2 and Sin3-Rpd3 chromatin remodeling complexes in transcriptional repression. *Mol. Cell. Biol.* 21, 6450–6460.

Ferrigno, P., Posas, F., Koepp, D., Saito, H., and Silver, P.A. (1998). Regulated nucleo/cytoplasmic exchange of HOG1 MAPK requires the importin beta homologs NMD5 and XPO1. *EMBO J.* 17, 5606–5614.

Fournier, M.L., Paulson, A., Pavelka, N., Mosley, A.L., Gaudenz, K., Bradford, W.D., Glynn, E., Li, H., Sardi, M.E., Fleharty, B., et al. (2010). Delayed correlation of mRNA and protein expression in rapamycin-treated cells and a role for Ggc1 in cellular sensitivity to rapamycin. *Mol. Cell. Proteomics* 9, 271–284.

Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. (2000). Genomic expression programs in the

- response of yeast cells to environmental changes. *Mol. Biol. Cell* **11**, 4241–4257.
- Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dümpelfeld, B., et al. (2006). Proteome survey reveals modularity of the yeast cell machinery. *Nature* **440**, 631–636.
- Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephore, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. *Nature* **425**, 737–741.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I.H. (2009). The WEKA data mining software: an update. *SIGKDD Explor.* **11**, 10–18.
- Handfield, L.F., Chong, Y.T., Simmons, J., Andrews, B.J., and Moses, A.M. (2013). Unsupervised clustering of subcellular protein expression patterns in high-throughput microscopy images reveals protein complexes and functional relationships between proteins. *PLoS Comput. Biol.* **9**, e1003085.
- Henriksen, P., Wagner, S.A., Weinert, B.T., Sharma, S., Bacinskaja, G., Rehman, M., Juffer, A.H., Walther, T.C., Lisby, M., and Choudhary, C. (2012). Proteome-wide analysis of lysine acetylation suggests its broad regulatory scope in *Saccharomyces cerevisiae*. *Mol. Cell. Proteomics* **11**, 1510–1522.
- Hood-DeGrenier, J.K. (2011). Identification of phosphatase 2A-like Sit4-mediated signalling and ubiquitin-dependent protein sorting as modulators of caffeine sensitivity in *S. cerevisiae*. *Yeast* **28**, 189–204.
- Howe, A.G., Zarembek, V., and McMaster, C.R. (2002). Cessation of growth to prevent cell death due to inhibition of phosphatidylcholine synthesis is impaired at 37 degrees C in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **277**, 44100–44107.
- Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K. (2003). Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691.
- Jones, T.R., Kang, I.H., Wheeler, D.B., Lindquist, R.A., Papallo, A., Sabatini, D.M., Golland, P., and Carpenter, A.E. (2008). CellProfiler Analyst: data exploration and analysis software for complex image-based screens. *BMC Bioinformatics* **9**, 482.
- Kaluarachchi Duffy, S., Friesen, H., Baryshnikova, A., Lambert, J.P., Chong, Y.T., Figeys, D., and Andrews, B. (2012). Exploring the yeast acetylome using functional genomics. *Cell* **149**, 936–948.
- Kemmeren, P., Sameith, K., van de Pasch, L.A., Benschop, J.J., Lenstra, T.L., Margaritis, T., O'Duibhir, E., Apweiler, E., van Wageningen, S., Ko, C.W., et al. (2014). Large-scale genetic perturbations reveal regulatory networks and an abundance of gene-specific repressors. *Cell* **157**, 740–752.
- Koh, J.L., Ding, H., Costanzo, M., Baryshnikova, A., Toufighi, K., Bader, G.D., Myers, C.L., Andrews, B.J., and Boone, C. (2010). DRYGIN: a database of quantitative genetic interaction networks in yeast. *Nucleic Acids Res.* **38**, D502–D507.
- Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* **11**, 319–324.
- Lee, S., Lim, W.A., and Thorn, K.S. (2013). Improved blue, green, and red fluorescent protein tagging vectors for *S. cerevisiae*. *PLoS ONE* **8**, e67902.
- Loewith, R., and Hall, M.N. (2011). Target of rapamycin (TOR) in nutrient signaling and growth control. *Genetics* **189**, 1177–1201.
- Loo, L.H., Laksameethanasan, D., and Tung, Y.L. (2014). Quantitative protein localization signatures reveal an association between spatial and functional divergences of proteins. *PLoS Comput. Biol.* **10**, e1003504.
- Lykke-Andersen, S., Brodersen, D.E., and Jensen, T.H. (2009). Origins and activities of the eukaryotic exosome. *J. Cell Sci.* **122**, 1487–1494.
- Mazumder, A., Tummeler, K., Bathe, M., and Samson, L.D. (2013). Single-cell analysis of ribonucleotide reductase transcriptional and translational response to DNA damage. *Mol. Cell. Biol.* **33**, 635–642.
- Nagaraj, N., Kulak, N.A., Cox, J., Neuhauser, N., Mayr, K., Hoerning, O., Vorm, O., and Mann, M. (2012). System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap. *Mol. Cell. Proteomics* **11**, M111.013722.
- Newman, J.R., Ghaemmaghami, S., Ihmels, J., Breslow, D.K., Noble, M., DeRisi, J.L., and Weissman, J.S. (2006). Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441**, 840–846.
- Parsons, A.B., Brost, R.L., Ding, H., Li, Z., Zhang, C., Sheikh, B., Brown, G.W., Kane, P.M., Hughes, T.R., and Boone, C. (2004). Integration of chemical-genetic and genetic interaction data links bioactive compounds to cellular target pathways. *Nat. Biotechnol.* **22**, 62–69.
- Platt, J. (1998). Fast training of support vector machines using sequential minimal optimization. In *Advances in Kernel Methods—Support Vector Learning* (MIT Press), pp. 185–208.
- Rees, J.S., Lowe, N., Armean, I.M., Roote, J., Johnson, G., Drummond, E., Spriggs, H., Ryder, E., Russell, S., St Johnston, D., et al. (2011). In vivo analysis of proteomes and interactomes using Parallel Affinity Capture (iPAC) coupled to mass spectrometry. *Mol. Cell. Proteomics* **10**, M110.002386.
- Robert, T., Vanoli, F., Chiolo, I., Shubassi, G., Bernstein, K.A., Rothstein, R., Botrugno, O.A., Parazzoli, D., Oldani, A., Minucci, S., and Foiani, M. (2011). HDACs link the DNA damage response, processing of double-strand breaks and autophagy. *Nature* **471**, 74–79.
- Sheff, M.A., and Thorn, K.S. (2004). Optimized cassettes for fluorescent protein tagging in *Saccharomyces cerevisiae*. *Yeast* **21**, 661–670.
- Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**, 635–646.
- Teixeira, D., and Parker, R. (2007). Analysis of P-body assembly in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **18**, 2274–2287.
- Tkach, J.M., Yimit, A., Lee, A.Y., Riffle, M., Costanzo, M., Jaschob, D., Hendry, J.A., Ou, J., Moffat, J., Boone, C., et al. (2012). Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. *Nat. Cell Biol.* **14**, 966–976.
- Tong, A.H., Evangelista, M., Parsons, A.B., Xu, H., Bader, G.D., Pagé, N., Robinson, M., Raghibzadeh, S., Hogue, C.W., Bussey, H., et al. (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* **294**, 2364–2368.
- Visintin, R., Craig, K., Hwang, E.S., Prinz, S., Tyers, M., and Amon, A. (1998). The phosphatase Cdc14 triggers mitotic exit by reversal of Cdk-dependent phosphorylation. *Mol. Cell* **2**, 709–718.
- Xie, M.W., Jin, F., Hwang, H., Hwang, S., Anand, V., Duncan, M.C., and Huang, J. (2005). Insights into TOR function and rapamycin response: chemical genomic profiling by using a high-density cell array method. *Proc. Natl. Acad. Sci. USA* **102**, 7215–7220.
- Yan, H., Merchant, A.M., and Tye, B.K. (1993). Cell cycle-regulated nuclear localization of MCM2 and MCM3, which are required for the initiation of DNA synthesis at chromosomal replication origins in yeast. *Genes Dev.* **7**, 2149–2160.
- Yang, X.J., and Seto, E. (2008). The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. *Nat. Rev. Mol. Cell Biol.* **9**, 206–218.