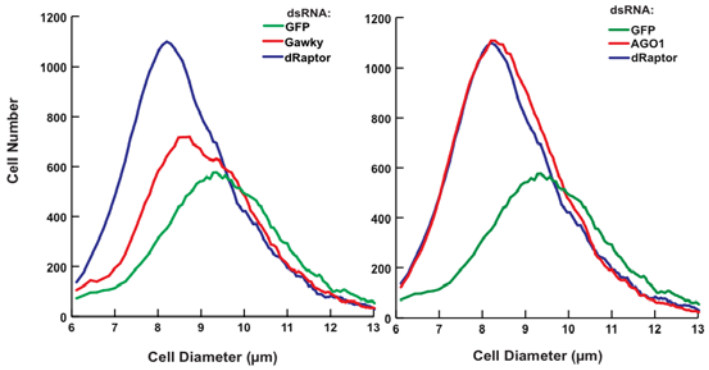
Supplemental figures and legends, Supplemental table legends, and Methods
Lindquist, et al. manuscript

Supplemental Figures

Supplemental Figure S1. Confirmed hit status in secondary screens, relative to rank order in primary screen. The pre-printed format of cell microarrays does not allow inclusion of dozens of assay-specific positive controls, precluding calculation of traditional assay quality statistics, so here we tested whether the top of each rank-ordered list of candidate genes (based on the primary screen) is enriched in confirmed hits (that is, confirmed in the secondary screens). All candidate genes tested in the secondary screens were rank-ordered according to their low (left) or high (right) Hit Score in the primary screen. A gene that had a confirmed phenotype in the secondary screens (as described in the text and Supplemental Methods) is denoted with a line whereas unconfirmed candidate genes are left blank. The confirmed genes are biased towards higher-ranking in the primary screen: 42% of confirmed low-pS6 genes and 56% of high-pS6 genes were in the top third of their respective candidate gene lists.



Supplemental Figure S2. Raw cell count histograms. These histograms reflect the raw cell counts that are presented as normalized in Fig. 6B. *Drosophila* S2 cells were transfected with dsRNAs targeting the indicated genes, and cell size was measured using a Coulter counter.

Supplemental Figures S3 and S4 are contained within the Supplemental Methods that follow.

Supplemental Table Legends

Supplemental Table S1: Genome screen results and analysis. The spreadsheet contains the raw results of the genome screen, a list of candidate high- and low-S6 hits from the genome screen, and an analysis of families of genes enriched in the genome screen, and primer sequences.

Supplemental Table S2: Results and analysis from secondary screen and subsequent follow-up experiments. The spreadsheet contains the raw results of the secondary screen, a list of the S6/S6K regulators confirmed via the secondary screen and subsequent followup experiments, the primer sequences used for RNAi in followup experiments, and the groups of genes from published studies that were used in our meta-analysis.

Other Supplemental Materials

The open-source high-throughput image analysis software CellProfiler was used to analyze images from the screens (http://www.cellprofiler.org). The following CellProfiler image analysis pipelines will be available on the journal's website:

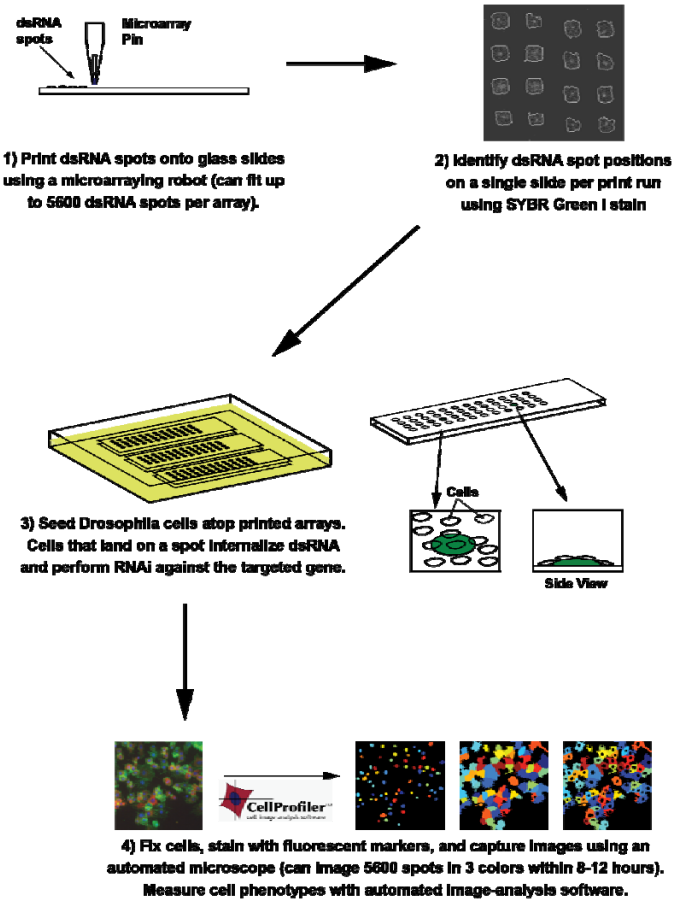
Pipeline 1. Rotates a SYBR test slide to align control spots.

Pipeline 2. Identifies SYBR-stained spots on a test slide.

Pipeline 3. Performs illumination correction for image analysis.

Pipeline 4. Analyzes images of cells.

Supplemental Methods



Supplemental Figure S3. Schematic of cell array printing and analysis.

Microarray printing and spot identification

An overview of the procedure is shown in Supplemental Figure S3. dsRNAs were drawn from 384-well source plates at a concentration of 0.3 ug/ul in a solution of 0.5M NaCl / 10uM Tris pH 7.0 and printed onto poly-lysine slides (Erie Scientific, Erie, NH) using a PixSys 5500 contact arrayer and SMP7 stealth microarray pins (Genomic Solutions). Preprints were performed on UltraGAPS slides (Corning). Spots of dsRNA had a diameter of ~220 µm and were spaced 450 µm apart from one another. The main grid of dsRNAs, which contained periodic cell-death spots

of *DIAP1* dsRNA, was flanked by *DIAP1* dsRNA marker spots at the corners, as well as spots of Rhodamine dissolved in poly-D-lactic acid. Printed arrays were placed in 5-slide boxes (Corning), sealed in Kapak pouches, and stored at -20°C until use. Arrays were printed in 40 replicates; one slide from each print run was stained with SYBR Green I (Molecular Probes), which binds dsRNA and fluoresces, in order to identify locations for later high-resolution microscopy. Slides were stained for 12 minutes in Coplin jars filled with 1:5,000 SYBR Green I diluted in 2x SSC (Ambion) / 0.1% Triton (Roche), then washed for 7 minutes in 2xSSC/0.1% Triton, washed for 2 minutes in 2xSSC, and spun dry for 15 seconds using a tabletop slide-centrifuge. SYBR-stained slides were imaged using an Axon slide-scanner at 10µm resolution with a 532nm laser. dsRNA spot positions were obtained from the resulting image using CellProfiler software (Carpenter et al. 2006; Lamprecht et al. 2007) (http://www.cellprofiler.org) and two image analysis pipelines (Supplemental Pipelines 1 and 2).

Creation of position lists in CellProfiler

First, the "Rotate" image analysis pipeline (Supplemental Pipeline 1) was employed by CellProfiler to set two array-flanking marker spots, either *DIAP1* spots or rhodamine spots, onto the same Y coordinate. This allowed for the left-corner marker spot to be defined as the origin and for the x-axis to be defined as the line through the left and right corner spots. Because a full-size array at 10µm resolution is about a 30-MB image file, the automated rotation is a computationally intensive process that taxed the MatLab engine underlying this version of CellProfiler. So, we used Adobe Photoshop to identify the angle of rotation required for each array through trial-and-error and fed this value to the CellProfiler Rotate pipeline; in our experience, an image would generally need rotation about a tenth of a degree counter-clockwise.

This process produced a rotated image in which, using the "Info" tab of Adobe Photoshop, we determined the distance X_w from the left-corner marker spot (the origin) to the right-corner marker spot, as well as the X- and Y-offsets (X_o, Y_o) from the left-corner marker spot to the top left spot on the main grid of the array. We subsequently used the "Identify" pipeline (Supplemental Pipeline 2) on the rotated image, specifying grid lines as 45 pixels apart and expected object size between 10 and 40 pixels in diameter. The "Identify" pipeline broke the SYBR-stained image into a grid and then attempted to identify one "spot" per grid-tile, based on a user-defined threshold of SYBR-stain intensity. Spots of irregular shape or size were filtered out; the positions of the centroids of successfully-identified spots were recorded in an output file. The accuracy of the identification pipeline was verified by examining the outlined spots in the resulting image file.

The final position list was produced as a comma-separated value file, by choosing ExportLocations from the "Tools" menu, specifying the CellProfiler Output file that was created in the Identify pipeline, specifying "Spots", then specifying the coordinates (in μm) of X_o,Y_o, a scale of 10 μm per pixel, either "Comb" or "Meander" acquisition pattern, and "Yes" to correct for uneven grids. (Note that it is possible to acquire images of the microscope slide in either or two orders. In the "Comb" order, the stage moves like a typewriter, capturing each row of spots in left-to-right order, while in the "Meander" order, rows 1,3,5,... are captured left-to-right and rows 2,4,6... are captured right-to-left. We used "Meander", in order to minimize large stage movements and focal-plane shifts during image acquisition.)

Because our microarray pins were spaced 4.5mm apart in the printhead while our spots were at 450 μm intervals, each pin printed a 10x10 block of dsRNA spots. Fewer than 1% of spots were lost due to printing failure. In our experience the spots within a 10x10 block are

spaced perfectly regularly, while significant deviations in the array grid occur at the interface between these blocks. Such an interface is visible in the scanned slide in Supplemental Figure S3. These irregularities are consistent from slide to slide, and it is possible to superimpose the scans of two SYBR-stained slides from a single print run such that their spots overlap completely; this is why we could sacrifice one slide for SYBR-staining to generate a position list that applied to the remaining 39 slides in a print run. Because the 10x10 blocks were internally regular, and because not every spot on a SYBR-stained slide could be identified properly, the "Correct for Uneven Grids" option allows positions where spots were not correctly identified to be forced into alignment with the successfully-aligned spots. The key step in generating an accurate position list, then, lies in trying different threshold intensities until each 10x10 block contains at least 10% of its spots identified properly.

Microarray seeding and cell culture. Microarray slides were equilibrated to room temperature and placed into 100-mm square dishes (Nalge Nunc). Each dish accommodates three slides; we used two cell microarrays and one blank slide (which served as a no-primary-antibody control). Cells were grown from an initial density of 20 million per T75 flask for 4 days prior to seeding, then seeded into array dishes as a thoroughly-mixed single-cell suspension of 10 million cells in 25 ml fresh medium. Cells were grown on microarrays for 72 hours in Schneider's Medium / 10%IFS, then moved to dishes of medium containing 1 mM $CuSO_4$ for induction of the human S6 reporter. For the primary screen and the "24 hour refed" arrays of the secondary screen, we moved cells to fresh Schneider's/10%IFS with $CuSO_4$ after 72 hours. For other growth conditions in the secondary screen, induction occurred by moving arrays transiently to fresh dishes of Schneider's/10%IFS, collecting the nutrient-depleted medium in which the arrays had

previously grown, filtering that medium, adding $CuSO_4$ to 1 mM, distributing the induction medium into new square dishes into which we moved the arrays for 24 hours' induction. Among those other growth conditions, four "+ rapa" arrays received an additional 1-hr dose of 3 ml IFS plus 5 µl 100 µM rapamycin in EtOH per dish (final concentration 20 nM rapamycin), four "serum-spike" arrays received a 1-hr dose of 3 ml IFS plus 5 µl EtOH vehicle per dish, and four "starved" arrays were not given any additional treatment. For the "GFP synthetic" and "TSC2 synthetic" arrays in the secondary screen, cells were initially seeded in 10-cm dishes at 20 million cells/dish and transfected with 50 µg dsRNA targeting GFP or TSC2, incubated for two days, then rinsed in Schneider's medium, seeded onto microarrays, and induced as in the primary screen. For the Kc167 arrays in the secondary screen, cells were grown for 3 days in T75 flasks from an initial density of 70M cells/flask, then seeded onto microarrays at a density of 15 million cells per dish, and grown on arrays for 72 hours prior to fixation.

Fixation and fluorescent staining of cell microarrays. Cells on arrays were rinsed once in PBS+ (Phosphate-Buffered Saline containing an additional 1 mM $MgCl_2$ and 1 mM $CaCl_2$), then fixed for 25 minutes in a fresh solution of 3.7% paraformaldehyde (from 16% stock, Electron Microscopy Sciences) / 4% sucrose (Sigma) in PBS^+. Fixed arrays were rinsed in PBS^+ and permeabilized for 30 minutes with 0.1% Triton X-100 (Roche) in PBS^+, then rinsed and placed in a humidified chamber and topped with 600 µl blocking solution, 1% BSA (Sigma) + 1% normal donkey serum (Jackson Immunolabs) in PBS^+. Arrays were incubated overnight at 4 degrees in 600 µl antibody-incubating solution (0.5% BSA in PBS^+) with primary antibody diluted 1:100. Arrays were rinsed in PBS^+, incubated in blocking solution for 20 minutes, and fluorescently stained for 1 hour with 1:500 Alexa 488-phalloidin (Invitrogen) and 1:500 Cy3-

anti-rabbit (Jackson Immunolabs) in antibody-incubating solution. Arrays were then rinsed in PBS^+ and stained for 1hr with 1:10,000 Hoechst 33342 dye in PBS+ (Invitrogen) prior to a final PBS^+ rinse, a slow rinse in 3:2 PBS^+:glycerol solution, mounting with Vectashield (Vector Labs), coverslipping, and sealing with clear nail polish. Stained slides were stored at 4 degrees Celsius in the dark until image acquisition.

Spot-finding, array alignment, and image acquisition

Images were acquired on an Axiovert 200M microscope using KS400 high-throughput image-acquisition software (Carl Zeiss). Briefly, our KS400 macro loaded information from two user-specified files for the exposure settings and the list of array spot-positions; asked the user to specify the size of the images and the distance Xw between two marker spots; asked the user to bring the cell nuclei into focus for calibrating the autofocus; and, finally, asked the user to orient the X-Y coordinate axes by bringing the stage to the left and right corner spots. The program then automatically captured images in up to three channels at each coordinate pair in the position-list file. In order to make sure that cell arrays were properly aligned, we first ran the image-acquisition routine on a short position-list containing cell-death (ds*DIAP1*) spots, and based on the images from this test run, we reran the imaging routine, making slight adjustments to where we specified the left and/or right corner spots. A slide was considered optimally aligned when the test images were centered around, or entirely contained within, the internal *DIAP1* spots. Once an optimal pair of corner spots was found, we ran the image-acquisition routine using the same corner spots and the full 5600-spot position list. The dsRNA spot diameter is ~220 μm, which accommodates an inscribed square of side length 155 μm. To ensure that the field of view fell completely within the printed spot, we set the field of view to a square with side

length 150 μm.

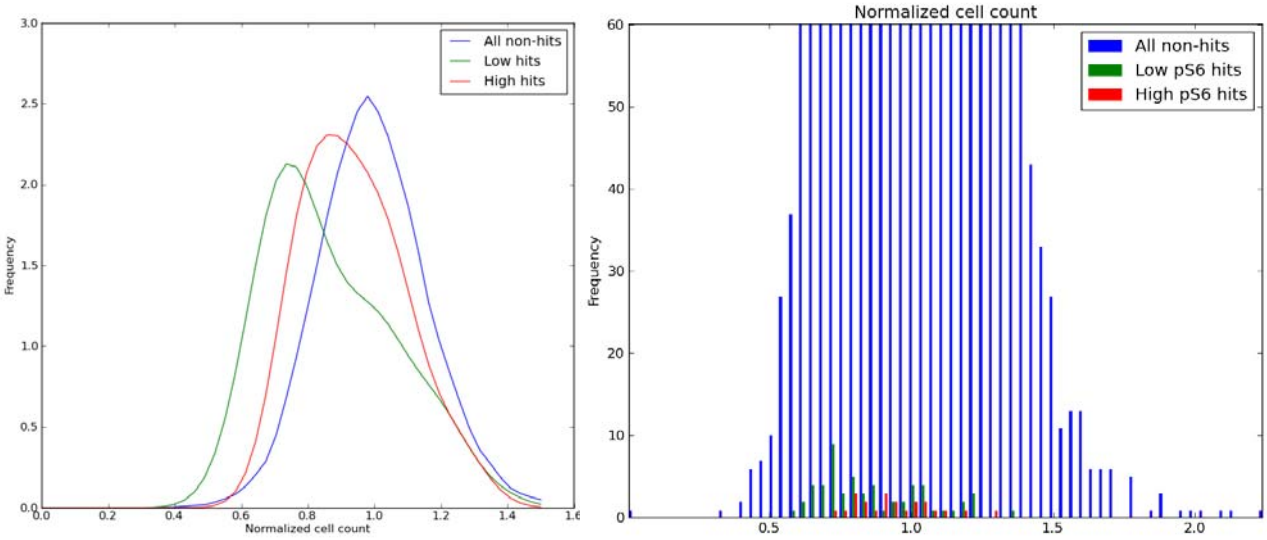
Quantifying pS6 immunofluorescence and statistical analysis of screen data

We processed the acquired image of each spot with two CellProfiler image-analysis pipelines. The first (Supplemental Pipeline 3) was an illumination-correction pipeline that corrects the variation in brightness over the imaging field and improves the accuracy of intensity measurements. Then we ran an image-analysis pipeline (Supplemental Pipeline 4) that identifies nuclei (based on the Hoechst stain) and cell borders (based on the phalloidin stain) and quantifies the pS6 phenotype. The pS6 phenotype was measured as the mean cytoplasmic fluorescence intensity of pS6 across all cells in the imaged population. This was calculated by computing a ratio where the numerator is total cytoplasmic pS6 signal, averaged across all cells in the population (Mean_Cytoplasm_Intensity_CorrRed_IntegratedIntensity), and the denominator is cytoplasm area, averaged across all cells in the population (Mean_Cytoplasm_Areashape_Area). This yields a metric that is not directly affected by changes in the number of cells in each sample. We tested alternate metrics, such as omitting the normalization to cytoplasm area, normalizing to cell area instead, and measuring pS6 signal across the entire image, and found these to be less robust at discriminating between controls. To confirm empirically that changes in cell number were not directly affecting measured levels of p-hS6, we later plotted histograms of cell count for the low-pS6 hits, high-pS6 hits, and non-hits from the primary screen (how the hits were determined is described later). Both low- and high-hits show slightly decreased cell count. This is typical in RNAi screens and in our case also indicates that there is not an artifactual relationship between decreased cell count and pS6 measurement (Supplemental Figure S4, left). Further, there are many samples that have higher and lower cell counts than the hits themselves,

which yet do not score as having a pS6 phenotype (Supplemental Figure S4, right).

Supplemental Figure S4. Cell counts per spot for low-pS6 hits, high-pS6 hits, and non-hits from the primary screen shown with a relative y-axis (left) and an absolute y-axis (right). "Normalized" cell count in this context refers to the spatial normalization across each slide.

In order to account for nonspecific background fluorescence from secondary antibody,



we imaged a no-primary control slide, computed its local median background fluorescence within the 5x5 window surrounding each spot location, and subtracted that background intensity from the raw fluorescence intensity of each screening slide. We subsequently adjusted for regional variation in antibody density by dividing each spot's fluorescence by the local median intensity within the 5x5 window surrounding that spot.

Given that our assay was designed to capture both high- and low-pS6 perturbations, rather than optimized for either, we expected that reproducibility would not be as high as previously seen for cell microarrays (Wheeler et al. 2004; Neumann et al. 2010). This was

indeed the case, with correlation coefficients ranging from 0.10-0.28 between replicates in the primary and secondary screens. We therefore performed the primary screen in quadruplicate and used a stringent scoring methodology wherein the score calculated for each gene excluded the best replicate for that gene, as follows.

For each genome array replicate, we placed each spot's normalized pS6 signal onto a log scale and computed the mean and standard deviation over the slide's 5600 spots. With these values, we calculated a z-score for each of the 5600 spots on the array. The z-score is the number of standard deviations an observed data point is from the mean[1]. We then produced separate rankings for low-pS6 and high-pS6 phenotypes (shown in Supplemental Table S1A) based on "hit scores" that were computed as follows. First, we computed p-values for each spot as the area underneath a normal distribution (with mean = 0 and standard deviation = 1) above the spot's z-score. We computed high-pS6 hit scores by taking the geometric mean of the p-values of the three *lowest* scoring spots, throwing out the remaining *highest* scoring spot, since there were four replicates of the genome array. We computed low-pS6 hit scores by taking the geometric mean of the p-values of the three *highest* scoring spots, throwing out the remaining *lowest* scoring spot. We felt this stringent approach, throwing out the strongest of the four replicate spots, was appropriate to reduce false positives. We found that the geometric mean showed a better enrichment for known TORC1 pathway components than averaging p-values or z-scores, possibly because it reduced bias against samples for which one replicate failed.

We took a practical approach to choosing hits for the secondary screens based on the

[1] Robust z-scores were not typical at the time of our original analysis. Retrospectively, we determined that the log-transformed primary screen data is normally distributed without dramatic outliers, such that the mean is nearly identical to the median and the SD and MAD exhibit the expected relationship. As well, robust and regular z-scores produce identical rankings and we did not use a pre-defined z-score cutoff in our screen, so the choice of methodology was not critical.

primary screen data: (1) Because about half of the 13,633 genes in our library are targeted by two or more dsRNA sequences, one set of genes was chosen for the secondary screens based on having more than one dsRNA with at least a moderate phenotype, defined as scoring within the top 1,500 spots for the high-pS6 and low-pS6 lists (1,500/22,400 total spots = 6.7%). This cutoff was chosen empirically, influenced by the number of genes that could fit on the secondary screening cell microarrays and by the presence of known pathway regulators in the lists. (2) A second set of genes was added to secondary screening based on having a single dsRNA with a strong phenotype, defined using empirically chosen cutoffs of 1% for low-pS6 and 0.6% for high-pS6. We selected the cutoffs differently for low- and high-pS6 because (a) there were more and larger gene families on the low-pS6 list, so we chose more low-pS6 hits than high-pS6 hits in order to capture these, and (b) there were more low-pS6 hits that scored far from the distribution of negative-control spots, exemplified by the fact that the magnitude of hit scores diminished sooner on the high-pS6 list than on the low-pS6 list. We surmise that our assay was better suited to identifying low-pS6 regulators, since the cells were grown in the presence of serum. Thus, the basal levels of pS6 were relatively elevated, somewhat limiting the ability to detect further increases. (3) A third set of genes was added where only a single dsRNA scored in the top 1500, but whose known biological functions suggested interesting hypotheses to test. Of those genes we tested in this set, the only one to pass secondary screening with multiple independent dsRNAs was *Akap200*. Interestingly, this gene had only one dsRNA in the genome library and might have passed the initial primary screening threshold if given equal opportunity.

In the secondary screens, each dsRNA was represented by 4 spots on the array (8 for some controls), and each condition included 3 or 4 arrays for a total of n=12 or 16 spots per dsRNA per condition. For analysis of the secondary screens' pS6 phenotypes, we normalized

each spot's pS6 signal by dividing by the local median intensity of only those GFP control spots within the 5x5 window surrounding that spot, to account for local staining variation. To determine whether a dsRNA had a significant effect on pS6 under a given growth condition, we calculated p-values and chose a threshold of 0.05 (relative to the GFP controls in that condition). The p-values for each dsRNA under each condition are computed by t-test comparison to the set of GFP spots under that condition (n=1200 GFP spots per array, e.g. n=4800 GFP spots for the "serum-spike" condition). We further increased the stringency and ruled out off-target effects by only counting genes as significant if all available dsRNAs targeting that gene scored as significant. For analysis of the secondary screens' cell count phenotypes, cell counts were subjected to the same local-median normalization as pS6 scores, in order to account for local variation in initial cell density.

Tips for microarray seeding and cell culture

In order to obtain an even distribution of cells across the microarrays in a dish, it was essential to rock array dishes gently and immediately after seeding in a tissue culture hood; then to move dishes carefully to the incubator; to rock dishes gently again immediately after placing them into the incubator; and to keep the incubator off for the first hour cells spend in the incubator, so that the incubator did not vibrate while cells settle onto the array. Also, it was essential to preserve high environmental humidity (≥70%) in the incubator where fly cells grow: if medium evaporated and gained in osmolarity, then cells might begin to detach.

Supplemental references

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**(10): R100.

Lamprecht, M.R., Sabatini, D.M., and Carpenter, A.E. 2007. CellProfiler: free, versatile software for automated biological image analysis. *Biotechniques* **42**(1): 71-75.

Neumann, B., Walter, T., Heriche, J.K., Bulkescher, J., Erfle, H., Conrad, C., Rogers, P., Poser, I., Held, M., Liebel, U. et al. 2010. Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature* **464**(7289): 721-727.

Wheeler, D.B., Bailey, S.N., Guertin, D.A., Carpenter, A.E., Higgins, C.O., and Sabatini, D.M. 2004. RNAi living-cell microarrays for loss-of-function screens in Drosophila melanogaster cells. *Nat Methods* **1**(2): 127-132.