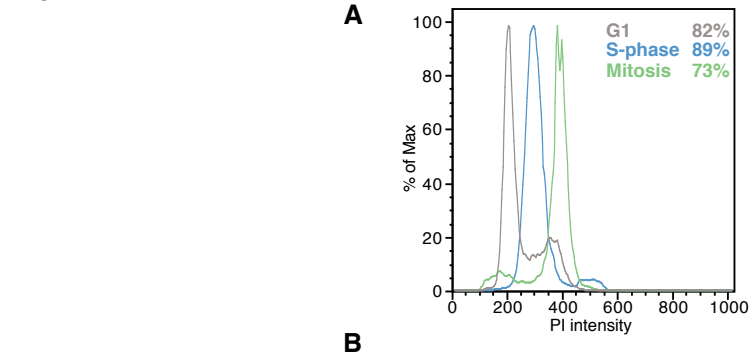
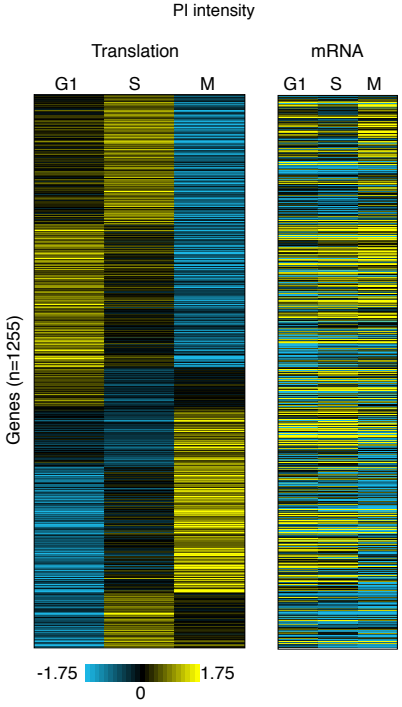
Supplemental Materials

Figure S1





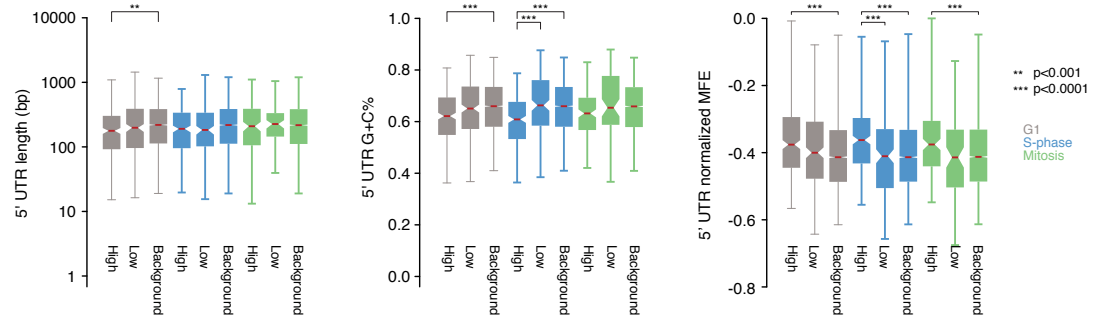


Figure S1. Related to Figure 1. **(A)** Flow cytometry histogram plotting the distribution of cells as a function of DNA content measured by propidium iodide (PI) fluorescence intensity. **(B)** Heat maps representing the translational regulation (left) and mRNA expression levels (right) for the 1255 translationally regulated genes within any phase of the cell cycle. Genes and cell cycle phases are clustered based on the level of normalized ribosome occupancy (mean-centered translational efficiency by gene, left). mRNA expression levels (right) were ordered based on the clustering of translational regulation. Scales and colors indicate the direction and magnitude of mRNA translation or levels, respectively. **(C)** Box and whisker plots denoting the 5'UTR length (left), %G+C content (middle), and minimum free energy (MFE) (right) for genes with higher, lower, or expected levels of ribosome occupancy. Y-axes are as labeled. Cell cycle phases are colored: Grey, G1; Blue, S-phase; Green, mitosis.

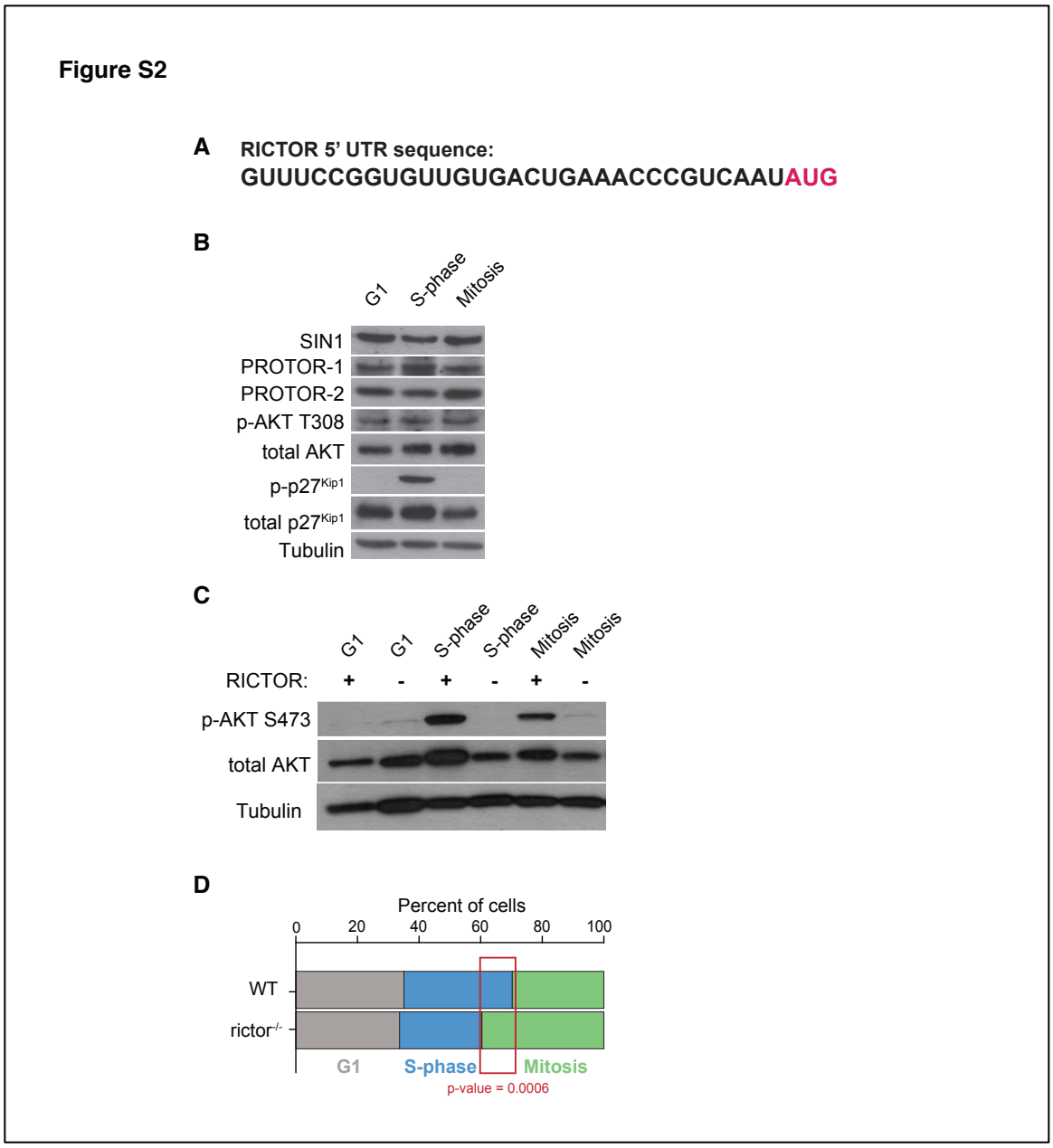


Figure S2. Related to Figure 2. (A) The sequence of the *RICTOR* 5'UTR. The start codon for the coding region is highlighted in red. **(B)** Western blots showing the levels of mTORC2 components and the phosphorylation status of AKT and downstream signaling targets. SIN1, PROTOR-1, and PROTOR-2 are components of mTORC2. Phosphorylation of AKT T308 is independent of mTORC2. $p27^{Kip1}$ is a phosphorylation target of AKT. Tubulin is a loading control. **(C)** Western blots

illustrating the phosphorylation status of AKT in synchronized wildtype and *rictor*$^{-/-}$ MEFs. **(D)** Cell cycle analysis of asynchronous MEFs with the percentage of cells in each phase of the cell cycle indicated along the Y-axis. The red box highlights the magnitude of the difference between the number of WT and *rictor*$^{-/-}$ MEFs in S-phase. Probability calculated by Student's t-test. Cell cycle phases are colored as labeled.

Figure S3. Related to Figure 4. (A) The sequence of the *NIPBL* 5'UTR. Potential AUG start codons are bold. The uORF corresponding to the peak of bound ribosome in Figure 4C is orange. The region of the CC > A mutation in Cornelia de Lange syndrome is underlined. **(B)** The sequence of the *WAPAL* 3'UTR immediately downstream of the stop codon. The region corresponding to the peak of ribosome binding is underlined. In-frame stop codons are shown in red. The predicted amino acid sequence of the putative C-terminal extension is shown below. The sequence up to the first in-frame stop codon is red.

Supplemental Tables

Table S1. Translationally regulated genes from each phase of the cell cycle, related to Figure 1.

Table S2. Genes with significant changes in translation between G1 and S-phase, G1 and mitosis, and S-phase and mitosis, related to Figure 1.

Table S3. Functional enrichments of genes that are translationally regulated during G1, S-phase, or mitosis, related to Figure 3.

Supplemental Experimental Procedures

Alignment and analysis of ribosome profiling data

Raw single-end sequencing reads were processed in the following manner. First, reads were clipped of the known adaptor sequence (CTGTAGGCACCATCAAT) and remaining 3' sequence with the FASTX-Toolkit, retaining only reads of 24bp post-clipping length or longer. Clipped reads were mapped to the human genome assembly (NCBI build 36) with Burrows-Wheeler Alignment (BWA; ver. 0.5.9-r16, default parameters)(Li and Durbin, 2009). Unaligned reads were then mapped to known splice junctions with Tophat (ver. 1.4.1) (Trapnell et al., 2009) using a transcriptome index created from version 11 of the Gencode gene set (Harrow et al., 2006). Provisional merged BAM files (Li et al., 2009) were created from aligned and unaligned reads from either the genome or spliced alignment and read groups were enforced across lanes and experiments. BAM files were then indexed, coordinate sorted, and had alignment metrics determined, all with the Picard suite (http://picard.sourceforge.net/). Only those protein-coding transcripts of levels 1

or 2 of Gencode support (verified or manually curated loci respectively) were used for all subsequent analyses. A unified gene model was created for each gene in which the union of coding sequence across all isoforms of a given gene was collapsed into a single model per gene. Gene-level translational regulation was inferred for only those genes estimated to be expressed in the Hela transcriptome, determined by comparing their observed expression level (calculated as a normalized rpkM expression estimate) to a null distribution of similar expression in randomly selected non-coding regions of the genome with a size distribution sampled from the empirical size distribution (Olshen et al. *submitted*). In total, 10,814 genes were retained for further analysis.

We quantify mRNA translation for genes (ribosome-given-mRNA levels) with a statistical framework called *Babel*. Briefly, this errors-in-variables regression model, the details of which are presented elsewhere (Olshen et al. *submitted*), assesses the significance of translational regulation for every gene within and between samples and conditions. For a transcript with a given abundance (mRNA counts), the expected level of ribosome occupancy (RPF counts) is inferred from trimmed least-squares regression (based on the monotonic relationship typified in figure 1). We model both mRNA levels and the level of bound ribosome given mRNA abundance by the negative binomial distribution. The over-dispersion of the latter is modeled using an iterative algorithm to prevent over-fitting. We then use a parametric bootstrap based on the modeling described above to estimate a p-value for every gene. It tests formally the null hypothesis that the level of bound ribosome is as expected from mRNA abundance. These p-values are estimated for a one-sided test

in which both low and high p-values are of interest; low p-values correspond to higher than expected ribosome-given-mRNA counts and high p-values correspond to the opposite.

Since Babel quantifies this mRNA translation (ribosome-given-mRNA levels) as a p-value in each sample of a given condition, we combine these p-values into a single assessment of the significance of translational regulation in each condition (phase of the cell cycle). Here, we developed an alternative methodology to Fisher's method first proposed by Edgington (Edgington, 1972). This approach, in which the combined p-value is a function of the arithmetic mean of individual p-values, maintains symmetry and convexity between each of two or more tests of hypotheses and their p-values. It produces a small combined p-value from consistently small p-values of ribosome-given-mRNA levels (and the reverse for two or more large p-values). Once p-values are combined, two-sided p-values can be estimated and genes are considered regulated at the translational level if their corresponding p-values are low after adjusting for multiple comparisons [in this study, those corresponding to a false discovery rate (FDR) less than 1% (Storey, 2002)].

Differences in ribosome occupancy given mRNA levels between conditions are assessed utilizing a statistic based upon using the Gaussian quantile function to transform the within condition p-values into z-statistics. Corresponding p-values are estimated by comparison to the Gaussian distribution and again corrected for multiple comparisons as described above. Significant differentially translationally regulated genes are those corresponding to an FDR < 5% in any of the pair-wise

comparisons of cell cycle phases. For visualization purposes, the conventional measure of translational efficiency was used and calculated as previously described (Ingolia et al., 2009).

Supplemental References

Edgington, E.S. (1972). Additive Method for Combining Probability Values from Independent Experiments. J Psychol *80*, 351-&.

Harrow, J., Denoeud, F., Frankish, A., Reymond, A., Chen, C.K., Chrast, J., Lagarde, J., Gilbert, J.G., Storey, R., Swarbreck, D*., et al.* (2006). GENCODE: producing a reference annotation for ENCODE. Genome Biol *7 Suppl 1*, S4 1-9.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science *324*, 218-223.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Storey, J.D. (2002). A direct approach to false discovery rates. J Roy Stat Soc B *64*, 479-498.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.