High-Throughput Translational Profiling with riboPLATE-seq

Jordan B. Metz[1,2,†], Nicholas J. Hornstein[1,2,3,†], Sohani Das Sharma[1,4], Jeremy Worley[1], Peter A. Sims[1,5,*]

1. Department of Systems Biology, Columbia University Irving Medical Center, New York, NY 10032, USA

2. Medical Scientist Training Program, Columbia University Irving Medical Center, New York, NY 10032, USA

3. Department of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA

4. Department of Biological Sciences, Columbia University, New York, NY 10027, USA

5. Department of Biochemistry & Molecular Biophysics, Columbia University Irving Medical Center, New York, NY 10032, USA

† Co-first authors

* Corresponding author

Contact Information:

Jordan Metz – jm3423@cumc.columbia.edu

Nicholas Hornstein – njh219@gmail.com

Sohani Das Sharma – sohani.dassharma@gmail.com

Jeremy Worley – jw3409@cumc.columbia.edu

Peter Sims – pas2182@cumc.columbia.edu

ABSTRACT

Protein synthesis is dysregulated in many diseases, but we lack a systems-level picture of how signaling molecules and RNA binding proteins interact with the translational machinery, largely due to technological limitations. Here we present riboPLATE-seq, a scalable method for generating paired libraries of ribosome-associated and total mRNA. As an extension of the PLATE-seq protocol, riboPLATE-seq utilizes barcoded primers for pooled library preparation, but additionally leverages rRNA immunoprecipitation on whole polysomes to measure ribosome association (RA). We demonstrate the performance of riboPLATE-seq and its utility in detecting translational alterations induced by inhibition of protein kinases.

KEYWORDS

RNA-seq, translation, ribosome, immunoprecipitation, mTOR, MNK

BACKGROUND

The cellular responses to many physiologic stimuli require new programs of protein production. Transcriptional regulation allows direct control of gene expression over a broad dynamic range, but cells can often more rapidly adjust protein expression levels through translational control. Consequently, alongside transcription factors and their associated regulatory networks, there are mechanisms of modulating the translation of specific genes. mTOR is an important example of a translational regulator that integrates many potential extracellular signals to regulate cellular metabolism and protein synthesis. Activated through the PI3K/Akt/mTOR signaling axis, mTORC1 phosphorylates eIF4E inhibitors (4E-binding proteins, or 4E-BPs), which releases eIF4E and promotes formation of the eIF4F complex in

the initial steps of translational initiation[1]. The actions of mTORC1 are mediated in a sequence-specific manner by 5' terminal oligopyrimidine (5'TOP) motifs, which are C/T-rich sequences in the 5' UTRs of mTORC1 target transcripts[2]. The mTOR protein, the 4E-BP/eIF4E axis, and the 5'TOP tract-containing genes (TOP genes) constitute a basic translational regulatory network.

Despite the attention garnered by profiling and modeling transcription control networks, less progress has been made in understanding systems-level translational control. This is in part due to technological limitations of current translational profiling protocols, which lack the scalability for coupling measurements of protein synthesis with a large number of perturbations. Early genome-wide studies of translational regulation combined polysome profiling and microarray analysis to quantify ribosome association on a gene-by-gene basis[3]. The combination of nuclease footprinting of ribosomes[4] and deep sequencing led to the development of ribosome profiling, which refines translational profiling by resolving the positions of bound ribosomes throughout the transcriptome with single-nucleotide resolution[5] More recent modifications expand on this concept, such as cell type specificity through recombinant tagging of ribosomes driven by cell type-specific markers (e.g. RiboTag)[6,7], increased sensitivity via ligation-free ribosome profiling[8], or targeted profiling in specific sub-cellular compartments[9]. Although these approaches are amenable to detailed mechanistic analysis of translational control in a small number of samples, they are prohibitively expensive and labor-intensive to scale for concurrent analyses of multiple perturbations across a larger sample set. The ideal technology for systems-level analysis would couple genome-wide perturbations to a genome-wide readout of translation, allowing direct observation of translational alterations in specific genes in response to a systematic screen of potential perturbations.

Here we present riboPLATE-seq, a scalable method for generating paired libraries of ribosome-associated and total RNA, which is based our recently reported Pooled Library Amplification for Transcriptome Expression (PLATE-seq) technology[10]. PLATE-seq allows highly-multiplexed RNA-seq, by

71 introducing sample-specific barcodes during reverse transcription to enable pooling of cDNA from
72 multiple individual samples at an early stage of library preparation, reducing both reagent and labor
73 costs. Furthermore, as PLATE-seq generates cDNA fragments strictly from the 3' ends of intact,
74 polyadenylated RNA via oligo-dT pulldown and priming, the resulting libraries saturate more rapidly
75 than those with full gene body coverage. Consequently, PLATE-seq is advantageous over conventional
76 RNA-seq in throughput of both library preparation and sequencing.

77 We take advantage of riboPLATE-seq for parallel, genome-wide translational profiling in 96 samples,
78 which would be technically challenging with conventional polysome or ribosome profiling. riboPLATE-
79 seq uses pan-ribosomal immunoprecipitation to isolate the ribosome-associated fraction of each species
80 of polyadenylated transcript, enabling inferences regarding gene-specific translation akin to polysome
81 profiling. Using PLATE-seq as a readout for ribosomal IP, riboPLATE-seq enables high-throughput
82 translational profiling and screening of potential translational regulators. While riboPLATE-seq measures
83 the abundance of ribosome-bound mRNA rather than nucleotide-resolved ribosome density (as in
84 ribosome profiling), it is highly scalable, inexpensive, and seamlessly compatible with automated liquid
85 handling.

86 In this study, we use riboPLATE-seq to interrogate translational regulation mediated by the
87 PI3K/Akt/mTOR and MAPK/ERK signaling pathways. By treating cells with inhibitory drugs targeting
88 mTOR, PI3K, and MNK1 in a multi-well plate, we sought to uncover specific per-gene signatures of
89 altered translation corresponding to loss of function in these kinases. Additionally, we tested pairwise
90 combinations of these inhibitors to characterize their potential interactions. We also generated
91 signatures of mTOR and MNK1 inhibition in ribosome profiling and RNA sequencing for comparison to
92 established methods of interrogating translation. Importantly, ribosome profiling provides a measure of
93 ribosome-mRNA association that is quantitatively distinct from that of riboPLATE-seq. The ratio of
94 aligned reads in ribosome profiling over aligned reads in RNA-seq libraries obtained from the same

biological sample corresponds to the average number of ribosomes bound per individual transcript. In contrast, the ratio of aligned reads in riboPLATE-seq over normal PLATE-seq corresponds instead to the fraction of the total transcribed pool of transcripts bound by ribosomes.

RESULTS

riboPLATE-seq Technology

riboPLATE-seq enables transcriptome-wide measurements of ribosome association in a multi-well plate format by combining pan-ribosomal immunoprecipitation (IP) with a low-cost technique for RNA-seq called PLATE-seq (Figure 1A). In PLATE-seq, we isolate polyadenylated RNA species with an oligo-d(T) capture plate, followed by incorporation of a well-specific barcode in poly(T)-primed reverse transcription. After mixing the barcoded cDNA libraries from each well of a plate, we conduct all subsequent library preparation steps on a single, pooled sample and sequence the resulting libraries to a modest depth (~4 million reads per well). Previous studies have used ribosome IP to isolate ribosome-bound mRNA from specific cell types *in vivo* with the translating ribosome affinity purification (TRAP)[11] and RiboTag[6] systems, relying on transgenic or recombination-driven epitope labeling of ribosomal proteins. In riboPLATE-seq, we use a native epitope in the 5.8S rRNA for pan-ribosomal IP. By comparing transcript abundance as measured by PLATE-seq both with and without ribosomal IP, we can measure gene-specific ribosome association across the transcriptome. In this way, riboPLATE-seq can extend the scalability of PLATE-seq from transcriptional to translational profiling.

To implement riboPLATE-seq, we divide polysome lysates from a multi-well plate experiment into two plates. We then subject one plate to indirect, pan-ribosomal IP on an automated liquid handling system with biotinylated anti-rRNA antibody y10b and streptavidin-coated magnetic beads. Finally, we generate PLATE-seq libraries from the immunoprecipitated polysomes from the first plate and total lysate from

the second plate as described previously. This design minimizes sample-to-sample noise due to batch effects by processing an entire plate as one batch, and simultaneous processing of a plate via automated liquid handling systems significantly reduces the time and effort required for processing large numbers of samples. Furthermore, per-sample reagent costs are substantially lower in PLATE-based library preparations. PLATE-seq generates 3'-end RNA libraries for under $6 per sample in reagents, while riboPLATE-seq requires additional expenses for automated ribosome IP totaling ~$7 per sample for reagents and disposables, all in 96-well plate format. A full riboPLATE-seq study performed on one 96-well plate costs less than $20 in materials per sample to generate paired libraries of ribosome-associated and total RNA, compared with $86 for ligation-free ribosome profiling and RNA sequencing.

riboPLATE-seq Translational Profiling Screen

To characterize the performance of riboPLATE-seq, we designed an experiment to identify the translational impact of inhibiting multiple components of mitogenic signaling in cancer cells. TS-543 glioma neurospheres harbor an activating mutation in PDGFRA, leading to increased mitogenic signaling activity[12] that exacerbates the effects of inhibiting these pathways. We treated cells seeded in a 96-well plate with inhibitors of the PI3K/Akt/mTOR and MAPK/ERK signaling axes for six hours (Figure 1B), as both pathways converge on the ribosome at eIF4E and are thought to influence the formation of the eIF4F complex. This complex, formed from the cap-binding protein eIF4E, the RNA helicase eIF4A, and the scaffold protein eIF4G, is required for cap-dependent translation initiation, and the association of eIF4E and eIF4G is a heavily regulated component of this process. PI3K activates AKT which phosphorylates mTOR, and activated mTOR facilitates eIF4F formation by phosphorylation of eIF4E-interacting proteins (4E-BPs), which then release eIF4E[1,2]. Separately, the MAPK signaling cascade activates MNK1, which phosphorylates eIF4E directly, increasing its affinity for the 5' m7G cap and stabilizing eIF4F[13]. By screening several members of these two pathways we sought to identify targets specific to each individual member and compare translational signatures between pathways. The drug

142 treatments in total consisted of two competitive mTOR inhibitors, PP242 and AZD-8055; an inhibitor of
143 PI3K upstream of mTOR, BKM120; a specific inhibitor of MNK1/2 activity, MNK-i1; and 4EGi-1, a 4E-BP
144 mimic that inhibits the association of eIF4E and eIF4G. We determined concentrations of these drugs
145 from an examination of the literature, ensuring values near the half-maximum inhibitory concentrations
146 (IC50) for the main substrates of the drugs in question: 625nM PP242[14], 50nM AZD-8055[15], 1μM
147 BKM120[16], 100nM MNK-i1[17], and 50μM 4EGi-1[18]. In order to analyze possible interactions between
148 kinases, we also treated samples with pairwise combinations of PP242, BKM120, and MNK-i1.

149 Previous studies of the effect of MAP kinase interacting kinases (MKNKs, a.k.a. MNKs 1/2) on
150 translational regulation utilized small-molecule inhibitors of these proteins, notably the compound
151 CGP57380[19,20]. However, this compound has been shown to be a nonspecific inhibitor of several
152 unrelated kinases, with effects on eIF4F formation independent of its effects on MNKs. CGP57380 has
153 low-micromolar IC50 values for MNK isoforms (0.87 μM/ 1.6μM for MNK1/MNK2, respectively) and
154 significantly inhibits other kinases at these concentrations, including MKK1, CK1, and BRSK2[21].
155 Additionally, CGP57380 concentrations below that which affects eIF4E phosphorylation may still
156 decrease proliferation and survival, and an increase in eIF4E:4EBP binding occurs at concentrations
157 below those impacting MNK1, indicating broad off-target effects impinging translational regulation[22].
158 Determination of the translational targets of MNK1 via ribosome or polysome profiling with this drug in
159 prior work is complicated by this lack of specificity, especially with regards to off-target effects directly
160 impacting translational machinery. In contrast, MNK-i1 has been recently identified as a highly specific
161 MNK inhibitor, with IC50 values of 0.023 μM and 0.016 μM for MNK1 and MNK2 respectively, and blocks
162 eIF4E phosphorylation without impacting other pathways converging on eIF4E[22]. We therefore sought to
163 clarify the effect of MNK1 on translation with this novel inhibitor.

For comparison, we performed ribosome profiling and RNA-seq on TS-543 neurospheres treated with PP242 or MNK-i1 in identical regimens to the riboPLATE-seq study, in order to assess the similarity of translational perturbations detected across experiment types. As riboPLATE-seq measures the fraction of an expressed transcript associated with ribosomes (hereafter referred to as "ribosome association" or RA) and ribosome profiling/RNA-seq measure the average number of ribosomes bound per transcript (conventionally defined as "translation efficiency" or TE), we expected these two methods to give quantitatively distinct results while identifying similar sets of targets for these translational regulators.

Performance of riboPLATE-seq

First, we assessed the quality of the pooled ribosome-associated riboPLATE-seq and normal PLATE-seq libraries in terms of library complexity and saturation. Figures 2A and 2B show saturation curves for riboPLATE-seq and PLATE-seq, respectively, demonstrating the dependence of these libraries' sensitivities on read depth. The two curves are comparable, with ~10-11K unique genes detected at saturating depth, though riboPLATE-seq requires about twice the number of aligned reads as PLATE-seq to achieve this saturation. As expected, PLATE-seq libraries contain more unique genes on average than riboPLATE-seq libraries despite a shallower sequencing depth. Asymmetric division of initial lysate volumes favoring ribosome IP over unmodified PLATE-seq (90%/10%) helped to combat this inherent inequality, generating sufficiently complex libraries in both cases.

Figure 2C highlights the differences between riboPLATE-seq and PLATE-seq in terms of library complexity and sequencing depth, and compares these with libraries generated by ligation-free ribosome profiling and conventional RNA-seq. On average, riboPLATE-seq detects approximately 9,800 unique genes in 1.4 million uniquely mapped reads, while PLATE-seq detects an average of 10,400 genes in 0.6 million reads per sample in this study. These measurements are comparable to the initial report characterizing PLATE-seq, in which Bush et al. detected an average of approximately 10,200 genes from

187 0.67 million uniquely mapped reads per sample[10]. In contrast, ligation-free ribosome profiling and total
188 RNA sequencing libraries downsampled to the respective median read depths of riboPLATE-seq and
189 PLATE-seq libraries still detect ~14,000 genes each, reflecting the inherent complexity of these libraries
190 even at reduced sequencing depths (Additional File #1: Supplementary Figure S1). At full depth,
191 ribosome profiling libraries detect an average of 14,000 genes per 2 million reads, while RNA-seq
192 detects 15,000 genes in 1.6 million reads per sample. In summary, both riboPLATE-seq and PLATE-seq
193 generate libraries of a lower overall complexity than ribosome profiling and RNA sequencing, and
194 require substantially fewer reads to achieve saturation.

195 To determine the specificity of pan-ribosomal IP for ribosome-bound RNA, we measured the depletion
196 of RNA species for which we expected little or no ribosome association (RA) in the riboPLATE-seq vs
197 PLATE-seq libraries. We measured depletion extrinsically, with respect to a set of polyadenylated spike-
198 in RNAs added after lysis, and intrinsically with respect to the set of highly-expressed, polyadenylated
199 non-coding RNA transcripts (ncRNA) contained in the UCSC “known genes” RefSeq annotation. Figures
200 2D-E summarize these two analyses.

201 To assess the depletion of spike-in RNA in riboPLATE-seq, we added polyadenylated RNA standards
202 (ERCC spike-ins) to half of the wells in a riboPLATE-seq experiment, after lysis but prior to ribosome IP.
203 Figure 2D shows the distribution of the log2-ratio of spike-in abundance in riboPLATE-seq to PLATE-seq,
204 demonstrating depletion of spike-ins associated with ribosomal IP across most wells. The wells exhibit a
205 median 4.5-fold depletion ratio, and 31/48 wells exhibit 4-fold depletion or greater. However, two wells
206 exhibit a modest (<2-fold) enrichment of spike-ins after ribosome IP, and the wide distribution of
207 depletion log2-ratios ranges from 11 to 0.7 (equivalent to ~1.6-fold relative enrichment). Aside from
208 potential non-specific pulldown of spike-ins, random re-initiation of free ribosomes on polyadenylated
209 transcripts in lysate could result in their capture by ribosome immunoprecipitation, resulting in their

enrichment in riboPLATE-seq. This might be minimized in future riboPLATE-seq studies by inclusion of GDPNP in the lysis buffer. As GTP hydrolysis is required in both start site selection and subunit joining steps of 80S initiation complex formation, inclusion of a non-hydrolyzable GTP analogue such as GDPNP would prevent re-initiation of free ribosomes in lysate[23]. In summary, inclusion of ERCC RNA spike-ins provides a valuable, internal quality control measurement to check IP fidelity in riboPLATE-seq.

As ribosome profiling has revealed low but significant levels of ribosome occupancy among ncRNA[24], we sought to contrast ribosome association between ncRNA and mRNA transcripts with riboPLATE-seq. We expected noncoding transcripts to be generally depleted in our riboPLATE-seq libraries compared with PLATE-seq libraries from the same samples. Indeed, we observed lower RA for the set of highly-expressed ncRNA transcripts than mRNA within the same sample (Figure 2E). Examining the relationship between RA and transcript abundance across ncRNA and mRNA gene sets also uncovered lower RAs for ncRNA than mRNA at all expression levels (Figure 2F), with similar patterns observed between translation efficiency (TE) and expression level in our ribosome profiling and RNA-seq data (Figure 2G). Combined with the observed spike-in depletion, our results are consistent with the depletion of RNA that is not bound to ribosomes by the pan-ribosomal immunoprecipitation implemented in riboPLATE-seq.

Pharmacological Screening of Mitogenic Signaling with riboPLATE-seq

After establishing the performance of the riboPLATE-seq, we sought to characterize its ability to detect differential expression and RA. In principal component analyses (PCA) of the PLATE-seq and riboPLATE-seq profiles, samples segregate according to the drug with which they were treated and related drugs co-cluster (Figure 3A, B). Principal component 1 (PC1) separates DMSO-treated controls from drug-treated samples, and PC2 separates samples treated with one kinase inhibitor from those treated with

4EGi1 or a combination of kinase inhibitors. Combination-treated samples co-cluster more readily with each other than with any of their singularly-treated counterparts due to separation on PC2.

We further calculated RA for all genes in all samples using riboPLATE- and PLATE-seq counts and differences in RA across the genome relative to DMSO-treated controls. PCA plots of RA and log-fold change in RA are shown in Figures 3C and 3D, respectively. In both, PC2 separates combination- and 4EGi1-treated samples from singularly-treated samples (PP242, BKM120, MNK-i1, or AZD8055 alone), as it does in the previous plots in Figure 3A and 3B, while PC1 separates drugs by the number of genes displaying significant perturbations in RA compared with controls. In all plots, clusters of samples treated with singular drugs are organized in a manner consistent with their expected effects. Clusters corresponding to the three inhibitors of the mTOR signaling axis (PP242, BKM120, and AZD8055) nearly overlap in Fig. 3A and are very closely spaced in Fig. 3B, while MNK-i1-treated samples cluster more closely to the DMSO-treated controls, suggestive of weaker effects.

This division between strong and weak inhibitors also appears in the RA-dependent plots in Figure 3C and 3D. In both, MNK-i1-treated samples have more negative PC1 values than any of those treated with mTOR axis inhibitors, while PP242-treated samples have the most positive PC1 values of any single kinase inhibitor. For PC2>0, the clusters corresponding to different singular kinase inhibitors are arranged in order of the number of significant RA perturbations they exhibit in comparison to DMSO controls (MNK-i1, AZD-8055, BKM120, then PP242; Spearman rank correlation p=0.0). Below the PC2=0 line, clusters which correspond to combination and 4EGi-1 treatments are similarly arranged along PC1 (4EGi-1, MNK-i1/BKM120, PP242/MNK-i1, PP242/BKM120; Spearman rank correlation p=0.0) (Table 2).

riboPLATE-seq Differential Translation Efficiency Analysis

In order to more rigorously analyze differential ribosome association as a function of drug treatment, we utilized DESeq2 to compare the replicates for each drug treatment to vehicle controls. We first identified

the total set of significantly differentially ribosome-associated genes across all singular drug treatments and generated a hierarchically-clustered heatmap of differential RA across conditions (Figure 4A). Signatures of differential RA due to treatment with PP242 (mTOR inhibitor), AZD8055 (mTOR inhibitor), and BKM120 (PI3K inhibitor), which target the PI3K/mTOR pathway, form a cluster in the column dendrogram. As expected, MNK-i1 and 4EGi-1 targets cluster separately. MNK-i1 targets MNK1 in a separate pathway converging on the ribosome at eIF4E, and 4EGi-1 is a broad inhibitor of eIF4E and cap-dependent translation in general.

The closely related differential-RA signatures for PP242, BKM120, and AZD8055 also include strong downregulation of the 5'TOP motif-containing genes, canonical translational targets of mTOR signaling, as indicated by the black tick marks in Figure 4A. In the rightmost two columns of Figure 4A, signatures of differential translation efficiency obtained via ribosome profiling and RNA-seq recapitulate the major patterns seen in differential RA. Both up- and down-regulated targets of PP242 are in good agreement between the two methods. The 5'TOP motif-containing genes exhibit low TE by ribosome profiling after PP242 treatment, whereas MNKi-1 treatment is far less effective on these genes based on ribosome profiling and leads to fewer differentially translated genes in general, consistent with riboPLATE-seq.

We used gene set enrichment analysis (GSEA) to identify gene ontologies associated with differential ribosome association under each drug treatment (Fig. 4B), based on effect size of RA change associated with each drug. Across all drugs, all ontologies with low family-wise error rate (FWER<0.001) have negative normalized enrichment scores (NES), indicating a predominately inhibitory effect of these drugs on ribosome association. All treatments downregulate genes related to translation initiation and ribosomal components, as expected for inhibition of translation-activating signaling pathways. MNK-i1 treatment inhibits these ontologies despite a relatively weak effect on the TOP motif-containing genes, suggestive of mTOR-independent inhibition of translation. Surprisingly, all kinase-dependent treatments

exhibit stronger, more consistent downregulation(s) of these genes than 4EGi-1; as this drug targets eIF4E directly, we expected TOP genes to be included in the set of its strongly-inhibited targets.

Attenuation of Perturbations to Ribosome Association in Drug Combinations

We expected that drug combinations would elicit greater changes in RA than the individual drug treatments alone. Specifically, we expected at least additivity if not outright synergy from simultaneous inhibition of the PI3K/Akt/mTOR and MAPK/ERK pathways, and a similar but perhaps less pronounced additivity of effects from inhibiting kinases in the same pathway (i.e. mTOR and PI3K). Surprisingly, we instead found a pattern of attenuation of the strongest effects of individual drugs when combined. Figure 5 compares the drug combination-treated samples to those treated with individual drugs, revealing attenuation of the strongest single-drug effects in combination. The volcano plots in Figure 5A-C highlight the significantly downregulated targets of the three individual kinase inhibitors PP242, MNK-i1, and BKM120, defined by thresholds on the significance and magnitude of RA change (FDR<0.01 and log-fold change in RA<-0.75, respectively). As a set, these genes exhibit comparatively diminished, less significant changes in ribosome association in combination-treated samples (Figure 5D-F), though drug combinations do not result in fewer significant RA changes overall (Table 2). The scatterplots in Figure 5G-I emphasize this unexpected result of combining translation inhibitors by explicitly plotting changes in effect size for significant targets of individual drugs in combined treatments. Both positive and negative effects are diminished overall in each pairwise combination, though we observed less attenuation among individual targets of PP242 and BKM120 under combination treatment with PP242+BKM120 (Additional File #1: Supplementary Table 1).

Motif-Based Target Classification in a Translation Control Network

Finally, we displayed the results of our study in network form and mapped occurrences of a known translational cis-regulatory element, the 5' TOP motif, across this network. Following the observation of

concordant regulation of canonical TOP genes in drug treatments impacting the mTOR signaling pathway, we first sought to expand the potential set of TOP genes within the strongly-perturbed genes in this study. We first obtained sets of canonical TOP genes and candidates containing previously uncatalogued 5'TOP tracts, a subset of which have known TOP-containing homologues in the mouse genome, from the comprehensive analysis of human transcription start sites performed by Yamashita et al[25]. This yielded a total set of 1,626 TOP gene candidates: within this set, 237 candidates have mouse homologues that are known TOP genes, and this subset overlaps substantially with the 83-member set of canonical TOP genes identified (54/83 mouse homologues/canonical TOP). We found these TOP candidates to behave similarly to canonical TOP genes in terms of perturbed RA. In the strip plots in Figure 6A, TOP genes and candidates within the significant targets (FDR<0.05) of each drug on the plate are color-coded, allowing comparison of their differential RA between conditions.

We then constructed a simple translational regulatory network from our riboPLATE-seq data and overlaid it with these canonical and novel TOP genes (Figure 6B). We considered the genes demonstrating significant reductions in RA due to treatment with PP242, MNK-i1, and BKM120 as the positive translational targets of mTOR, PI3K, and MNK1, respectively, as these drugs are specific inhibitors of these kinases. Here, we used a typical threshold for significance (FDR<0.05), but additionally required target genes to have at least 20 average normalized read counts across all samples.

As expected, the targets of mTOR are enriched heavily for TOP genes and candidates. Nearly all known TOP genes and candidates with mouse homology, as well as a significant majority of remaining TOP gene candidates, are targets of mTOR. The largest fraction of each set is found in either the joint targets of mTOR and PI3K or the exclusive targets of mTOR. The largest set of targets belongs to mTOR (228 exclusive/386 total), followed by PI3K (57 exclusive/213 total) and then MNK1 (21 exclusive/43 total), and TOP genes and candidates comprise a larger fraction of mTOR targets than PI3K or MNK1 targets.

The set of targets common to all three kinases is also highly enriched for curated and novel TOP genes (Fisher's exact test p=0.00006), suggesting a subset of TOP genes impacted by MNK1, though this three-way intersection is vastly smaller and less significantly-enriched for these genes than the two-way intersection of mTOR and PI3K targets (Fisher's exact test p=6x10-24). Furthermore, the targets exclusive to either MNK1 or PI3K are not significantly enriched for TOP genes and candidates (Fisher's exact test p=0.760 and p=0.219, respectively) in contrast to the exclusive targets of mTOR (p=5x10-8). Supplementary Table 2 details a more comprehensive statistical analysis of the network in Figure 6B, which considers TOP genes and candidates separately, while the calculations above consider enrichment across the combined set of genes and candidates (Additional File #1: Supplementary Table 2). This combined regulatory network suggests a relatively minor effect of MNK1 on TOP gene translation, and underscores the dependence of the effect of PI3K on TOP genes on mTOR downstream.

DISCUSSION

Ribosome association is frequently used to infer translational activity. This can be measured by sucrose gradient fractionation of intact RNA in polysome profiling, or of digested monosomes and their ribosome-protected footprints in ribosome profiling. Translation efficiency, defined as the rate of protein production per transcript, is approximated differently in these two methods. In polysome profiling, it is calculated as the ratio of transcripts that sediment in "heavy" vs "light" fractions, similar to the ratio of ribosome association in riboPLATE-seq. Ribosome profiling refines this measurement with its focus on ribosome footprints, calculating instead a per-transcript ribosome occupancy with additional information about position, regional density, and ribosome arrest[26]. riboPLATE-seq sacrifices the specific positional information provided by ribosome profiling for a general measurement of ribosome association, obtained by IP rather than sucrose gradient fractionation. With pooled library construction,

greater throughput is possible with riboPLATE-seq than with either ribosome profiling or polysome profiling.

Though it is more scalable than ribosome profiling, riboPLATE-seq is not without limitations. The lack of resolution of individual ribosome positions means the method cannot resolve location-specific effects, such as the effect of ribosome association in 5' leader sequences on translation in downstream coding sequences[7]. More generally, riboPLATE-seq is insensitive to translational regulation at the levels of elongation or termination due to its inability to distinguish active from stalled ribosomes, a limitation common to many ribosome-centric measurements[27] including ribosome profiling.

In this study, we interrogated translational regulation in mitogenic signaling pathways in cancer cells. Focal amplification of PDGFRA in TS-543 glioma neurospheres leads to constitutive activation of several members of these pathways, including ERK, Akt, and PI3K[12]. We observed the expected results of mTOR inhibition on translation with riboPLATE-seq, including decreased ribosome association with TOP genes, and further clarified targets for the kinases PI3K and MNK1. We found PI3K to target a subset of the TOP genes impacted by mTOR, without a strong impact on known TOP genes or candidate transcripts separate from that shared with mTOR. This suggests the effect of PI3K on the TOP genes may be wholly mediated by mTOR, consistent with the known organization of their signaling pathway. In contrast, treatment with either the highly-specific MNK inhibitor MNK-i1 or the eIF4E inhibitor 4EGi-1 did not significantly impact the TOP genes. Despite both drugs impacting the effector through which mTOR is thought to mediate TOP gene translation, eIF4E, off-target effects of commonly-used MNK inhibitors in past studies[22] may have overemphasized previous observations to this effect, and potential off-target effects of 4EGi-1[28] complicate interpretation of its specific, eIF4E-dependent targets.

Pairwise combinations of inhibitors did not generate additive effects on ribosome association, instead triggering what we interpreted as compensatory regulation and attenuating the strongest effects of individual inhibitors. Surprisingly, combination treatment with PP242 and BKM120 attenuated their individual targets the least, despite the seeming redundancy of these drugs targeting the same pathway. Though still attenuated overall, a greater portion of the individual targets of PP242 and BKM120 were enhanced by combination treatment. This might suggest that PI3K inhibition does not saturate inhibition of its downstream effectors, including mTOR. These surprising findings highlight the utility of explicitly testing combined perturbations and the need for scalable measurement strategies like riboPLATE-seq.

CONCLUSION

This study serves as a proof-of-concept for larger-scale perturbation screens of potential translational regulators. Here, riboPLATE-seq revealed signatures of specific translational targets for kinases in related signaling pathways. Our results are consistent with the known structure of these pathways, including the previously established mechanism by which mTOR controls translation of the TOP motif-containing genes. However, the majority of the ~500 known protein kinases remain unstudied at the level of translational regulation. The technology described here could enable a more comprehensive screen of protein kinases and/or RNA binding proteins, allowing inference of translational regulatory networks and *de novo* identification of regulatory motifs important to these networks that might be validated by high-resolution techniques like ribosome profiling and CLIP-seq. We anticipate that the ability to dissect these networks at scale will advance our understanding of translational regulation and the design of specific therapies for diseases involving aberrant translation.

METHODS

Tissue Culture and Drug Treatment

We seeded TS-543 neursopheres (passage #11) on a 96-well plate (Corning, #3799) at a density of 7,500 cells per well (50,000 cells/mL) in 150uL NS-A complete medium (containing 10% v/v NeuroCult NS-A Proliferation Supplement, 20ng/mL EGF, 10ng/mL bFGF, and 2ug/mL heparin) (STEMCELL Technologies #05751). We incubated the plate of cells for 36 hours prior to the start of the experiment at 37°C and 5% CO2 in a tissue culture incubator. We separately prepared stock solutions of PP242 (Tocris, #4257), MNK-i1, NVP-BKM120 (Selleck, S2247), AZD8055 (Selleck, S1555), and 4EGi-1 (Tocris, #4800) in DMSO vehicle (Sigma, #472301). After dilution with NS-A basal culture medium (without supplement, cytokines, or heparin) (STEMCELL Technologies #05751), we administered the drugs or pure DMSO to the experimental and control wells, respectively, in 1uL doses. Final concentrations were 50nM AZD8055, 625nM PP242, 1µM BKM120, 100nM MNK-i1, and 50µM 4EGi-1, including in pairwise combination-treated samples. Drug treatment proceeded for 6 hours in the tissue culture incubator prior to lysis.

Cell Lysis

Following treatment, we centrifuged the plate of TS-543 for 7 minutes at 1800RPM on a Sorvall Legend XTR at room temperature and removed supernatants by aspiration. Placing the plate on ice, we resuspended the pelleted cells in each well in 30uL of polysome lysis buffer (20mM Tris-HCl, pH=7.4, 250mM NaCl, 15 mM $MgCl_2$),0.1mg/mL cycloheximide, 0.5% Triton X-100, 1mM DTT, 0.5U/mL SUPERase-In (ThermoFisher, AM2696), 0.024U/mL TURBO DNase (Life Technologies, AM2222), 1x Protease Inhibitor (Sigma, P8340)), mixed 5 times by pipetting, and rested the plate on ice for 5 minutes. We then centrifuged the plate for 5 minutes at 1400RPM at 4°C to remove bubbles before performing a quick freeze-thaw, placing the plate first in a -80°C freezer and then resting at room temperature for 5 minutes each. Following an additional 10 minutes rest on ice, we viewed the plate under a microscope to check the extent of cell lysis. At this point, we added standard RNA spike-ins (ERCC Spike-In Mix #1, 1

418 uL/well of 1:5000 diluted stock) (Life Technologies, #4456740) to half of the wells for spike-in depletion
419 experiments. We then prepared a new 96-well plate containing 3.5uL 2x TCL buffer (Qiagen, #1070498)
420 per well, to which we transferred 3.5uL of lysate (approximately 10% total volume).

421 **Automated Pan-Ribosome Immunoprecipitation**

422 To the remaining lysate, we added 1 uL of SUPERase-in (ThermoFisher, AM2696) and 1 uL of biotinylated
423 y10b antibody (ThermoFisher, MA516060) to each well, then sealed the plate and allowed it to incubate
424 while gently shaking for 4 hours at 4°C. During this incubation, we washed 500uL of Dynabeads MyOne
425 Streptavidin C1 streptavidin-coated magnetic beads (ThermoFisher, #65001) 3 times with polysome
426 wash buffer (20mM Tris-HCl (pH 7.4), 250mM NaCl, 15mM MgCl2, 1mM DTT, 0.1mg/mL cycloheximide,
427 0.05% v/v Triton X-100), using 1mL per wash and resuspending in 500uL. We added 5uL of washed
428 beads to each well, then incubated while gently shaking at 4°C for an additional hour. After this short
429 incubation, we placed the plate on a magnet, removed and reserved supernatants, and washed the
430 wells 3 times with 200uL per well of polysome wash buffer supplemented with 1uL/mL SUPERase-in on
431 the Biomek 4000 automated liquid handling system.

432 Following the final wash, we resuspended the beads in 15uL of ribosome release buffer (20mM Tris-HCl
433 (pH 7.4), 250mM NaCl, 0.5% Triton X-100, 50mM EDTA) per well. During a 15-minute incubation at 4C
434 on a Peltier module, with continuous pipet mixing on the Biomek 4000 in order to maximize elution, we
435 distributed 15uL of 2x TCL buffer to each well of a new 96-well plate. Finally, we replaced the sample
436 plate on the magnet and transferred eluants to the TCL-containing plate.

437 **PLATE-seq Library Preparation and Sequencing**

438 The plates of ribosome-associated and previously reserved total lysate in TCL buffer were submitted to
439 the Columbia Genome Center for processing by the previously-described PLATE-seq method of RNA-seq
440 library preparation[10], which involves poly-A selection of transcripts, incorporation of sequence barcodes

in poly(T)-primed reverse transcription, and pooling for subsequent library preparation steps, generating a single 3'-end RNA-seq library from each 96-well plate. We pooled total and ribosome-associated PLATE-seq libraries and sequenced the combined libraries on the Illumina NextSeq 550 with a 75-cycle high-output kit. With paired-end sequencing, the first read corresponds to the 3' end of a transcript, and the second read contains the barcode identifying the library in which the read was obtained.

Ribosome Profiling and RNA Sequencing

We seeded TS-543 neurospheres in a 6-well plate at a starting density of 50,000 cells/mL in 2 milliliters of NS-A complete medium per well, and allowed the plate to rest for 36 hours. After preparing PP242 and MNK-i1 solutions in DMSO as above, we treated two wells each with 625nM PP242, 1.0µM MNK-i1, or DMSO vehicle for 6 hours in the tissue culture incubator. Following treatment, we transferred samples to 15mL conical vials for centrifugation at 640 RCF for 7 minutes, then removed supernatants and added 400 uL polysome lysis buffer (recipe above). After mixing by rapid pipetting, we transferred samples to 1.8mL microcentrifuge tubes, rested them on ice for 5 minutes, and triturated by 5 passages through a 23-gauge needle. Following a clarifying spin of 11K RCF for 10 minutes at 4C on a benchtop centrifuge, we transferred supernatants to a new set of microcentrifuge tubes and discarded pellets. Finally, we prepared ligation-free ribosome profiling and total RNA-seq libraries from these clarified polysome lysates following the instructions provided with their respective kits (smarter-seq smRNA-seq kit, Takara-Clontech; NEBnext Ultra-Directional II), augmented with our previously-published ligation-free ribosome profiling protocol[8]. We sequenced 6 ribosome profiling libraries or up to 12 RNA-seq libraries in one NextSeq 550 high-output 75-cycle kit.

Both PLATE-seq and ligation-free ribosome profiling library preparation protocols are available on our laboratory website[29].

Read Alignment and Data Analysis

Using a custom pipeline, we first trim reads of trailing polyA sequence and adapters with fastx-clipper[30] (fastx-toolkit v0.0.14), then align the multiplexed reads to the hg19 assembly of the human genome and known RNA spike-in sequences with STAR v2.4.0. We then demultiplex the aligned reads to their original riboPLATE- or PLATE-seq libraries according to their barcodes (available in BarcodesHiSeq.txt). We use a similar pipeline to process and align ribosome profiling and RNA sequencing libraries. Additional processing steps are required for ribosome profiling in order to trim polyA tails and adapters with fastx_trimmer[30], then remove ribosomal RNA-aligned reads with bowtie2[31] (v2.1.0). We use featureCounts[32] (from Subread v1.4.6) to count the number of fragments aligned to each gene in each library, counting all exon-aligned reads as valid.

Definition of Gene Sets of Interest

As PLATE- and riboPLATE-seq depend on isolation of RNA by poly(T) pulldown, they can only be used to measure polyadenylated transcripts. We defined a set of questionably-polyadenylated transcripts by the union of the set of non-polyadenylated and variably-polyadenylated genes identified in a screen of polyadenylation status across the transcriptome[33]. We removed these genes from consideration in our study to leave only consistently polyadenylated transcripts. We additionally obtained a set of known 5' terminal oligopyrimidine motif-containing genes (TOP genes), as well as novel TOP candidates with and without known TOP-containing analogues in mice, from a comprehensive search of transcription start sites[25].

Regularized Logarithm Normalization and Outlier Removal

After subsetting the count matrices for all libraries to remove counts for reads aligned to non-polyadenylated and spike-in transcripts, we constructed an overall count matrix of all 192 libraries for all 96 samples. We loaded this matrix into DESeq2 with corresponding column data describing the sample ID, library type (ribo or RNA PLATE-seq), and drug treatment for each library. We then used the

regularized logarithm function in DESeq2 with default parameters to obtain a variance-stabilized, log-scale transformation of raw counts for all libraries.

We used this transformed count matrix to perform two-dimensional principal component analyses (PCA) in Python, utilizing the scikit-learn package for analysis and matplotlib for visualization. We removed a total of 15 outliers from the PCA plots of ribosome-associated RNA, total RNA, and rlog-ribosome association, which we defined as the log-difference between the two measurements (rlog(riboPLATE)-rlog(PLATE)). We then constructed a reduced matrix of raw counts corresponding to the remaining 81 samples (see Table 1) and again performed rlog normalization on this matrix in DESeq2. We used this final matrix of rlog-normalized counts for non-outlier samples, as well as the raw count version of this matrix, for RNA quality control and principal component analyses in Figures 2 and 3, respectively.

Differential count analysis in DESeq2

We performed differential expression and differential ribosome association analyses using the DESeq2 package in R[34]. First we constructed a column data table identifying the sample ID, drug treatment, and library type (riboPLATE vs PLATE) for each library. After loading this table and the raw count matrix for non-outlier samples, we performed simultaneous analyses for all drug treatments vs DMSO-treated controls using the following design formula:

```
design = ~condition + condition:ind.n + condition:type
```

In this design formula, condition refers to drug treatment status, type refers to riboPLATE-seq or PLATE-seq library preparation, and ind.n corresponds to sample ID nested within condition, in accordance with the DESeq2 vignette[35] (section *Group-specific condition effects, individuals nested within groups*). Pairing of riboPLATE- and PLATE-seq libraries from the same sample in this way allows DESeq2 to correct for sample-to-sample noise while calculating group-wise effects of drug treatment.

After setting PLATE-seq as the reference level for library type and DMSO as the reference level for drug treatment condition, we executed the DESeq2 function using fitType=local, then retrieved results for each drug treatment and comparison of interest. For differential ribosome association, we used function calls of the following format:

```
res <- results(dds, name=paste0("typeribo.condition",cond))
```

Where dds is the DESeqDataSet object and cond is any of the drug treatments. The interaction between library type and drug treatment is equivalent to the calculation of ribosome association (RA). For differential expression, we considered only changes in gene abundance between PLATE-seq libraries independent of riboPLATE-seq. This required function calls of the following format:

```
res <- results(dds, name=paste0("condition",cond))
```

Where dds is the DESeqDataSet object and cond is any of the drug treatments. This isolates the main effect of drug treatment, defined across total RNA PLATE-seq libraries only.

For ribosome profiling and RNA sequencing, we followed a similar workflow, utilizing a design formula without sample pairing:

```
design = ~condition + type + condition:type
```

where condition corresponds to drug treatment (MNKi1, PP242, or DMSO) and type corresponds to RNA-seq vs ribosome profiling. We then set the base level of `type` to RNA and the base `condition` to DMSO, and executed DESeq2 with `fitType=local`. Finally, we retrieved results specific to the interaction effect, here equal to the calculated change in translation efficiency (TE) as a result of drug treatment, with the following two function calls:

```
res_242 <- results(dds, name = "typeribo.conditionPP242")
```

```
res_MNK <- results(dds, name = "typeribo.conditionMNKi1")
```

Identification of Perturbed Gene Ontologies by Gene Set Enrichment Analysis

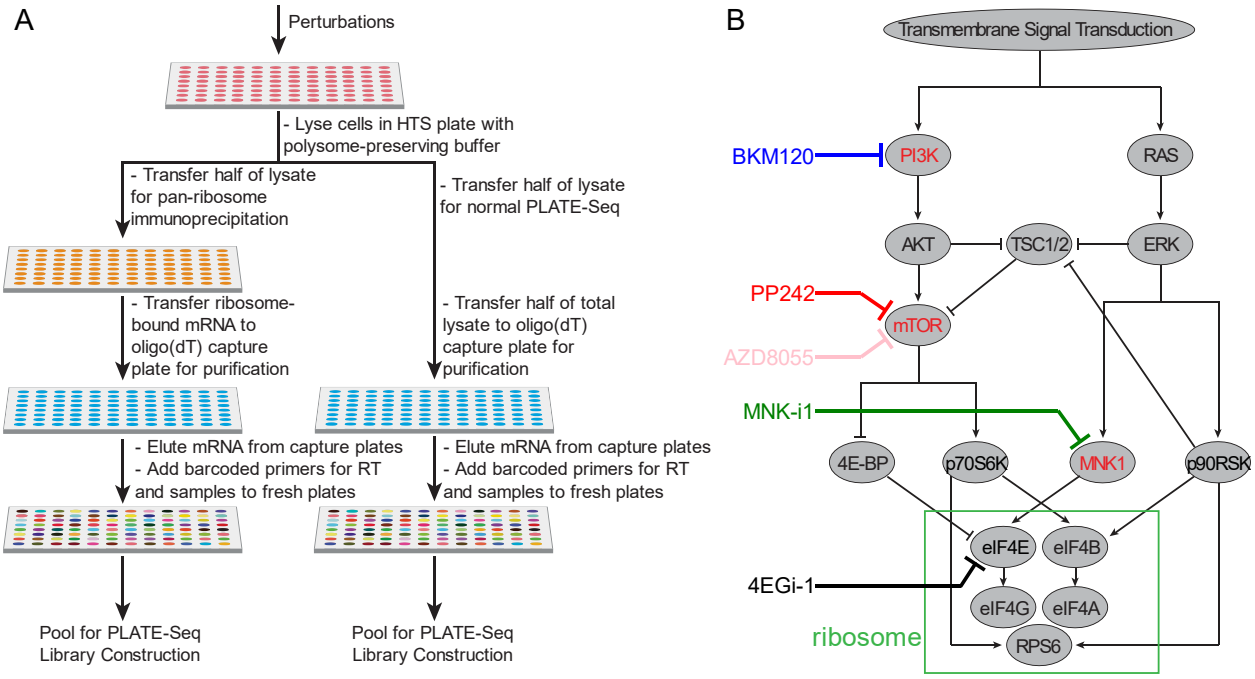
We constructed ranked lists for gene set enrichment analysis (GSEA) using the results of these differential expression and ribosome association analyses. We first removed any gene with invalid results reported by DESeq2 in any drug-vs-control comparison (i.e. genes with assigned p-value 'NA') from consideration. Next, we created a ranked list of genes for each drug treatment, using each gene's log-fold change in RA as a rank statistic. We then performed a preranked GSEA on the C5 collection of gene ontology terms against these ranked lists, using the Preranked option in the GSEA-P desktop application[36], with default parameters and scoring method set to 'classic'.

Network Visualization

To create a basic network, we interpreted the abundant genes exhibiting highly significant reductions in RA under treatment with kinase inhibitors (FDR<0.01, baseMean>=20) as positive targets of the kinases in question. We loaded these gene sets into CytoScape[37] (v2.7.1) as individual networks for each kinase, merged the three networks, and used the yFiles[38] Organic automatic layout to organize the resulting merged network. We then color-coded the sets of canonical and novel TOP motif-containing genes present in the network, based on the lists obtained from Yamashita et al[25].

Data Visualization and Code

We generated plots and diagrams using matplotlib (v3.0.2) and Jupyter Notebook (v5.0.0)[39,40]. Our analyses use NumPy[41] (v1.13.3) for data manipulation, SciPy[42] (v0.14.0) for statistical tests, scikit-learn[43] (v0.19.1) for PCA. We additionally generated strip plots and heatmaps using Seaborn[44] (v0.9.0) in Python.



553

554 **Figure 1: Overviews of the protocol and experimental design of the study performed**. A) Schematic
555 diagram of the riboPLATE-seq protocol, from lysis in a multi-well plate to pooled library preparation. The
556 right-hand side mirrors the original PLATE-seq protocol. In this workflow, an oligo-dT-grafted plate
557 captures polyadenylated RNA that can be reverse-transcribed with barcoded adapters, generating a
558 plate of cDNA that may be pooled for library construction. The left side incorporates a pan-ribosome IP
559 before PLATE-seq pooling and library preparation, generating instead a pooled library of ribosome-
560 associated RNA. B) Simplified structure of the signaling pathways under study and the specific protein
561 targets considered. The PI3K/AKT/mTOR signaling axis at left converges with the MAPK/ERK pathway at
562 right on eIF4E, early in the process of ribosome assembly (green box). The figure also outlines the
563 inhibitors used in this study and their specific targets within these pathways. NVP-BKM120 is a PI3K
564 inhibitor (blue), both AZD-8055 and PP242 are mTOR inhibitors (pink and red, respectively), MNK-i1 is a
565 MNK1 inhibitor (green), and 4EGi-1 is a direct eIF4E inhibitor (black).

566

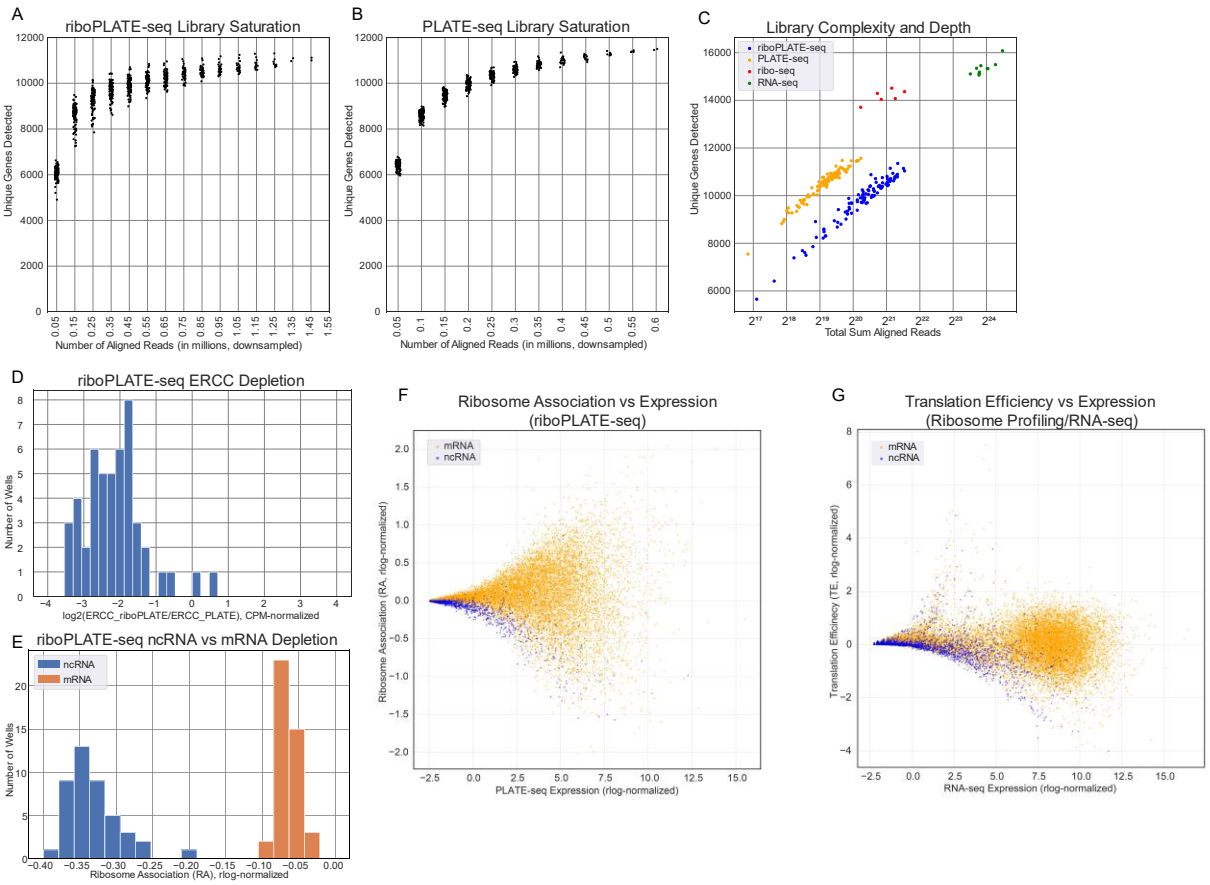
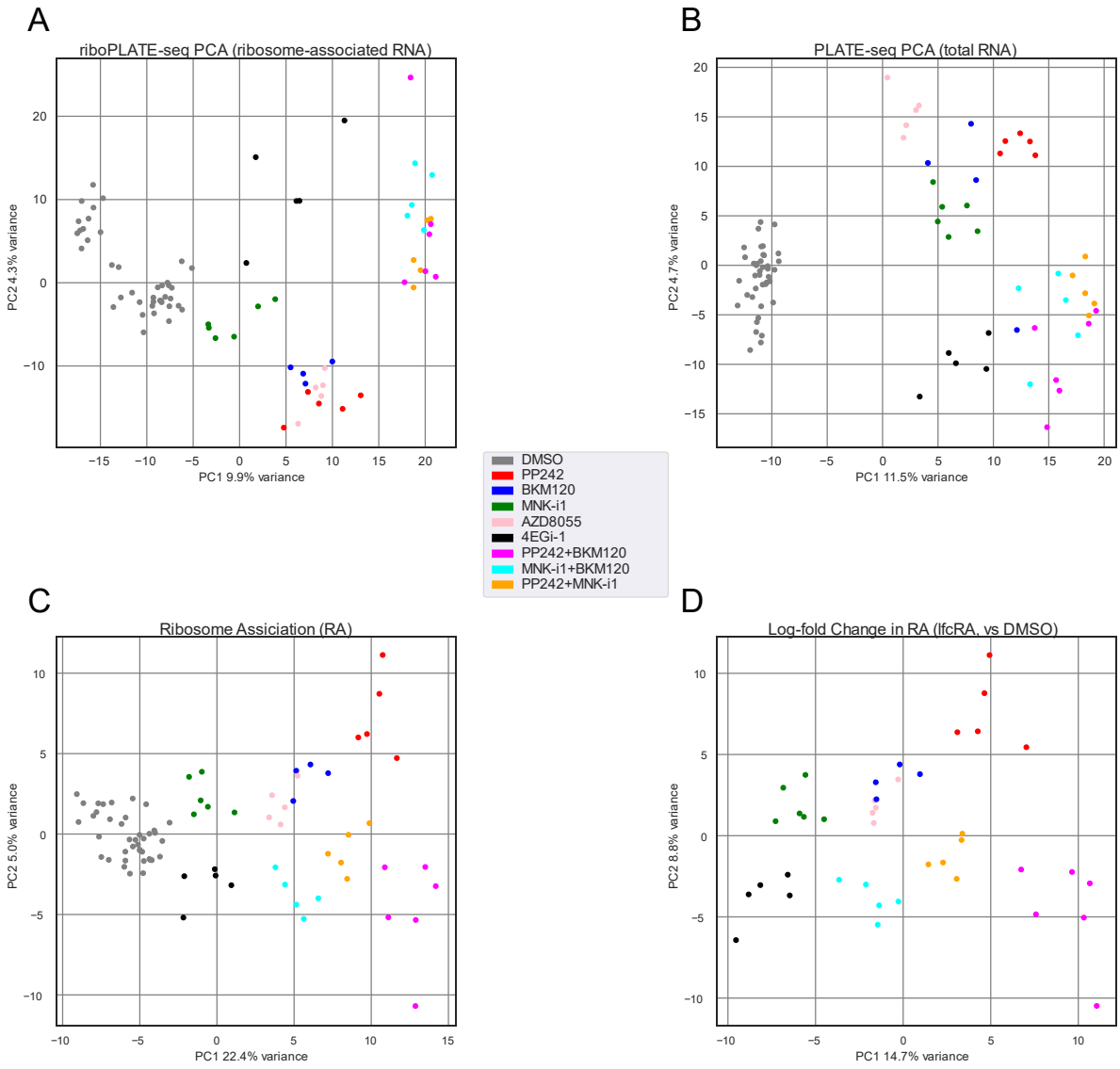


Figure 2: RNA-seq quality control in riboPLATE-seq. A-B) Library saturation strip plots for ribosome-associated (riboPLATE-seq) and total RNA (PLATE-seq) libraries in this study. In each, the Y axis shows the number of unique genes detected in each sample at each subsampled read depth on the X axis, excluding libraries smaller than the subsampling depth. With ~10-11,000 unique genes detected, riboPLATE-seq and PLATE-seq are comparably saturated. C) Scatter plots emphasizing the relationship between library size and complexity across library types. The Y axis represents the number of unique genes detected within a library; the X axis represents its size in summed gene counts. PLATE-seq and riboPLATE-seq are very similarly distributed, with PLATE-seq generating slightly more complex, smaller libraries than riboPLATE-seq, while ribo- and RNA-seq generate larger, more complex libraries. D) Depletion of RNA spike-ins due to ribosome IP for 48 spike-in-containing samples. Depletion is calculated per-sample as the $\log_2$-ratio of the sum of all spike-in-aligned counts in the riboPLATE-seq

579 library over the same sum in the sample's paired PLATE-seq library, using counts-per-million
580 normalization (CPM). Most libraries show significant depletion of spike-ins after IP (median $\log_2$ ratio -
581 2.16; Wilcoxon signed-rank test $p=1.97*10^{-9}$). E) Ribosome association (RA) of noncoding RNA (ncRNA,
582 blue) vs mRNA (orange) across sample wells in riboPLATE-seq. RA is calculated as the ratio of rlog-
583 normalized riboPLATE- to PLATE-seq counts for each gene in each sample well; distributions are
584 calculated over annotated genes with more than 100 raw counts across 96 PLATE-seq libraries. ncRNA
585 are depleted from riboPLATE- relative to PLATE-seq, with significantly lower RA than mRNA (mean RA -
586 0.09 for mRNA vs -0.34 for ncRNA; Mann-Whitney U test $p=7.28*10^{-16}$) F,G) Scatterplots of rlog-RA (F)
587 and TE (G) as functions of transcript abundance in DMSO-treated controls, determined by PLATE-seq or
588 total RNA-seq aligned counts. In both, points represent noncoding (blue) or messenger (orange) RNAs,
589 with abundance on the X axis and RA or TE on the Y axis. In riboPLATE-/PLATE-seq and to a lesser extent
590 in ribo-/RNA-seq, ncRNA exhibit lower RA and TE than comparably abundant mRNA.



591

592 **Figure 3: Two-component Principal Component Analysis (PCA) of rlog-normalized riboPLATE-seq count**

593 **data, color-coded by drug treatment.** A) ribosome-associated RNA PLATE-seq (riboPLATE-seq), B)

594 expressed RNA/conventional PLATE-seq, C) riboPLATE-seq rlog ribosome association (RA, riboPLATE-seq-

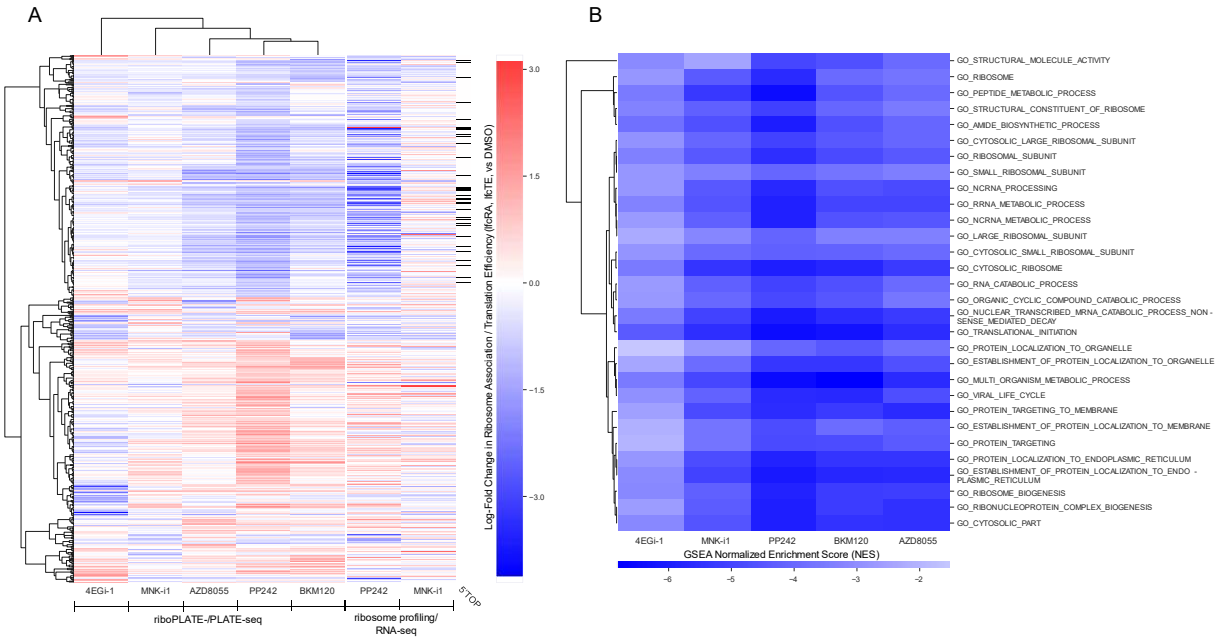
595 PLATE-seq), D) difference in rlog-RA between each sample and the average across DMSO-treated

596 samples. For RA-dependent measurements in (C,D), the domain of the PCA was restricted to the genes

597 with significant changes in RA reported by DESeq2 for any drug treatment relative to DMSO, with an

598 additional abundance threshold (FDR<0.01, baseMean>=20; 872 genes total). Drug treatments elicit

599 changes consistent enough to yield clustering behavior among samples treated with the same drug in all
600 four analyses, as well as co-clustering of related drug treatments (e.g. BKM120, PP242, and AZD8055).
601 Separation is also apparent between combination treatments and their constituent, individual drugs in
602 each plot. Additionally, in (C,D), the average PC1 coordinate value of a cluster appears correlated with
603 the number of significant perturbations in RA detected for that cluster's drug vs DMSO, though this
604 relationship does not extend across the horizontal line PC2=0 (Spearman rank correlation coefficients ρ
605 = 1.0, p=0.0 across subgroups; ρ = 0.97, p=3.3*10^{-5} for RA and ρ = 0.90/p=0.002 for lfcRA considering all
606 samples).



607

608 **Figure 4: Translational signatures of drug treatments observed with riboPLATE-seq and comparison to**
609 **ribosome profiling/RNA-seq.** A) Hierarchically-clustered heatmap of log-fold change in RA for each drug
610 treatment vs DMSO, calculated by DESeq2. The Y axis is comprised of a restricted set of genes exhibiting
611 significant differences in RA due to drug treatment (p<0.01), with greater than 20 mean normalized
612 counts per sample (baseMean>20, calculated by DESeq2 to account for sequencing depth), across the

set of drugs on the X axis. The column dendrogram demonstrates the greatest similarity in effect among the three inhibitors of the PI3K/Akt/mTOR signaling axis, with lesser correlations to MNK-i1 and 4EGi-1. At right, not included in the dendrogram, are signatures of differential translation efficiency (TE) across this same set of genes, generated via ribosome profiling and RNA sequencing of PP242-, MNK-i1-, and DMSO-treated samples. PP242 generates a strong signature with correlation across RA and TE regimes (Spearman ρ=0.76, p=5.00*10^{-90}); by contrast, the signature for MNK-i1 treatment is weak and poorly correlated across datatypes (Spearman ρ=0.089, p=0.055). In the rightmost column, the canonical terminal oligopyrimidine motif-containing genes (TOP genes) present in the Y axis gene set are identified with tick marks, showing consistent downregulation of these genes in RA and TE under treatment with PI3K/Akt/mTOR axis inhibitors. B) Heatmap of GSEA normalized enrichment scores (NES) of highly significant differentially-enriched gene ontology terms from the MSigDB C5 collection (biological process, molecular function, and cellular component). The Y axis consists of a highly conservative set of gene ontology terms identified as significantly differentially enriched in RA (GSEA FWER<0.001) due to treatment with any of the drugs on the X axis. At this level of significance, all NES are negative, indicating downregulation of their associated gene ontologies. Downregulated ontologies are also largely concerned with translational machinery, the ribosome, and protein synthesis in general, reflective of the translation-inhibitory principal effects of the drugs used.

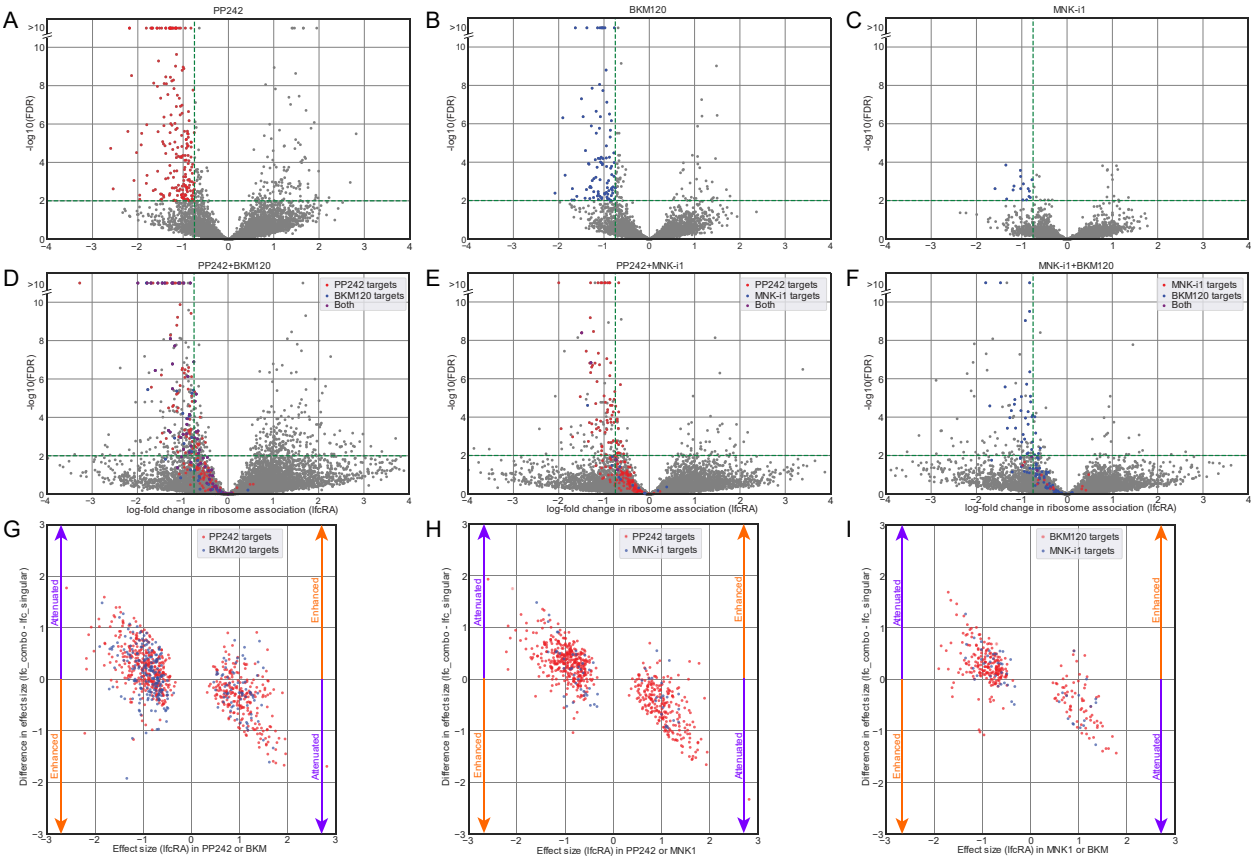
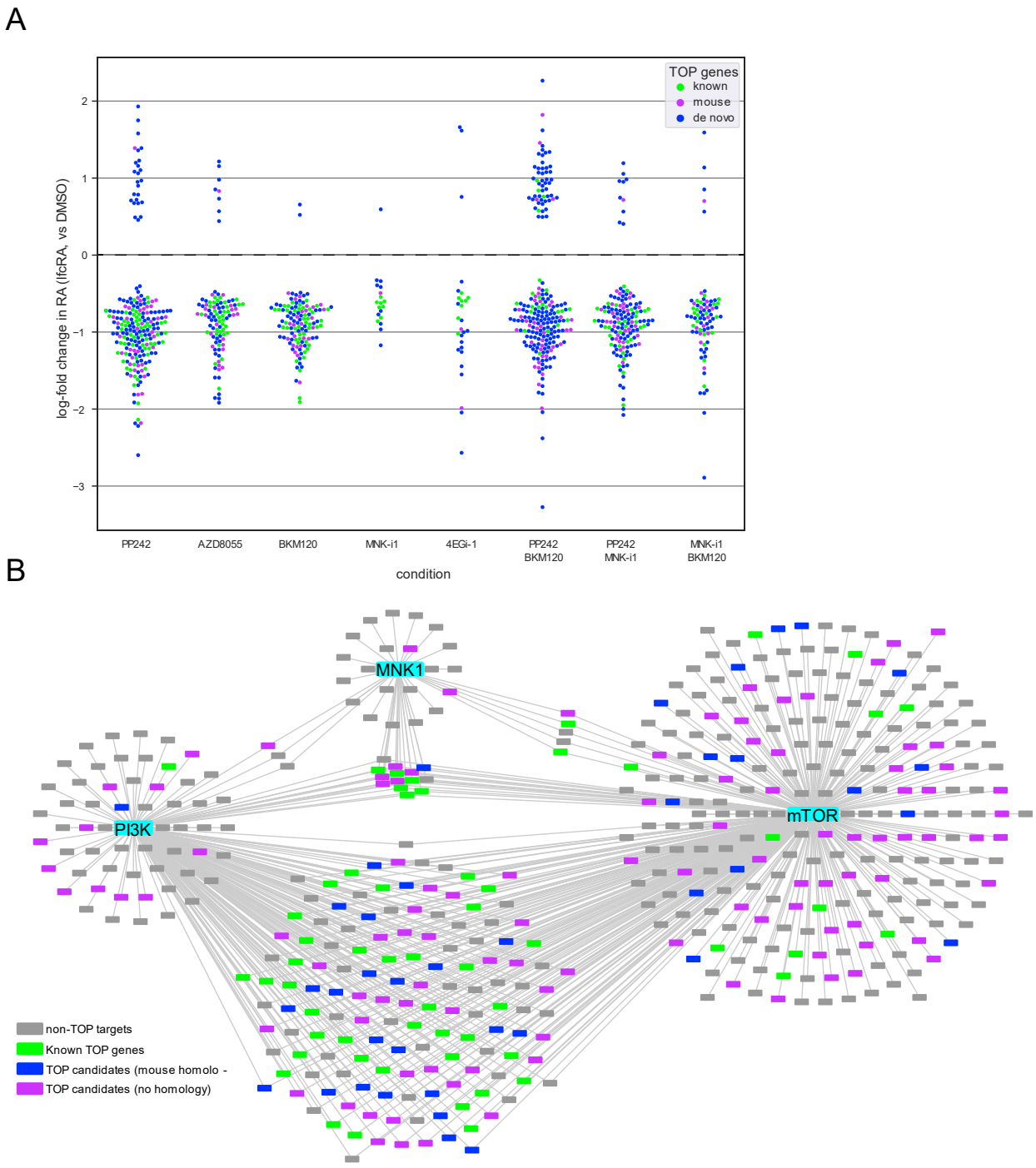


Figure 5: Effects of drug treatments in combination. A-C) plots of the per-gene changes in RA for PP242, BKM120, and MNK-i1 treatments vs DMSO, calculated by DESeq2. In each, the Y axis corresponds to significance and the X axis is effect size in terms of log-fold change in RA, with green dotted lines to represent the thresholds of p=0.01 and lfcRA<-0.75. Genes meeting both thresholds are color-coded red or blue, and represent the major inhibitory targets of these individual drugs. D-F) Volcano plots of the three pairwise combinations of the drugs in A-C, highlighting the specific targets of the individual drugs. In each, these specific targets have largely shifted down and rightward, indicating less significant and smaller perturbation in RA of these targets in combination. G-I) Scatterplots comparing the effect under single-drug treatment and the difference between combination and single-drug effects for the same target sets from (A-C). These plots show the relationship between initial effect size under single drug treatment and the degree of attenuation or amplification in this effect under combined treatment,

642 excluding targets that change sign (Additional File #1: Supplementary Table 1 for details). The majority
643 of targets fall in the first and third quadrants of each plot, indicating attenuation of most targets, though
644 a greater fraction of targets in quadrants 2 and 3 for the combination of PP242+BKM120 suggests
645 additivity in some of their effects. The plots additionally demonstrate a consistent pattern of increased
646 attenuation (e.g. increased differences between combination and singular effect size) with increasing
647 single-treatment effect size (ρ = -0.48, p = $1.9*10^{-32}$ for PP242+BKM120; ρ = -0.63, p = $9.7*10^{-15}$ for
648 PP242+MNK-i1; ρ = -0.55, p = $5.2*10^{-45}$ for BKM120+MNK-i1).



649

650 **Figure 6: Network visualization of RA perturbations and association with known regulatory motifs.** A)

651 TOP genes and candidates significantly perturbed in RA by drug treatments. Strip plots along the X axis,

652 labeled for each drug treatment in our riboPLATE-seq study, contain log-fold changes in RA (Y axis) for

653 the genes exhibiting significant RA perturbations (FDR<0.05) under each treatment relative to DMSO

controls, excluding non-TOP-containing genes. TOP candidates behave similarly to canonical TOP genes, exhibiting decreased RA under treatment with mTOR axis inhibitors (PP242, AZD8055, BKM120) while MNK-i1 and 4EGi-1 elicit fewer significant alterations in these sets. B) Network representation of targets of mTOR, PI3K, and MNK1, interpreted as the sufficiently-abundant genes exhibiting significant decreases in RA under treatment with PP242, BKM120, and MNK-i1, respectively (FDR<0.05, baseMean > 20). Targets are color-coded to identify them as canonical 5'TOP motif-containing genes (green), TOP gene candidates with known mouse homologues (blue) or no known homology (purple), or genes with no known TOP motif (gray). The mTOR-exclusive targets are enriched for TOP genes and candidates (Fisher's exact test $p=4.92*10^{-8}$), as are the common targets of mTOR and PI3K ($p=5.74*10^{-24}$) and the targets common to all three kinases (Fisher's exact test $p=5.88*10^{-5}$). In contrast, MNK1 targets very few genes, and proportionally fewer of them are TOP genes or candidates (Fisher's exact test on MNK1-exclusive targets p=0.760). Detailed analysis of TOP gene enrichment provided in Supplementary Table 2 (Additional File #1: Supplementary Table 2).

Drug Treatment	Total Samples	Remaining Samples (outliers removed)
AZD8055	6	5
BKM120	6	4
DMSO	48	40
MNK-i1	6	6
4EGi-1	6	5
PP242	6	5
MNK-i1+BKM120	6	5
PP242+MNK-i1	6	5

PP242+BKM-120	6	6

668 **Table 1: Experimental group sizes and outlier removal**. The left column describes the total number of
669 samples of each drug treatment condition on the plate, while the right column describes the number of
670 samples remaining after trimming PCA outliers.

671

Drug (riboPLATE-seq)	**lfcRA<0**	**lfcRA>0**
PP242	400	320
BKM120	216	88
AZD8055	155	61
MNK-i1	45	29
4EGi-1	107	30
PP242+BKM120	524	604
PP242+MNK-i1	313	174
MNK-i1+BKM120	244	110
(ribo-seq/RNA-seq)	**lfcTE<0**	**lfcTE>0**
PP242	1560	1481
MNK-i1	1254	1109

672 **Table 2**: **Significant perturbations (FDR<0.05) in ribosome association (RA) and translation efficiency**
673 **(TE) as a function of drug treatment, calculated with DESeq2, separated by direction of effect**. The first
674 eight rows pertain to riboPLATE-seq data and calculated changes in RA, while the last two relate to
675 ribosome profiling/RNA-sequencing and calculated changes in TE. The first column represents the

676 number of genes exhibiting significant decreases in RA/TE (log-fold change in RA/TE < 0), while the
677 second column represents increases in RA/TE (lfcRA/lfcTE > 0).

678

679 **ADDITIONAL FILES**

680 Additional File #1: Supplementary Tables 1 and 2, Supplementary Figure 1.
681 Supplementary_Information.docx (268KB)

682

683 **DECLARATIONS**

684 **Ethics approval and consent to participate**

685 Not applicable

686 **Consent for publication**

687 Not applicable

688 **Availability of data and materials**

689 Sequencing data for this project is available on the Gene Expression Omnibus (GEO), under
690 accession ID GSE139238 [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE139238]. The
691 jupyter notebook containing all of the analyses performed for this paper and the accompanying files
692 necessary for its functions are available for download from our laboratory website[29].

693 **Competing interests**

694 The authors declare that they have no competing interests.

695 **Funding**

PAS was funded by NIH/NCI R33CA202827.

Authors' Contributions

PAS and NJH initially proposed the riboPLATE-seq protocol, and NJH performed exploratory and foundational work in its development. JBM and JW developed the workflow for automated ribosome immunoprecipitation. JBM executed the riboPLATE-seq drug screen with SDS, and performed all computational and statistical analyses. SDS prepared all ribosome profiling and RNA-sequencing libraries in this study in addition to growing, treating, and lysing cells for the riboPLATE-seq experiment. JBM and PAS wrote the paper. All authors read and approved the final manuscript prior to submission.

Acknowledgements

We would like to thank Charles Karan and Ronald Realubit at the High-Throughput Screening Facility of the JP Sulzberger Columbia Genome Center for their help in executing PLATE-seq library preparations. We thank MRC Technology for sharing the MNK-I1 compound.

REFERENCES

1. Zoncu, R., Efeyan, A. & Sabatini, D. M. mTOR: from growth signal integration to cancer, diabetes and ageing. *Nat. Rev. Mol. Cell Biol.* **12**, 21–35 (2011).
2. Thoreen, C. C. *et al.* A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* **485**, 109–113 (2012).
3. Johannes, G., Carter, M. S., Eisen, M. B., Brown, P. O. & Sarnow, P. Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. *Proc. Natl. Acad. Sci.* **96**, 13118–13123 (1999).
4. Steitz, J. A. Polypeptide Chain Initiation: Nucleotide Sequences of the Three Ribosomal Binding Sites in Bacteriophage R17 RNA. *Nature* **224**, 957–964 (1969).

5. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S. & Weissman, J. S. Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science* **324**, 218–223 (2009).

6. Sanz, E. *et al.* Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. *Proc. Natl. Acad. Sci.* **106**, 13939–13944 (2009).

7. Gonzalez, C. *et al.* Ribosome Profiling Reveals a Cell-Type-Specific Translational Landscape in Brain Tumors. *J. Neurosci.* **34**, 10924–10936 (2014).

8. Hornstein, N. *et al.* Ligation-free ribosome profiling of cell type-specific translation in the brain. *Genome Biol.* **17**, 149 (2016).

9. Shigeoka, T. *et al.* Dynamic Axonal Translation in Developing and Mature Visual Circuits. *Cell* **166**, 181–192 (2016).

10. Bush, E. C. *et al.* PLATE-Seq for genome-wide regulatory network analysis of high-throughput screens. *Nat. Commun.* **8**, 105 (2017).

11. Heiman, M., Kulicke, R., Fenster, R. J., Greengard, P. & Heintz, N. Cell type-specific mRNA purification by translating ribosome affinity purification (TRAP). *Nat. Protoc.* **9**, 1282–1291 (2014).

12. Szerlip, N. J. *et al.* Intratumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor response. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 3041–3046 (2012).

13. Waskiewicz, A. J. Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2. *EMBO J.* **16**, 1909–1920 (1997).

14. Feldman, M. E. *et al.* Active-Site Inhibitors of mTOR Target Rapamycin-Resistant Outputs of mTORC1 and mTORC2. *PLoS Biol.* **7**, (2009).

15. Pike, K. G. *et al.* Optimization of potent and selective dual mTORC1 and mTORC2 inhibitors: The discovery of AZD8055 and AZD2014. *Bioorg. Med. Chem. Lett.* **23**, 1212–1216 (2013).

16. Koul, D. *et al.* Antitumor Activity of NVP-BKM120--A Selective Pan Class I PI3 Kinase Inhibitor Showed Differential Forms of Cell Death Based on p53 Status of Glioma Cells. *Clin. Cancer Res.* **18**, 184–195 (2012).
17. Genheden, M. *et al.* BDNF stimulation of protein synthesis in cortical neurons requires the MAP kinase-interacting kinase MNK1. *J. Neurosci. Off. J. Soc. Neurosci.* **35**, 972–984 (2015).
18. Wang, W. *et al.* 4EGI-1 induces apoptosis and enhances radiotherapy sensitivity in nasopharyngeal carcinoma cells via DR5 induction on 4E-BP1 dephosphorylation. *Oncotarget* **7**, 21728–21741 (2016).
19. Grzmil, M. *et al.* MAP Kinase-Interacting Kinase 1 Regulates SMAD2-Dependent TGF- Signaling Pathway in Human Glioblastoma. *Cancer Res.* **71**, 2392–2402 (2011).
20. Shi, Y. *et al.* MNK1-Induced eIF-4E Phosphorylation in Myeloma Cells: A Pathway Mediating IL-6-Induced Expansion and Expression of Genes Involved in Metabolic and Proteotoxic Responses. *PLoS ONE* **9**, e94011 (2014).
21. Bain, J. *et al.* The selectivity of protein kinase inhibitors: a further update. *Biochem. J.* **408**, 297–315 (2007).
22. Beggs, J. E. *et al.* The MAP kinase-interacting kinases regulate cell migration, vimentin expression and eIF4E/CYFIP1 binding. *Biochem. J.* **467**, 63–76 (2015).
23. Algire, M. A., Maag, D. & Lorsch, J. R. Pi Release from eIF2, Not GTP Hydrolysis, Is the Step Controlled by Start-Site Selection during Eukaryotic Translation Initiation. *Mol. Cell* **20**, 251–262 (2005).
24. Guttman, M., Russell, P., Ingolia, N. T., Weissman, J. S. & Lander, E. S. Ribosome Profiling Provides Evidence that Large Noncoding RNAs Do Not Encode Proteins. *Cell* **154**, 240–251 (2013).
25. Yamashita, R. *et al.* Comprehensive detection of human terminal oligo-pyrimidine (TOP) genes and analysis of their characteristics. *Nucleic Acids Res.* **36**, 3707–3715 (2008).

765 26. Woolstenhulme, C. J., Guydosh, N. R., Green, R. & Buskirk, A. R. High-Precision Analysis of
766 Translational Pausing by Ribosome Profiling in Bacteria Lacking EFP. *Cell Rep.* **11**, 13–21 (2015).
767 27. Arava, Y. *et al.* Genome-wide analysis of mRNA translation profiles in Saccharomyces cerevisiae.
768 *Proc. Natl. Acad. Sci.* **100**, 3889–3894 (2003).
769 28. Fan, S., Li, Y., Yue, P., Khuri, F. R. & Sun, S.-Y. The eIF4E/eIF4G Interaction Inhibitor 4EGI-1 Augments
770 TRAIL-Mediated Apoptosis through c-FLIP Down-regulation and DR5 Induction Independent of
771 Inhibition of Cap-Dependent Protein Translation. *Neoplasia* **12**, 346-IN7 (2010).
772 29. Sims Laboratory at Columbia University.
773 http://www.columbia.edu/~pas2182/index.php/technology.html.
774 30. Gordon, A. & Hannon, G. FASTX-Toolkit. http://hannonlab.cshl.edu/fastx_toolkit/index.html.
775 31. Langmead, B. Aligning short sequencing reads with Bowtie. *Curr. Protoc. Bioinforma.* **32**, 11.7. 1-
776 11.7. 14 (2010).
777 32. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning
778 sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2013).
779 33. Yang, L., Duff, M. O., Graveley, B. R., Carmichael, G. G. & Chen, L.-L. Genomewide characterization of
780 non-polyadenylated RNAs. *Genome Biol.* **12**, R16 (2011).
781 34. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq
782 data with DESeq2. *Genome Biol.* **15**, 550 (2014).
783 35. Analyzing RNA-seq data with DESeq2.
784 http://bioconductor.org/packages/devel/bioc/vignettes/DESeq2/inst/doc/DESeq2.html.
785 36. Subramanian, A., Kuehn, H., Gould, J., Tamayo, P. & Mesirov, J. P. GSEA-P: a desktop application for
786 Gene Set Enrichment Analysis. *Bioinformatics* **23**, 3251–3253 (2007).
787 37. Shannon, P. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction
788 Networks. *Genome Res.* **13**, 2498–2504 (2003).

789 38. Wiesel, R., Eiglsperger, M. & Kaufmann, M. yFiles - Visualization and Automatic Layout of Graphs.
790 19.
791 39. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90 (2007).
792 40. Thomas, K. *et al.* Jupyter Notebooks -- a publishing format for reproducible computational
793 workflows. *Stand Alone* 87–90 (2016) doi:10.3233/978-1-61499-649-1-87.
794 41. van der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy Array: A Structure for Efficient Numerical
795 Computation. *Comput. Sci. Eng.* **13**, 22–30 (2011).
796 42. Jones, E., Oliphant, T., Peterson, P. & others. SciPy.org — SciPy.org. *SciPy: Open Source Scientific*
797 *Tools for Python* https://www.scipy.org/ (2001).
798 43. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Mach. Learn. PYTHON* 6.
799 44. Michael Waskom *et al. mwaskom/seaborn: v0.9.0 (July 2018)*. (Zenodo, 2018).
800 doi:10.5281/zenodo.1313201.
801