The human *RBM10* gene dually encodes a repressor of ribosome biogenesis that downregulates cell proliferation

Xiongwen Cao[1,2], Cecelia M. Harold[3], Carson J. Bryant[4], Shu-Jian Zheng[1,2], Alexandra Khitun[1,2], Susan J. Baserga[3,4,5], Sarah A. Slavoff[1,2,4]

[1]Department of Chemistry, Yale University, New Haven, Connecticut 06520, United States

[2]Institute of Biomolecular Design and Discovery, Yale University, West Haven, Connecticut 06516, United States

[3]Department of Genetics, Yale University School of Medicine, New Haven, Connecticut 06520, United States

[4]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520, United States

[5]Department of Therapeutic Radiology, Yale University School of Medicine, New Haven, Connecticut 06520, United States

*Correspondence: sarah.slavoff@yale.edu

Abstract

Ribosome biogenesis in eukaryotes is a highly regulated, essential cellular process. However, few repressors of ribosome biogenesis have been identified. Here, we identify and define the function of MINAS-60 (*MI*croprotein that *N*egatively regulates *AS*sembly of the pre-*60*S ribosomal subunit), a 130-amino acid protein co-encoded with the human pre-mRNA splicing regulator and tumor suppressor protein RBM10. MINAS-60 localizes to the nucleolus, where it associates with multiple pre-60S assembly factors. Depletion of MINAS-60 increases the amount of mature 60S ribosomal subunit in the cytoplasm and, consequently, upregulates global protein synthesis and cell proliferation.

Mechanistically, we show that MINAS-60 represses late-stage pre-60S assembly and export of the mature 60S ribosome subunit to the cytoplasm. Together, these results implicate MINAS-60 as a repressor of ribosome biogenesis that acts as a checkpoint for the maturation of the large subunit (LSU). More broadly, the *RBM10* transcript encodes two sequence-independent proteins that inhibit cell proliferation via different molecular mechanisms, expanding existing models of multicistronic human gene functions.

Introduction

Expression of thousands of previously unannotated small open reading frames (smORFs, typically defined as ORFs comprising fewer than 100 codons and, in some studies, up to 150 codons[1]) has recently been revealed in mammalian cells[2]. These smORFs are found in long non-coding RNAs, 5' and 3' untranslated regions (UTRs) of mRNAs, and frame-shifted ORFs overlapping protein coding sequences (CDS), the latter of which are termed alternative ORFs or “alt-ORFs”[3]. A rapidly increasing number of smORF-encoded proteins (SEPs, also known as micropeptides and microproteins) and alt-ORF-encoded proteins (alt-proteins)[2] have been shown to play important roles in vertebrate biology. For example, the human SEPs BRAWNIN and MOCCI function in oxidative phosphorylation[4,5], and MP31 is a tumor suppressor that regulates lactate metabolism in glioblastoma[6]. Fewer alt-proteins have been characterized, including alt-FUS, which cooperates with FUS in formation of cytotoxic aggregates[7], and alt-RPL36, which we have shown regulates the PI3K-AKT-mTOR pathway[8]. Recently, genome-scale CRISPR screens revealed that hundreds of smORFs regulate human cell growth and survival[9,10]. These studies demonstrate that defining the function of bioactive SEPs and alt-proteins represents a major opportunity to gain insights into biological complexity.

Currently, there is proteomic and ribosome profiling evidence for tens of thousands of uncharacterized human smORFs and alt-ORFs[3,11]. However, the vast majority of alt-ORFs remain completely unstudied because it is challenging

to separate their functions from the canonical protein CDS in which they are nested. It thus remains unclear whether alt-ORFs are broadly biologically important. In this work, we expand our understanding of this class of genes by discovering an alt-ORF that overlaps the human RBM10 CDS and encodes a repressor of ribosome large subunit (LSU) biogenesis.

Ribosome biogenesis is a highly spatially and temporally regulated cellular process essential for growth and development[12-15]. In humans, ribosome biogenesis starts in the nucleolus with the transcription of a 47S precursor rRNA (pre-rRNA) by RNA polymerase I (RNAPI). The 47S pre-rRNA is chemically modified and joined with ribosome assembly factors and ribosomal proteins to form the 90S pre-ribosomal particle. Endonucleolytic cleavage of the 47S pre-rRNA subsequently generates the pre-40S and pre-60S particles, which undergo individual maturation and quality-control steps prior to export to the cytoplasm. Pre-60S ribosomal particles have recently been probed by cryo-electron microscopy (cryo-EM)[16,17], revealing a sequential maturation process involving quality control checkpoints for 5S ribonucleoprotein particle incorporation and rotation[18], active site formation, and removal of internal transcribed spacer 2 (ITS2) prior to export from the nucleus to the cytoplasm. In the cytoplasm, the final steps of maturation occur to produce large 60S and small 40S subunits, which associate to form translation-competent ribosomes. Dysregulated ribosome biogenesis has been linked to numerous human disorders, including cancer[19], Alzheimer's disease[20], and congenital disorders termed ribosomopathies[21,22].

To guarantee faithful protein translation, the cell has to safeguard ribosome integrity. For example, the Rio1-Nob1-Pno1 network establishes a checkpoint to safeguard against the release of immature 40S subunits into translating ribosomes[23]. Likewise, a cryo-EM structure study of pre-export human pre-60S complexes revealed an unidentified protein, "protein X", that prevents the incorporation of the pre-60S nuclear export factor NMD3 into the late pre-60S

ribosomal subunit, hypothetically inhibiting LSU export to the cytoplasm[24]. Nevertheless, due to a lack of information about protein X, it is currently unclear whether it functions as a checkpoint inhibitor of pre-60S nucleocytoplasmic export.

In this study, we provide molecular and proteomic evidence for the function of an alt-protein encoded by an alt-ORF nested in the human RBM10 CDS. We name this alt-protein MINAS-60, or *MI*croprotein that *N*egatively regulates *AS*sembly of the pre-*60*S ribosomal subunit. We show that MINAS-60 localizes to the nucleolus, associates with LSU assembly factors GTPBP4 and MRTO4, and co-fractionates with pre-60S complexes in nuclear extracts. Finally, we engineer MINAS-60 knockdown and rescue cell lines to demonstrate that loss of MINAS-60 increases cytoplasmic 60S ribosome subunit levels, global protein synthesis, and cell proliferation. This is independent of the function of the canonical RBM10 protein, which has been previously shown to inhibit cell proliferation through its role in alternative pre-mRNA splicing[25-29]. These results implicate MINAS-60 as a rare negative regulator of late nuclear steps in LSU biogenesis, and demonstrate that the human *RBM10* gene encodes two proteins that downregulate cell proliferation via distinct molecular mechanisms.

Results

RBM10 alternatively encodes an endogenously expressed, unannotated protein localized to the nucleolus

Motivated by the hypothesis that alt-proteins undergoing active synthesis could play important cellular roles, we set out to develop a chemoproteomic approach to identify newly translated alt-proteins. We leveraged bio-orthogonal non-canonical amino acid tagging (BONCAT), an *in vivo* labeling strategy to identify nascent proteins[30]. BONCAT utilizes the methionine analogue azidohomoalanine (AHA), which can be metabolically incorporated into all newly synthesized proteins by the endogenous protein translation machinery. Labeled proteins can then be derivatized with click chemistry to append a biotin tag for streptavidin

enrichment. We performed BONCAT enrichment on the whole proteome, then followed with our previously reported proteogenomic strategy for small proteome enrichment and liquid chromatography/tandem mass spectrometry-based identification of unannotated microproteins and alt-proteins[31]. Using this new strategy, we identified a tryptic peptide that mapped uniquely to an alternative open reading frame (alt-ORF) of human *RBM10* transcript variant 1 (*NM_005676.5*) in HEK 293T cells, which encodes an alternative protein we name MINAS-60 (Figure 1a, b, and Supplementary Data 1). A different peptide-spectral match mapping to the MINAS-60 alt-ORF was previously detected in a re-analysis of published mass spectral data from colorectal cancer samples, further supporting expression of MINAS-60 in human cell lines as well as patient-derived tissues, but the alt-protein was not further defined in that study[32].

To confirm translation and examine subcellular localization of MINAS-60, the cDNA sequence comprising the 5'UTR of *RBM10* transcript variant 1 through the stop codon of the putative alt-ORF was cloned into a mammalian expression vector with an HA tag appended to the C-terminus of MINAS-60. This construct produced two anti-HA immunoreactive bands (~15 kDa and ~20 kDa apparent molecular weight, due to multiple in-frame start codons, vide infra) from the MINAS-60 reading frame when transiently transfected into HEK 293T cells (Figure 1c). The translation of MINAS-60 was confirmed by immunostaining, and we observed that over-expressed MINAS-60 co-localized with a nucleolar protein, fibrillarin (Figure 1d).

To validate expression of endogenous MINAS-60 from the *RBM10* genomic locus, we generated two independent Cas9-directed knock-in (KI) HEK 293T cell lines with a 3×GFP11-FLAG-HA tag appended to the 3' end of the MINAS-60 alt-ORF[33]. Western blotting revealed that both KI cell lines exhibited four anti-HA immunoreactive bands, which were absent from control cells, indicating the endogenous expression of multiple isoforms of MINAS-60 that contain the same C-terminus (Figure 1e). To further confirm the endogenous expression of MINAS-

60, we performed immunostaining with the KI cells. As shown in Figure 1f, endogenously expressed MINAS-60 co-localizes with fibrillarin, consistent with the over-expression results. Taken together, these results indicate that MINAS-60 is endogenously expressed and localizes to the nucleolus.

MINAS-60 is initiated from multiple start codons

Because bottom-up proteomics did not directly identify the MINAS-60 N-terminus, we experimentally identified the start codons that initiate MINAS-60 translation. Alt-ORFs have been reported to initiate at upstream, non-AUG start codons[8,34,35] and internal, AUG start codons[7,36], so we tested both possibilities. We searched the upstream *RBM10* transcript sequence in-frame with the MINAS-60 tryptic peptide, and found two upstream near-cognate start codons: A_{383}TC and A_{386}GG (numbered relative to the first nucleotide of the cDNA), as well as seven internal AUG start codons (numbered ATG1 – ATG7) (Supplementary Figure 1a). Deletion of the two upstream near-cognate start codons did not prevent expression of MINAS-60 (Supplementary Figure 1b, compare lane 5 with lane 1). Mutation of ATG1 to TAA abolished the expression of the ~20 kDa band, and led to aberrant initiation at an internal site that is not utilized in the wild-type construct (Supplementary Figure 1c, compare lane 2 with lane 1), suggesting that the ~20 kDa MINAS-60 initiates at ATG1. Mutation of ATG5 or ATG6 to TAA reduced, while mutation of ATG7 to TAA abolished, the expression of the ~15 kDa band (Supplementary Figure 1c, compare lanes 6, 7, 8 with lane 1), suggesting that the ~15 kDa MINAS-60 could initiate at ATG7; ATG5/6 may initiate translation of short isoforms that are not resolved by SDS-PAGE, or may play regulatory roles.

To further confirm these results, we cloned the MINAS-60 coding sequence starting from ATG1 or ATG6 into a mammalian expression vector. The ATG1 construct produced two bands co-migrating with the wild-type products, and the ATG6 construct produced a band co-migrating with the short isoform (Supplementary Figure 1d). These results suggest that translation of MINAS-60

from transcript variant 1 of *RBM10* likely initiates at ATG1 and ATG7, and based on rescue experiments below, this construct recapitulates MINAS-60 function after *RBM10* knockdown (vide infra). We speculate that the additional MINAS-60 isoform diversity observed in KI cells could be generated by alternative splicing, analogous to PTBP3[36].

MINAS-60 is associated with nucleolar LSU assembly factors

Because many SEPs characterized to date bind to and regulate other proteins[37], we performed a co-immunoprecipitation (co-IP) with the nuclear lysates from MINAS-60-FLAG KI cells, and HEK 293T cells as a control. Two major bands were specifically present in the KI co-IP after SDS-PAGE that, upon analysis via label-free quantitative proteomics[38], yielded 17 proteins enriched >30-fold over control HEK 293T cells. GO analysis of these 17 hits shows that the top 2 enriched biological processes are ribosome biogenesis and ribosomal large subunit biogenesis (Supplementary Figures 2a-c, and Supplementary Data 2). To obtain a global picture of MINAS-60-associated proteins, we subsequently performed co-IP with the MINAS-60-FLAG KI cells and wild-type controls and analyzed the entire molecular weight range with quantitative proteomics; while ribosomal protein detections in this experiment limited sensitivity toward other proteins in the absence of fractionation, we observed clear enrichment of four LSU assembly factors: GTPBP4, MRTO4, BRIX1 and NOP2 (Figure 2a and Supplementary Data 3). Co-IP followed by western blotting confirmed the association of these four assembly factors with MINAS-60 (Figure 2b). The association of MINAS-60 with GTPBP4 and MRTO4 did not depend on RNA, because these associations largely survived RNaseA treatment. In contrast, co-purification of MINAS-60 with BRIX1 and NOP2 was severely diminished after treatment with RNaseA, suggesting that their association with MINAS-60 was likely indirect and RNA-dependent (Figure 2b). We therefore hypothesized that MINAS-60 could associate with late pre-60S particles containing GTPBP4 and MRTO4 in the nucleolus to regulate ribosome biogenesis[16,24].

Because GTPBP4 and MRTO4 are associated with pre-60S ribosomal particles[17], we examined whether MINAS-60 co-fractionates with the same high molecular weight species. We performed sucrose gradient fractionation of nuclear extracts of HEK 293T cells stably expressing MINAS-60-FLAG-HA, followed by western blotting analysis. As shown in Figure 2c, a fraction of over-expressed MINAS-60 co-sedimented with ribosome assembly factors in high-molecular weight fractions coincident with pre-ribosomal particle peaks, though these species were not sufficiently resolved to enable identification of the specific intermediate(s) with which MINAS-60 associated. Combined with our co-IP and immunofluorescence data, these results are consistent with a role for MINAS-60 in nucleolar pre-ribosomal particles.

MINAS-60 downregulates global protein synthesis and cell proliferation

We hypothesized that MINAS-60 could regulate ribosome biogenesis via its association with GTPBP4 and MRTO4. Ribosome biogenesis is required for global protein synthesis, which promotes cell growth and proliferation. As a result, ribosome biogenesis is commonly upregulated in rapidly dividing cancer cells[19,39]. We therefore reasoned that, if MINAS-60 regulates ribosome biogenesis, its absence should result in changes to protein synthesis and cell proliferation. In order to test this hypothesis, we required a system to query the function of MINAS-60 independent of RBM10, despite their co-encoding on the same transcript. To this end, we knocked down (KD) *RBM10* in HEK 293T cells with two different shRNAs, which silences the entire mRNA and, therefore, both proteins (RBM10 and MINAS-60). To deconvolute phenotypic effects specific to MINAS-60 in the KD, as well as to exclude off-target effects of shRNA, we generated rescue cell lines stably expressing MINAS-60 (Rescue_MINAS-60) or RBM10 (Rescue_RBM10) on the KD background. qRT-PCR and western blotting analysis revealed that *RBM10* mRNA is efficiently silenced by both shRNAs, and the KD cells successfully re-express MINAS-60 or RBM10 after rescue (Supplementary Figure 3). Noting that the RBM10 rescue construct could hypothetically be subject to leaky translation to produce both RBM10 and

MINAS-60, we also rescued the KD with an RBM10 construct bearing an $A_{398}TG$ to TAA mutation that eliminates the MINAS-60 start codon while preserving RBM10 translation (Supplementary Figure 3).

To test the effect of MINAS-60 expression on global protein synthesis, we labeled nascent peptides in the control, RBM10 KD, Rescue_MINAS-60, Rescue_RBM10 and Rescue_RBM10(A_{398}TG-TAA) cell lines with puromycin followed by anti-puromycin western blotting[40]. As shown in Figure 3a-b, depletion of the entire *RBM10* mRNA led to a significant increase in global protein synthesis, and this increase was rescued by reintroduction of MINAS-60. Partial rescue by RBM10 reintroduction was observed, which was not present in the RBM10(A_{398}TG-TAA) rescue cells that cannot express MINAS-60. Similar results were observed for a second shRNA targeting RBM10 (Figure 3c-d). Taken together, these results indicate that MINAS-60 downregulates global protein synthesis, and that RBM10 is unlikely to play a role in this process. We speculate that the partial rescue observed with RBM10 could have been due to leaky expression of MINAS-60 from the internal alt-ORF in the RBM10 construct.

Because protein synthesis, cell growth and proliferation are linked[14,19], we asked whether MINAS-60 regulates cell proliferation. As shown in Figure 4a-b, *RBM10* depletion led to a significant increase in cell proliferation, consistent with a published report using HeLa cells[29]. Remarkably, this increase was rescued by reintroduction of MINAS-60 alone, and partially rescued by RBM10 reintroduction. Similar results were observed for a second shRNA targeting RBM10 (Figure 4c-d). These results are consistent with a model in which MINAS-60 inhibits ribosome biogenesis, subsequently downregulating cellular protein synthesis and cell proliferation.

MINAS-60 expression increases in early S phase

It has been reported that transcription of rDNA, the first step of ribosome biogenesis, oscillates during the cell cycle, nearly ceasing during M phase,

increasing during G1 phase, maximizing during S and G2 phases in human cells[41-44]. If MINAS-60 regulates ribosome biogenesis, we hypothesized that the protein level and/or the nucleolar localization of MINAS-60 would change during the cell cycle in a similar manner. Immunostaining of synchronized MINAS-60 KI HEK 293T cells revealed that MINAS-60 increased at early S phase (double thymidine block release time points 0 h and 2 h) and co-localized with nucleolar protein fibrillarin, peaking at the 2 h time point. MINAS-60 expression then decreased by late S phase (release time points 4 h and 6 h), and was very low during G2/M phase (release time points 8 h and 9 h). At G1 phase (release time points 14 h and 16 h), the nucleolar MINAS-60 staining intensity again increased (Figure 5a). These results were confirmed via western blotting (Figure 5b). MINAS-60 expression is therefore coordinated with ribosome biogenesis activity during the cell cycle.

MINAS-60 inhibits the late-stage maturation and export of 60S ribosome subunit

Eukaryotic ribosome biogenesis can be divided into sequential processes[45]: pre-rRNA transcription; chemical modification and processing of the pre-rRNA, both of which occur in the nucleolus; folding, assembly and maturation of the pre-ribosomal subunits in the nucleolus and nucleus; and export and final maturation of ribosome subunits in the cytoplasm. Because MINAS-60 co-purified with proteins involved in various steps of LSU biogenesis, we wished to determine the step in this process that it regulates.

To determine whether MINAS-60 controls pre-rRNA transcription, we performed qRT-PCR targeting the primary pre-rRNA (47S/45S/30S) using previously published primers[40] in lysates from control HEK 293T, RBM10 KD, Rescue_MINAS-60 and Rescue_RBM10 cell lines. No significant differences in primary pre-rRNA were observed, suggesting that MINAS-60 does not regulate RNA polymerase I transcription (Supplementary Figure 4). To determine whether MINAS-60 regulates the processing of LSU pre-rRNA, we performed northern

blotting analysis with the control, RBM10 KD, Rescue_MINAS-60 and Rescue_RBM10 cell lines using a probe complementary to ITS2, which detects all LSU pre-rRNA processing products, including 41S, 32S and 12S pre-rRNAs[40]. As shown in Supplementary Figure 5, no significant differences were observed between these cell lines, suggesting that MINAS-60 does not regulate the processing of LSU pre-rRNA.

We therefore examined the role of MINAS-60 in LSU assembly and export. To determine whether MINAS-60 regulates nucleocytoplasmic export of the 60S ribosomal subunit, we quantified the ratio of cytoplasmic vs. nuclear RPL29-GFP[46] stably expressed in the control, RBM10 KD, Rescue_MINAS-60 and Rescue_RBM10 cell lines. As shown in Figures 6a-c, depletion of RBM10 increased the ratio of cytoplasmic to nuclear RPL29-GFP (Figure 6b, compare column 2 with column 1), which can be rescued by reintroduction of MINAS-60 (Figure 6b, compare column 3 with column 1), and partially rescued by RBM10 reintroduction, likely due to leaky MINAS-60 translation (Figure 6b, compare column 4 with column 1). Similar results were observed for a second shRNA targeting RBM10 (Supplementary Figure 6a-c). These observations are consistent with an increase in nucleocytoplasmic export of pre-60S subunits in the absence of MINAS-60. As a control, to determine whether MINAS-60 regulates nucleocytoplasmic export of pre-40S ribosomal subunit, we quantified the ratio of cytoplasmic vs. nuclear RPS2-GFP[46] stably expressed in the control, RBM10 KD, Rescue_MINAS-60 and Rescue_RBM10 cell lines. As shown in Supplementary Figure 7, no significant changes were observed between these cell lines, suggesting MINAS-60 does not regulate the maturation or export of 40S ribosomal subunits. We therefore concluded that MINAS-60 negatively regulates either the assembly or export of pre-60S ribosomal subunits.

The observation of upregulated 60S subunit export in MINAS-60 knockdown predicted that the absence of MINAS-60 should lead to an increase in mature cytoplasmic 60S ribosomal subunit. To test this, we performed cytoplasmic

polysome profiling. As shown in Figure 6d-e, knockdown of RBM10 increased the ratio of mature 60S/40S ribosome subunits, and this increase can be rescued by reintroduction of MINAS-60 (Figure 6e, compare column 3 with column 1), and partially rescued by RBM10 (Figure 6e, compare column 4 with column 1). Similar results were observed for a second shRNA targeting RBM10 (Supplementary Figure 8). Taken together, these results suggest that MINAS-60 decreases mature large ribosomal subunits by regulating the assembly or export of pre-60S particles.

Lastly, we speculated that MINAS-60 functions as a checkpoint inhibitor in pre-60S assembly prior to export from the nucleus, which would suggest that the protein composition of LSU precursors should change in cells lacking MINAS-60. We stably expressed MRTO4-FLAG or BRIX1-FLAG, LSU biogenesis factors present in multiple intermediate and late, or early and intermediate, pre-60S particles, respectively (Figure 7a)[16], in control or RBM10 KD HEK 293T cells, to enable affinity purification of pre-ribosomal particles, followed by quantitative proteomics and western blotting to examine the pre-LSU protein composition change between these two cell lines. Quantitative proteomics revealed clear and statistically significant increases in several late LSU assembly factors co-purified with MRTO4 in RBM10 KD cells (Figure 7b, Supplementary Data 4). These increases were further confirmed by western blotting (Figure 7c). Similar results were observed for BRIX1 co-IP, though to a lesser extent (Supplementary Figure 9). These results suggest that remodeling of pre-60S particles toward more mature stages occurs in the absence of MINAS-60, consistent with MINAS-60 acting as an inhibitor for LSU maturation and/or export.

Discussion

In this work, we have identified a previously unannotated alternative protein, MINAS-60, that is co-encoded with human RBM10 in the +1 reading frame of *RBM10* transcript variant 1. MINAS-60 localizes to the nucleolus and associates with several pre-60S assembly factors, including GTPBP4 and MRTO4, in high

molecular-weight complexes, consistent with a role in the LSU assembly pathway. Loss of MINAS-60 promotes the export of pre-60S particles into cytoplasm, leading to increases in cytoplasmic 60S subunits, global protein synthesis, and cell proliferation. To our knowledge, MINAS-60 is the first protein reported to inhibit LSU biogenesis, and we suspect that it functions as a checkpoint to ensure correct pre-60S assembly prior to nuclear export. Furthermore, MINAS-60 is only the second human microprotein reported to localize to the nucleolus to date. A previous study identified nucleolar microprotein C11orf98, which interacts with nucleolar proteins NPM1 and NCL, but no role for this factor in ribosome biogenesis has been investigated[47].

In order to achieve a balance between cellular growth requirements and energy-intensive ribosome production[13], and to ensure that improperly assembled ribosome subunits do not lead to mistranslation[23], ribosome biogenesis needs to be precisely monitored in cells. However, few repressors of ribosome biogenesis have been reported, and the ones that have been identified predominately act on pre-ribosomal RNA transcription[48-50]. Here, we report MINAS-60 as a repressor of LSU biogenesis. Recently, a structural study revealed that an unidentified protein (protein X) blocks the incorporation of nuclear export factor NMD3 into the late pre-60S particle, hypothetically suppressing pre-60S export into the cytoplasm[24]. Protein X binds in the immature peptidyl transfer center and directly contacts GTPBP4 as well as helix 89 of the rRNA. Based on the observation that 1) both MINAS-60 and protein X likely function as repressors of LSU biogenesis at the late assembly and export stage; and 2) both MINAS-60 and protein X associate with GTPBP4, we speculate that MINAS-60 functions by a similar mechanism to protein X. It is also possible that protein X and MINAS-60 are the same protein, and that protein X could not be identified due to its absence from the UniProt database.

In contrast to smORF-encoded proteins, few alt-ORFs nested within canonical protein CDS have been characterized in molecular detail. This work, combined

with previous literature[7,51], expands the recent finding that a single human transcript can encode overlapping, sequence-independent, yet functionally related proteins. Importantly, the *RBM10* gene plays important roles in human physiology and disease. RBM10 (RNA Binding Motif 10) is an RNA binding protein that regulates alternative pre-mRNA splicing[29,52]. Null mutations in the *RBM10* gene are found in patients with TARP syndrome[53], an X-linked inherited pathology associated with malformation of multiple organs and significant early-life mortality. The *RBM10* gene was also found to be among the most frequently mutated genes in lung adenocarcinoma samples, and RBM10 has been shown to inhibit cancer cell proliferation[54]. The finding that *RBM10* dually encodes MINAS-60, which also downregulates cell proliferation via repression of ribosome biogenesis, opens the question of whether this alt-protein contributes to *RBM10* mutation associated disease phenotypes, like one at another recently described multicistronic human loci[3].

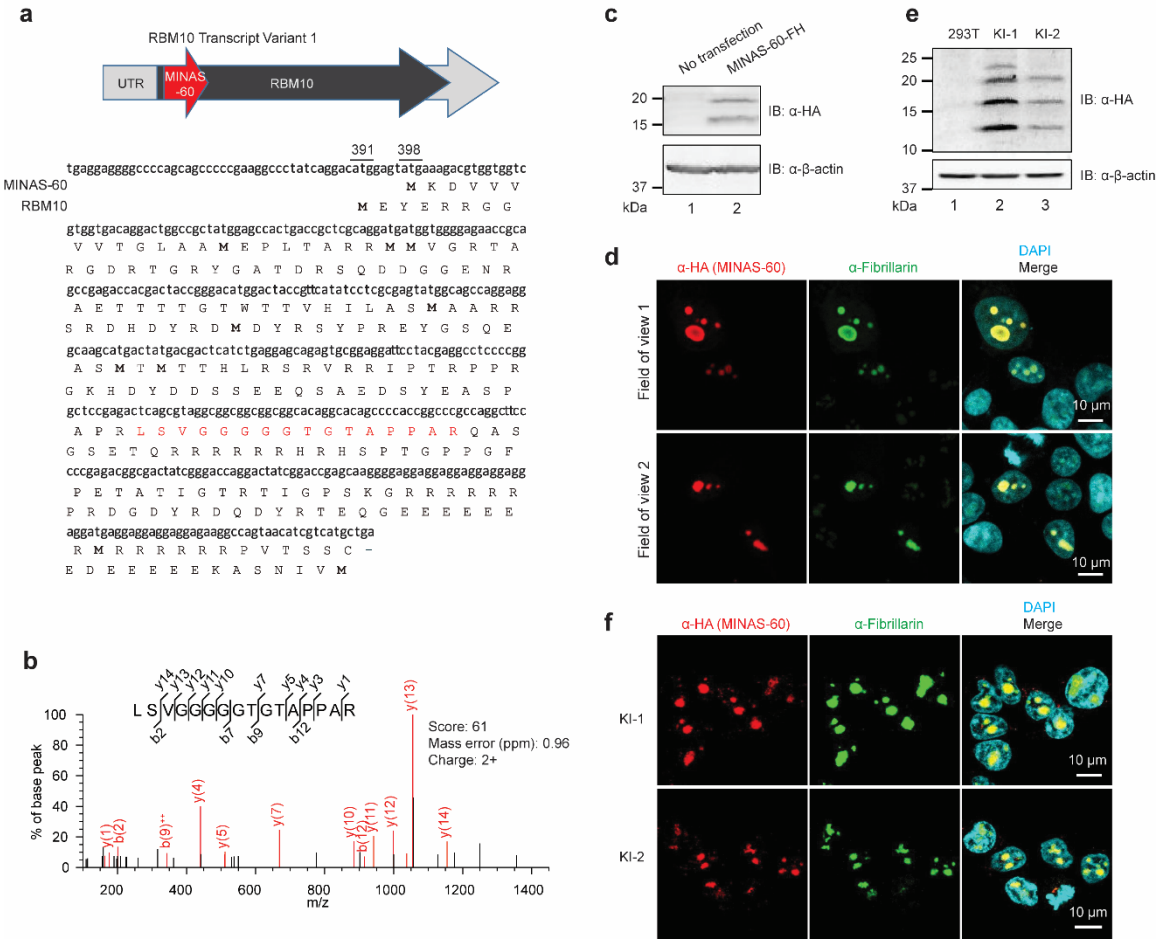


Figure 1. MINAS-60 is endogenously expressed and localizes to the nucleolus of human cells. a Top: a schematic representation of human *RBM10* transcript variant 1; light gray arrow, 5' and 3' untranslated regions (UTR); red, alternative open reading frame (alt-ORF) encoding MINAS-60; black, annotated RBM10 coding sequence. Bottom: the cDNA sequence of human *RBM10* transcript variant 1 is provided from the 5' through the stop codon of the MINAS-60 reading frame. The protein sequences of MINAS-60 and RBM10 are indicated below. The start codons of RBM10 (391) and MINAS-60 (398) are numbered above the cDNA sequence. Highlighted in red is the tryptic peptide of MINAS-60 detected by LC-MS/MS in this study. **b** MS/MS spectrum of the tryptic peptide of MINAS-60 detected by BONCAT-LC-MS/MS peptidomics in HEK 293T cells. **c** Over-expression of a cDNA construct spanning *RBM10* transcript variant 1 from the complete 5'UTR through the MINAS-60 stop codon , with a dual FLAG and

HA tag appended to the C-terminus of MINAS-60 (MINAS-60-FH) in HEK 293T cells, was followed by immunoblotting (IB) with the antibodies indicated to the right, with untransfected HEK 293T cells served as a control. Data are representative of three biological replicates. **d** HEK 293T cells transfected with the MINAS-60 expression construct were immunostained with anti-HA (red), anti-fibrillarin (FBL, green), and DAPI (cyan). Scale bar, 10 μm. Data are representative of three biological replicates. **e** Western blotting of control (lane 1) or MINAS-60 epitope tag CRISPR/Cas9 knock-in (KI, lanes 2 and 3) HEK 293T cells with the antibodies indicated to the right. Data are representative of three biological replicates. **f** Immunostaining of two MINAS-60 KI HEK 293T cell lines with anti-HA (red), anti-FBL (green), and DAPI (cyan). Scale bar, 10 μm. Data are representative of three biological replicates.

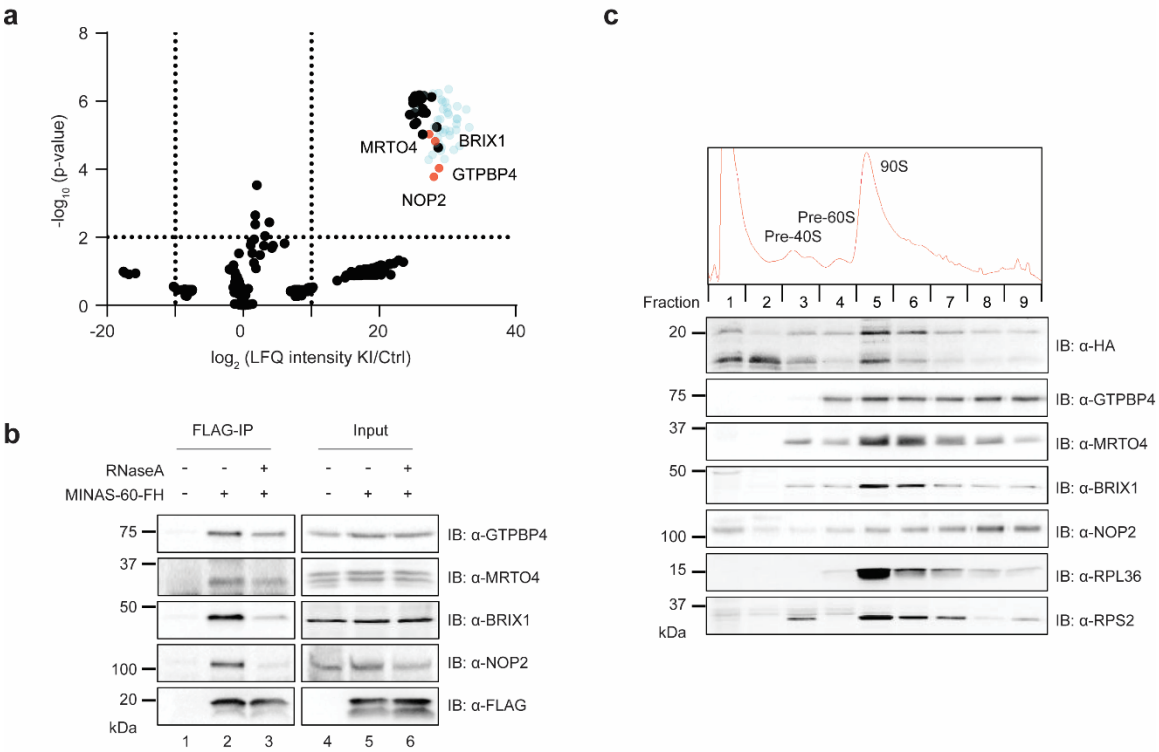


Figure 2. MINAS-60 associates with nucleolar LSU assembly factors. a Volcano plot of quantitative proteomics (*N* = 3) of anti-FLAG pulldown from MINAS-60 KI (KI) or control (Ctrl) HEK 293T nuclear lysates. Ribosomal proteins are indicated in blue. Enriched LSU assembly factors are indicated in red and gene names are labeled. For complete quantitative proteomics results, see Supplementary Data 3. **b** HEK 293T cells were transfected with MINAS-60-FLAG-HA (lanes 2 and 3) or vehicle (lane 1), and immunoprecipitation (FLAG-IP) was performed in absence (lanes 1 and 2) or presence (lane 3) of RNaseA, followed with immunoblotting (IB) with antibodies indicated on the right. Cell lysates (4%) before IP (input, lanes 4-6) were used as loading controls. **c** Top: Sucrose-gradient sedimentation analysis of nuclear lysates containing ribosome precursor complexes (pre-40S, pre-60S and 90S pre-ribosome) from HEK 293T cells stably expressing MINAS-60-FLAG-HA. Bottom: Western blot analysis of fractions numbered at the top with antibodies indicated on the right.

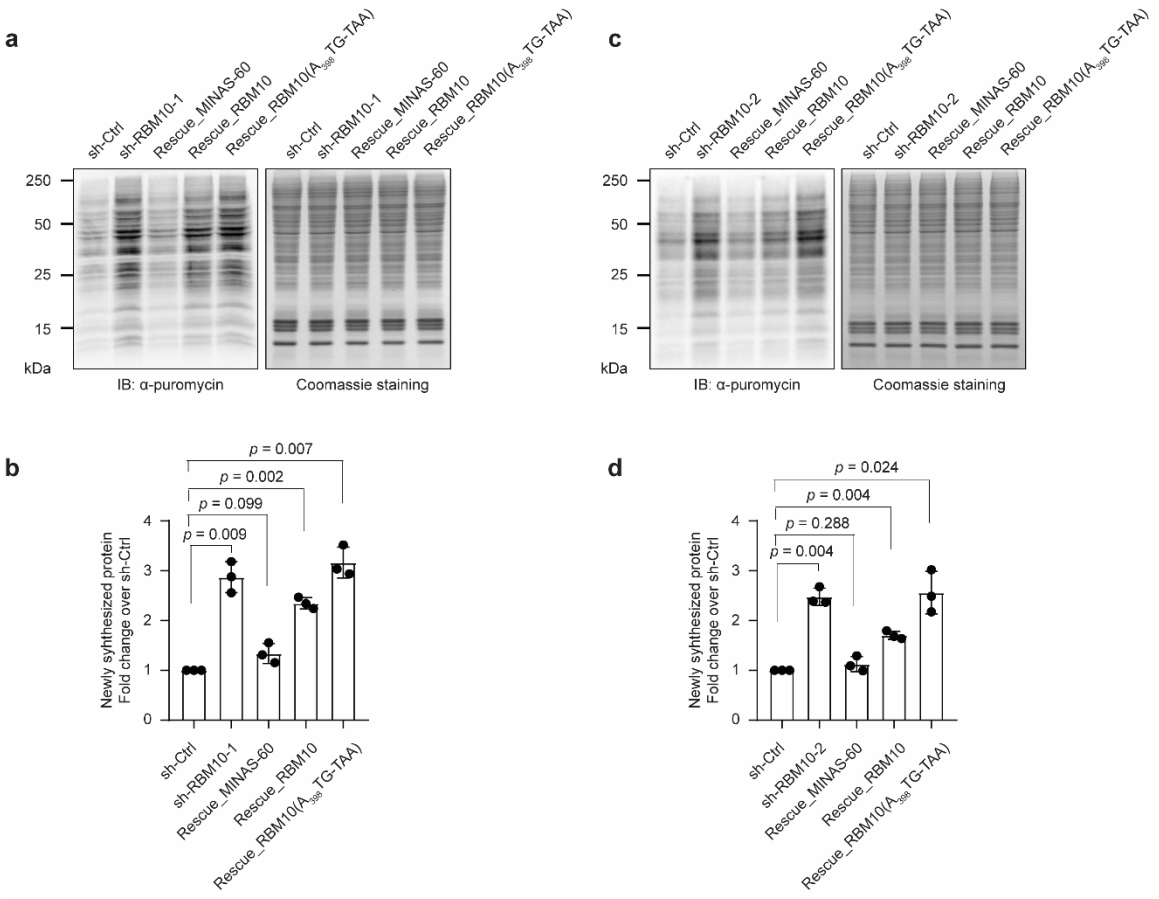
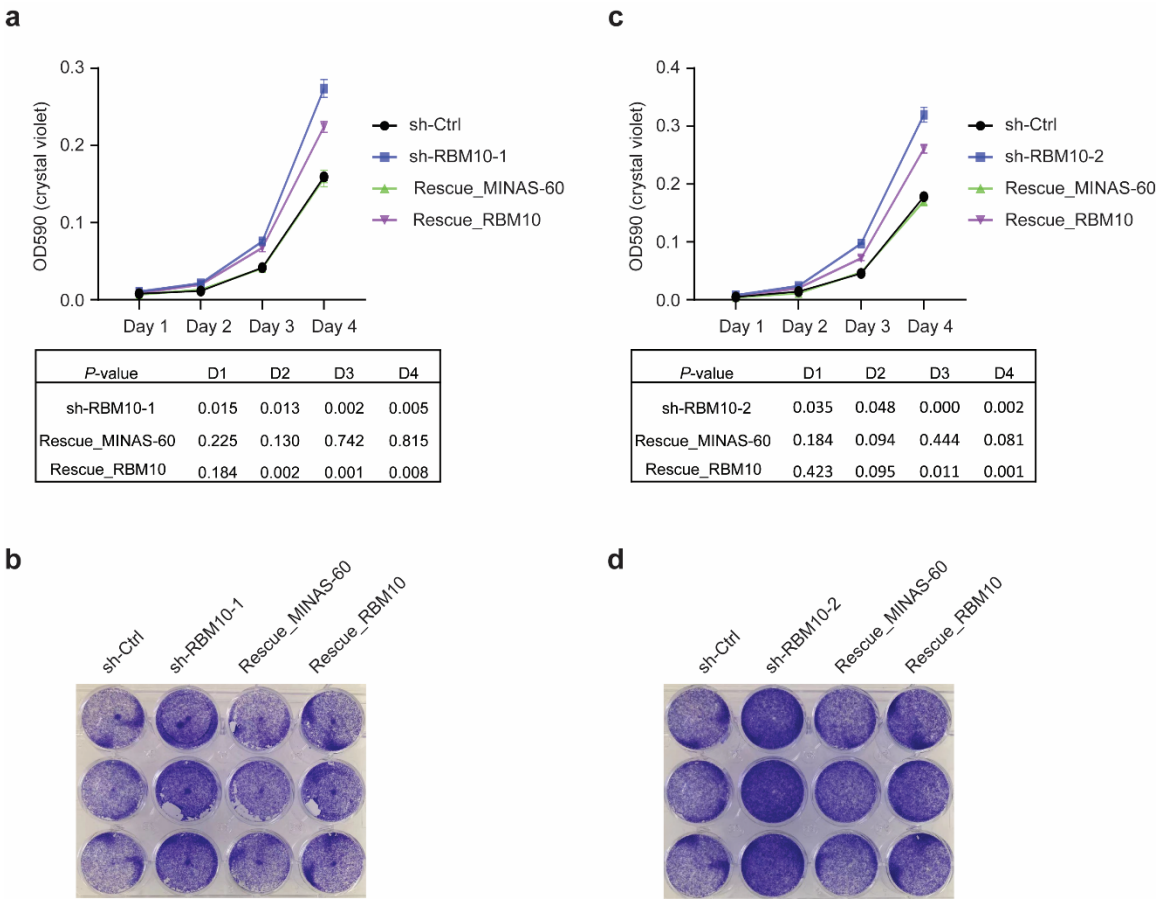


Figure 3. MINAS-60 downregulates global protein synthesis. a, c HEK 293T cells stably expressing empty pLKO.1 vector control (sh-Ctrl), two independent *RBM10* shRNAs (sh-RBM10-1 (**a**) or sh-RBM10-2 (**c**)), rescue with MINAS-60 (Rescue_MINAS-60), rescue with RBM10 (Rescue_RBM10), or rescue with RBM10 bearing an $A_{398}TG$ to TAA mutation that cannot express MINAS-60 (Rescue_RBM10($A_{398}TG$-TAA)) were treated with 1 μM puromycin for 1 hour at 37℃ before harvesting and western blotting with anti-puromycin antibody. Coomassie staining served as a loading control. **b, d** ImageJ was used to quantify the relative puromycin incorporation for cells indicated at the bottom relative to sh-Ctrl from three biological replicates. Data represent mean values ± s.e.m., and significance was evaluated with two-tailed *t*-test.



P-value	D1	D2	D3	D4
sh-RBM10-1	0.015	0.013	0.002	0.005
Rescue_MINAS-60	0.225	0.130	0.742	0.815
Rescue_RBM10	0.184	0.002	0.001	0.008

P-value	D1	D2	D3	D4
sh-RBM10-2	0.035	0.048	0.000	0.002
Rescue_MINAS-60	0.184	0.094	0.444	0.081
Rescue_RBM10	0.423	0.095	0.011	0.001

Figure 4. MINAS-60 downregulates cell proliferation. a, c Growth curve of control (sh-Ctrl), *RBM10* knockdown (sh-RBM10-1 (**a**) or sh-RBM10-2 (**c**)), rescue with MINAS-60 (Rescue_MINAS-60) and rescue with RBM10 (Rescue_RBM10) HEK 293T cells at the indicated number of days ($N = 3$). Data represent mean values ± s.e.m., and significance was evaluated with two-tailed *t*-test and shown below. **b, d** Crystal violet staining of the cell lines described above on day 4.

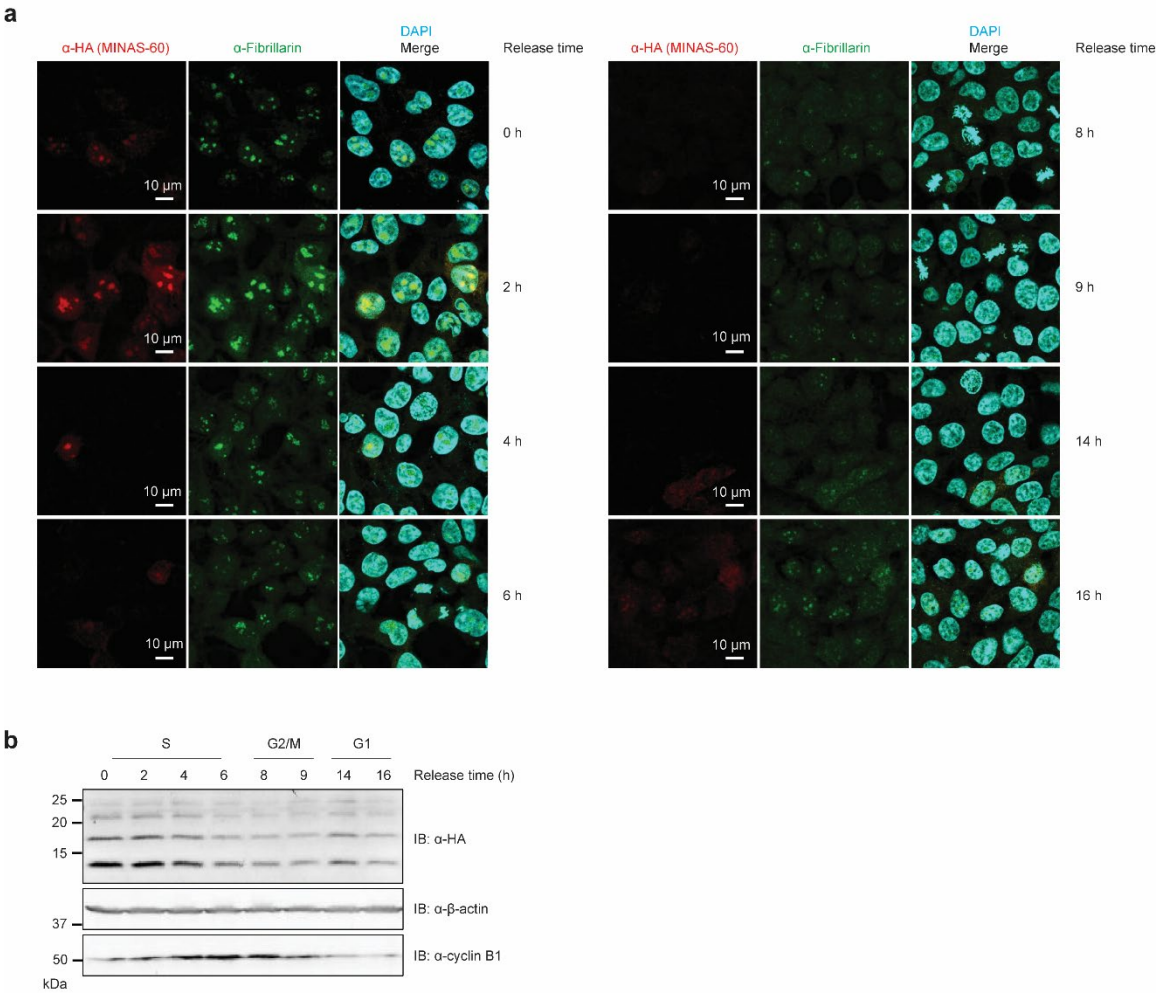


Figure 5. Nucleolar-localized MINAS-60 expression increases in early S phase. a Immunostaining analysis of synchronized MINAS-60 KI cells released from the G1/S boundary by a double thymidine block at the indicated time points with anti-HA (red), anti-FBL (green), and DAPI (cyan). Scale bar, 10 μm. **b** Western blotting analysis of synchronized MINAS-60 KI cells released from the G1/S boundary at the indicated time points with antibodies indicated on the right. All data are representative of three biological replicates.

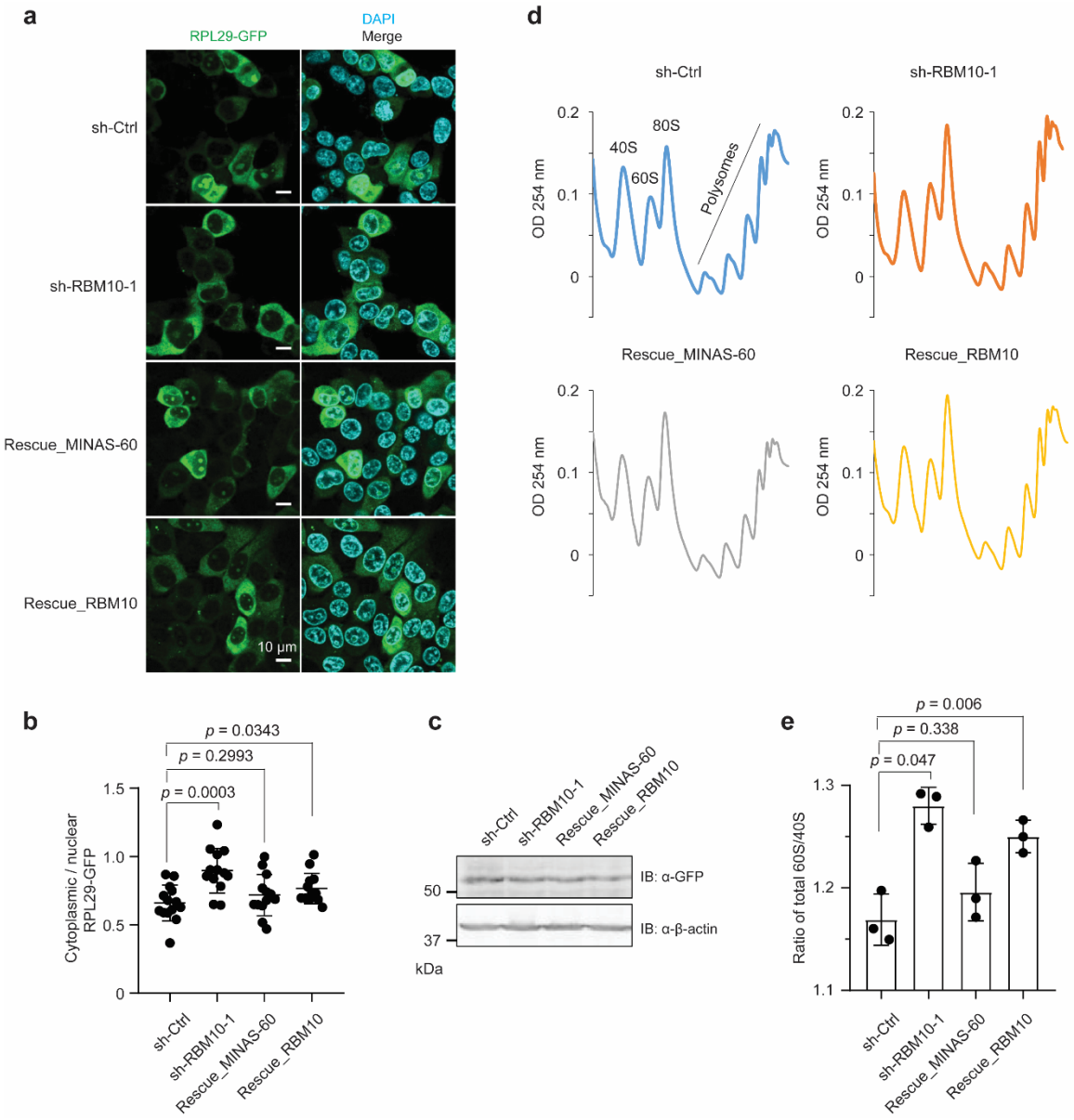


Figure 6. MINAS-60 inhibits LSU export. a Confocal live-cell imaging of control (sh-Ctrl), *RBM10* knockdown (sh-RBM10-1), rescue with MINAS-60 (Rescue_MINAS-60), or rescue with RBM10 (Rescue_RBM10) HEK 293T cells stably expressing RPL29-GFP. Scale bar, 10 μm. Data are representative of three biological replicates. **b** Quantitation of the ratio of RPL29-GFP intensity in the cytoplasm vs. nucleus in the four cell lines described above. At least 13 fields of view were analyzed, totaling > 350 cells for each measurement. Data represent mean values ± s.e.m., and significance was evaluated with two-tailed *t*-

test. **c** Western blot of the four cell lines described above with antibodies indicated on the right for comparison of RPL29-GFP expression. Data are representative of three biological replicates. **d** Sucrose-gradient sedimentation analysis of ribosomal fractions (40S, 60S, 80S and polysomes) of cytoplasmic lysates from control (sh-Ctrl), *RBM10* knockdown (sh-RBM10-1), rescue with MINAS-60 (Rescue_MINAS-60) or rescue with RBM10 (Rescue_RBM10) HEK 293T cells. **e** Quantification of the ratio of cytoplasmic 60S to 40S subunits after sucrose gradient fractionation. The area under each peak was measured using ImageJ following a previously published method[55].

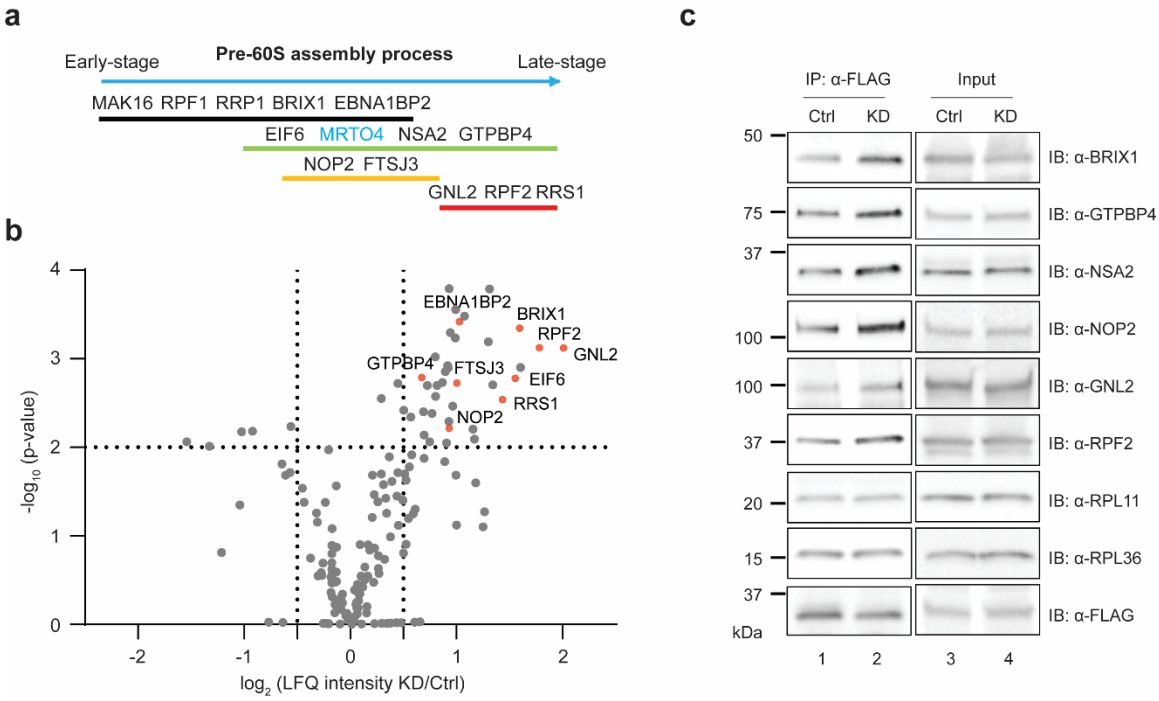


Figure 7. *RBM10* silencing promotes pre-60S assembly. a Schematic representation of pre-60S assembly factors associated with different states based on a pre-60S structure report[16]. The bait protein MRTO4 is indicated in blue. **b** Volcano plot of quantitative proteomics (*N* = 4) of MRTO4-FLAG pulldown from HEK 293T cells stably expressing MRTO4-FLAG and control shRNA (Ctrl), or MRTO4-FLAG and *RBM10* shRNA (KD), to quantify relative changes in MRTO4-FLAG co-IP of LSU assembly factors in *RBM10* KD over control HEK 293T cells. Increased assembly factors are indicated in red and gene names are labeled. For complete quantitative proteomics results, see Supplementary Data 4. **c** MRTO4 FLAG-IP and western blotting with antibodies indicated on the right using the two cell lines described above. Cell lysates (4%) before IP (input) were used as the loading control. Data are representative of three biological replicates.

Online methods

Data analysis. Two-tailed *t*-test were performed using Microsoft Excel or GraphPad Prism, and *F*-tests were performed to evaluate equal variance between samples.

Antibodies. Primary antibodies for western blotting include the following: anti-FLAG (Sigma, F3165 or Cell Signaling, 14793); anti-HA (Invitrogen, 71-5500); anti-β-actin (Invitrogen, BA3R); anti-GTPBP4 (Abclonal, A4565); anti-MRTO4 (ThermoFisher, 20194-1-AP); anti-NSA2 (Abclonal, A14475); anti-NOP2 (Cell Signaling, 25017); anti-BRIX1 (Abclonal, A14481); anti-RPF2 (Abclonal, A17224); anti-GNL2 (Abclonal, A13191); anti-RPL11 (Cell Signaling, 18163); anti-RPL36 (Bethyl Laboratories, A305065A-M); anti-RPS2 (Invitrogen, PA5-30160); anti-RBM10 (Abcam, ab72423); anti-puromycin (Kerafast, EQ0001); anti-GFP (Abcam, ab183734); anti-cyclin B1 (Cell Signaling, 4138). Immunoprecipitation was performed with anti-FLAG M2 affinity gel (Sigma, A2220). Secondary antibodies for western blotting are goat anti-rabbit IgG horseradish peroxidase conjugate (Rockland, 611-1302) and goat anti-mouse IgG horseradish peroxidase conjugate (Rockland, 610-1319-0500). Primary antibodies for immunostaining are rabbit anti-HA (Invitrogen, 71-5500) and mouse anti-Fibrillarin (abcam, ab4566). Secondary antibodies for immunostaining are goat anti-rabbit IgG Alexa fluor 568 (Invitrogen, A11011) and goat anti-mouse IgG Alexa fluor 647 (Invitrogen, A21235).

Cloning and genetic constructs. A DNA sequence comprising the full 5'UTR of human *RBM10* transcript variant 1 through the stop codon of MINAS-60 was amplified by PCR with a dual FLAG and HA epitope tag appended to the 3' end of the MINAS-60 coding sequence from an in-house library of reverse-transcribed HEK 293T cDNAs, then subcloned into pcDNA3. Deletion or mutation constructs bearing a dual FLAG and HA tag were generated by ligating PCR products into BamHI and EcoRI cloning sites in the pcDNA3 vector. For generation of HEK 293T cells stably expressing MINAS-60, a dual FLAG and HA

tag were appended to the 3' end of MINAS-60 by PCR, and the dually tagged coding sequence was then cloned into pLJM1. The cDNA clone expressing RPS2 was purchased from Addgene (a gift from Thomas Tuschl), and the coding sequences of MRTO4, BRIX1 and RPL29 were amplified by PCR from an in-house HEK 293T cDNA library, then subcloned into pJLM1 for producing lentivirus. RNAi constructs were made by synthesizing oligonucleotides encoding a 21 bp short hairpin RNA that targets RBM10 (shRNA1, CTTCGCCTTCGTCGAGTTTAG; shRNA2, TCCAACGTGCGCGTCATAAAG), then subcloned into pLKO.1. The empty pLKO.1 vector control was purchased from Sigma (SHC001). qPCR primer sequences are provided in Supplementary Table 1.

Cell culture, lentivirus production and stable cell line generation. HEK 293T cells were purchased from ATCC and early-passage stocks were established in order to ensure cell line identity. Cells were maintained up to only 10 passages. HEK 293T cells were cultured as previously described[8]. To produce lentivirus and generate stable cell lines, HEK 293T cells were co-transfected using polyethyleneimine (Polysciences, 23966) with expression construct in pLJM1, along with pMD2.G and psPAX2, and growth media were replaced after 7-8 h. 48 h post-transfection, media containing viruses was harvested, filtered through a 0.45-μm filter, and infection was performed by mixing with two volumes of fresh media containing suspended HEK 293T cells. 24 h post-infection, the growth media was replaced. 48 h post-infection, stably expressing cells were selected with 4 μg/mL puromycin for 2 days. Early stocks of stable cell lines were established after selection. Stable cell lines were released from puromycin for 2 days prior to use in experiments.

Immunostaining and live-cell imaging. HEK 293T cells were plated on glass coverslips and transfected the next day. Forty-eight hours later, the cells were fixed in 10% formalin for 15 min at room temperature (RT), permeabilized with PBS containing 0.2% (v/v) TritonX-100, then incubated with primary antibodies

for 18 h at 4℃. After washing with PBS, the cells were incubated with secondary antibodies and DAPI for 1 h at RT, washed with PBS and mounted with Mowiol (Sigma, 81381) before viewing.

HEK 293T cells stably expressing RPL29-GFP or RPS2-GFP were grown to 70% confluency on coverslips in 12-well plates. Coverslips were inverted and imaged in pre-warmed DMEM with 10% FBS, 1% penicillin-streptomycin in MatTek imaging dishes. Confocal imaging was performed on a Leica SP8 LS confocal microscope with 63× oil immersion objective with atmosphere-controlled stage at 37℃. Nuclear/cytoplasmic ratios of RPL29 and RPS2 were measured using the ImageJ Intensity Ratio Nuclei Cytoplasm Tool (RRID:SCR_018573; https://github.com/MontpellierRessourcesImagerie/imagej_macros_and_scripts/wiki/Intensity-Ratio-Nuclei-Cytoplasm-Tool).

Immunoprecipitation and proteomics. Control HEK 293T cells or MINAS-60 KI cells were grown to 80-90% confluency in 15 cm dishes. Cells were harvested and suspended in 1 mL nuclear isolation buffer (10 mM HEPES-KOH pH 7.4, 100 mM KCl, 5 mM $MgCl_2$ with 0.5% NP40 and Roche Complete protease inhibitor cocktail tablets (Roche, 11873580001)), and incubated on ice for 10 min, followed by centrifugation at 3,000 g, 4℃, 3 min. The nuclear pellets were suspended in 1 mL lysis buffer (Tris-buffered saline (TBS) with 1% Triton X-100 and Roche Complete protease inhibitor cocktail tablets), followed with sonication and immunoprecipitation as previously described[8]. After the final wash, elution was in 40 µL of 3× FLAG peptide (Sigma, F4799), at a final concentration of 100 µg/mL in lysis buffer at 4℃ for 1 h. The eluted proteins were subjected to SDS-PAGE separation prior to LC-MS/MS analysis.

Gel slices, containing either resolved protein bands or entire lanes, were digested with trypsin at 37℃ for 14-16 h. The resulting peptide mixtures were extracted from the gel, dried, subjected to ethyl acetate extraction to remove residual detergent, de-salted with peptide cleanup C18 spin column (Agilent

Technologies, 5188-2750), then re-suspended in 35 µl 0.1% formic acid (FA), followed by centrifugation at 21,130 g, 4℃, 30 min. A 5 µL aliquot of each sample was injected onto a pre-packed column attached to a nanoAcquity UPLC (Waters) in-line with a Thermo Scientific™ Q Exactive™ Plus Hybrid QuadrupoleOrbitrap™ mass spectrometer (Thermo Scientific) and a 130-min gradient was used to further separate the peptide mixtures as follows (solvent A: 0.1% FA; solvent B: acetonitrile with 0.1% FA): Isocratic flow was maintained at 0.1 µL/min at 1% B for 40 min, followed by linear gradients from 1% B to 6% B over 2 min, 6% B to 24% B over 48 min, 24% B to 48% B over 5 min, 48% B to 80% B over 5 min. Isocratic flow at 80% B was maintained for 5 min, followed by a gradient from 80% B to 1% B over 5 min, and isocratic flow at 1% B was maintained for 10 min. The full MS was collected over the mass range of 300-1,700 m/z with a resolution of 70,000 and the automatic gain control (AGC) target was set as 3×10^6. MS/MS data was collected using a top 10 high-collisional energy dissociation method in data-dependent mode with a normalized collision energy of 27.0 eV and a 1.6 m/z isolation window. MS/MS resolution was 17,500 and dynamic exclusion was 90 seconds.

For identification of alt- and microproteins, ProteoWizard MS Convert was used for peak picking and files were analyzed using Mascot. Oxidation of methionine and N-terminal acetylation were set as variable modifications, and a previously reported[35] three-frame translation of assembled transcripts from HEK 293T mRNA-seq was used as the database. For co-IP proteomics searches and quantitative analysis, files were analyzed using MaxQuant, oxidation of methionine and N-terminal acetylation were set as variable modifications, and human UniProt plus MINAS-60 was used as the database for searching. For all analysis, a mass deviation of 20 p.p.m. was set for MS1 peaks, and 0.6 Da was set as maximum allowed MS/MS peaks with a maximum of two missed cleavages. Maximum false discovery rates (FDR) were set to 1% both on peptide and protein levels. Minimum required peptide length was five amino acids. Protein quantitation was accomplished by calculating the LFQ intensity ratio of KI

or KD pulldown to negative control samples using MaxQuant (version 1.6.8.0) with standard parameters.

BONCAT (bio-orthogonal non-canonical amino acid tagging) and streptavidin pulldown. HEK 293T cells were grown to 80-90% confluency in 15 cm dishes, and were immersed in methionine-free DMEM (Corning, 17-204-CI) for 30 min before addition of 4 mM AHA (Click Chemistry tools, 1066-1000) in methionine-free DMEM with 10% FBS. After a 2 h incubation, the cells were washed twice with cold PBS, harvested, and lysed in 2 mL 1% (w/v) SDS in PBS by boiling for 10 min and then brought up to 0.1% SDS by adding 18 mL PBS with 0.2% TritonX-100. In a separate tube 200 µL of 70 mM $CuSO_4$ was combined with 20 µL of 100 mM Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA, Click Chemistry tools, 1061-100) and 70 µL of 500 mM Tris(2-carboxyethyl)phosphine hydrochloride, then incubated for 10 min at RT. 20 µL of 25 mM biotin-alkyne (BALK, Click Chemistry tools, 1266-5) was added to the mixture and then combined with the 20 mL cell lysate, followed with rotation at RT for 6 h. The supernatant was then passed through a 3 kDa Amicon Ultra 15 mL column, and diluted 1000 times with RIPA buffer (10 mM Tris-HCl (pH=7.4), 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 140 mM NaCl) to remove excess BALK. A 50 µL aliquot of streptavidin magnetic beads (NEB) was washed twice with 1 mL RIPA buffer before incubation with 1 mL of biotinylated protein for 1 h at RT with gentle rotation. After binding, beads were washed twice with 1 mL RIPA buffer, once with 1 M KCl, once with 0.1 M sodium carbonate, once with 2 M urea, and twice with RIPA lysis buffer. Bound proteins were eluted by boiling in 1×SDS loading buffer containing 20 mM DTT and 2 mM biotin for 10 min. After elution, proteins were electrophoresed on a 16% tricine gel, and gel bands corresponding in the 2-15 kDa size range were excised for proteomic analysis.

Generation of knock-in cell lines. MINAS-60 3xGFP11-FLAG-HA KI HEK 293T cells were generated using CRISPR-Cas9. Guide RNAs (gRNAs) were designed

with the guide design tool from the Zhang lab (crispr.mit.edu) to target the RBM10 genomic region gRNA 5'- TGTCGGCCAGGATTCCTACG-3'. Double-stranded DNA oligonucleotides corresponding to the gRNAs were inserted into pSpCas9(BB)-2A-GFP vector (Addgene, as a gift from F. Zhang, MIT, Cambridge, MA). A donor plasmid containing 300 bp homology left-arm and 300 bp homology right-arm sequence around the stop codon of MINAS-60, which are separated with 3xGFP11-FLAG-HA tag and BamHI / NotI restriction sites was synthesized by GenScript, and a DNA sequence containing pGK promoter and hygromycin resistance gene were subcloned into the donor plasmid using the BamH1 and NotI restriction sites. An equal mixture of the gRNA and donor plasmids were transfected into HEK 293T cells using polyethyleneimine, and hygromycin selection was performed 2 days post-transfection. MINAS-60-3xGFP11-FLAG-HA KI cells were confirmed by genomic DNA PCR and sequencing.

Puromycin incorporation. SUnSET was used to measure protein synthesis[56]. Briefly, HEK 293T cells were grown to 80-90% confluency in 6 well plates, then growth media was replaced with media containing 1μM puromycin and cells were incubated at 37℃ in a humidified atmosphere with 5% CO2 for 1 h according to a previously published protocol[57]. The cells were then washed once with PBS, harvested and analyzed with western blotting.

Sucrose gradient profiling. HEK 293T cells were seeded in 15 cm dishes at 2.5×10^7 cells per dish and cultured 24 h. Cells were then treated with 100 μg/mL cycloheximide for 5 min, washed with cold PBS containing 100 μg/mL cycloheximide twice, harvested and flash frozen until further processing.

For cytoplasmic polysome profiling, after thawing on ice, cells were lysed in polysome lysis buffer (20 mM HEPES-KOH pH 7.4, 100 mM KCl, 5 mM $MgCl_2$, 100 μg/mL cycloheximide, 1 mM DTT with 1% TritonX-100, Roche Complete protease inhibitor cocktail tablets and Ribonuclease Inhibitors (Promega N2511)),

and incubated on ice for 10 min, followed by centrifugation at 21,130 g, 4℃, 10 min. The supernatants were normalized according to absorbance (A260) and layered onto 12 mL 10-50% sucrose gradients (20 mM HEPES-KOH pH 7.4, 100 mM KCl, 5 mM $MgCl_2$, 100 µg/mL cycloheximide, 1 mM DTT with Roche Complete protease inhibitor cocktail tablets and Ribonuclease Inhibitors), followed with centrifugation in an SW-41Ti rotor at 252,878 g, 4℃, 3 h, then sampled using a Biocomp Gradient profiler (model 251) with constant monitoring of optical density at 254 nm using standard parameters. Data analysis was performed with Excel.

For sucrose gradient sedimentation analysis of nuclear lysates, thawed cells were suspended in 1 mL nuclear isolation buffer (10 mM HEPES-KOH pH 7.4, 100 mM KCl, 5 mM MgCl2 with 0.5% NP40, Roche Complete protease inhibitor cocktail tablets and Ribonuclease Inhibitors), and incubated on ice for 10 min, followed by centrifugation at 3,000 g, 4℃, 3 min. The nuclear pellets were suspended in 1 mL nuclear lysis buffer (20 mM HEPES-KOH pH 7.4, 300 mM KCl, 5 mM MgCl2, 1 mM DTT with 1% TritonX-100, Roche Complete protease inhibitor cocktail tablets and Ribonuclease Inhibitors), followed with sonication (50% intensity, 5 s pulse with 25 s rest, 5×, MICROSON XL 2000), and centrifugation at 21,130 g, 4℃, 10 min. The supernatants were normalized by A260 and layered onto 12 mL 10-50% sucrose gradients, centrifuged and sampled as described above.

Crystal violet staining and cell cycle synchronization. 4×10^4 HEK 293T cells were seeded in 12-well plates in triplicate, and fixed in 10% formalin for 15 min at RT every 24 h. After washing with ddH_2O, cells were stained with 0.1% crystal violet in methanol for 30 min at RT in the dark, followed with three ddH_2O washes and dried. The cells were then immersed in 1 mL 10% acetic acid with shaking for 20 min. 20 µL of the solution was combined with 80 µL ddH_2O in a 96-well plate, and the optical density at 590 nm was monitored with Synergy™ HT.

For cell cycle synchronization, 1.5×10^5 HEK 293T cells were seeded in 12-well plates and cultured overnight. Cells were then treated with 2 mM thymidine for 16 h, followed with two PBS washes, and incubated with fresh media for 9 h before the second 2 mM thymidine block for 14 h following a previously published protocol[58]. Cells were washed with PBS, then incubated with fresh media to release from the G1/S boundary, and fixed in 10% formalin for immunostaining or harvested for western blotting.

Northern blot

Total RNA was extracted from HEK 293T cells using TRIzol reagent. To determine changes in levels of LSU pre-rRNA intermediates, 3 µg of total RNA was run on a 1% agarose/1.25% formaldehyde gel in a 1.5 M tricine/ 1.5 M triethanolamine buffer, followed by an overnight transfer to a Hybond XL nylon membrane (GE Healthcare, RPN 303S) in 10× saline-sodium citrate transfer buffer after a brief 15 min soak in a 0.5 M sodium hydroxide solution. Membranes were then exposed to UV (254 nm) to immobilize the RNA, followed by incubation with denatured yeast tRNA for 1 h at 42℃, and hybridized overnight at 37℃ with 5' end radiolabeled oligonucleotide probe (P4 5'-CGGGAACTCGGCCCGAGCCGGCTCTCTCTTTCCCTCTCCG-3') in a solution of 7.5× Denhardt's solution, 5× sodium chloride-sodium phosphate-EDTA buffer with 0.1% SDS, as previously reported[50]. Membranes were also hybridized with a 7SL probe (7SL 5'-TGCTCCGTTTCCGACCTGGGCCGGTTCACCCCTCCTT-3') as a loading control.

Acknowledgements

We thank Franziska Bleichert and all members of the Slavoff and Baserga labs for helpful conversations. This work was supported by a Searle Scholars Program Award, an Odyssey Award from the Richard and Susan Smith Family Foundation, and start-up funds from Yale University West Campus (to S. A. S.). X.C. was supported in part by a Rudolph J. Anderson postdoctoral fellowship

from Yale University. A.K. was in part supported by an NIH Predoctoral Training Grant (5T32GM06754 3-12). S.J.B, C.J.B. and C.M.H. were supported by R35 GM131687. C.M.H. was supported by an NSF GFRP.

Author contributions

X.C., C.M.H., C.J.B., S.Z. and A.K. designed and performed experiments and analyzed data. S. A. S. and S. J. B. designed experiments and analyzed data. X. C. and S. A. S. wrote the manuscript, and all authors edited and approved the final version of the manuscript.

References

1. Basrai, M.A., Hieter, P. & Boeke, J.D. Small open reading frames: beautiful needles in the haystack. *Genome Res* **7**, 768-71 (1997).
2. Orr, M.W., Mao, Y., Storz, G. & Qian, S.B. Alternative ORFs and small ORFs: shedding light on the dark proteome. *Nucleic Acids Res* **48**, 1029-1042 (2020).
3. Brunet, M.A., Levesque, S.A., Hunting, D.J., Cohen, A.A. & Roucou, X. Recognition of the polycistronic nature of human genes is critical to understanding the genotype-phenotype relationship. *Genome Res* **28**, 609-624 (2018).
4. Lee, C.Q.E. et al. Coding and non-coding roles of MOCCI (C15ORF48) coordinate to regulate host inflammation and immunity. *Nat Commun* **12**, 2130 (2021).
5. Zhang, S. et al. Mitochondrial peptide BRAWNIN is essential for vertebrate respiratory complex III assembly. *Nat Commun* **11**, 1312 (2020).
6. Huang, N. et al. An Upstream Open Reading Frame in Phosphatase and Tensin Homolog Encodes a Circuit Breaker of Lactate Metabolism. *Cell Metab* **33**, 128-144 e9 (2021).
7. Brunet, M.A. et al. The FUS gene is dual-coding with both proteins contributing to FUS-mediated toxicity. *EMBO Rep* **22**, e50640 (2021).
8. Cao, X. et al. Alt-RPL36 downregulates the PI3K-AKT-mTOR signaling pathway by interacting with TMEM24. *Nat Commun* **12**, 508 (2021).
9. Chen, J. et al. Pervasive functional translation of noncanonical human open reading frames. *Science* **367**, 1140-1146 (2020).
10. Prensner, J.R. et al. Noncanonical open reading frames encode functional proteins essential for cancer cell survival. *Nat Biotechnol* (2021).
11. Brunet, M.A. et al. OpenProt 2021: deeper functional annotation of the coding potential of eukaryotic genomes. *Nucleic Acids Res* **49**, D380-D388 (2021).
12. Klinge, S. & Woolford, J.L., Jr. Ribosome assembly coming into focus. *Nat Rev Mol Cell Biol* **20**, 116-131 (2019).
13. Pena, C., Hurt, E. & Panse, V.G. Eukaryotic ribosome assembly, transport and quality control. *Nat Struct Mol Biol* **24**, 689-699 (2017).
14. De Keersmaecker, K., Sulima, S.O. & Dinman, J.D. Ribosomopathies and the paradox of cellular hypo- to hyperproliferation. *Blood* **125**, 1377-82 (2015).

15. Nerurkar, P. et al. Eukaryotic Ribosome Assembly and Nuclear Export. *Int Rev Cell Mol Biol* **319**, 107-40 (2015).
16. Kater, L. et al. Visualizing the Assembly Pathway of Nucleolar Pre-60S Ribosomes. *Cell* **171**, 1599-1610 e14 (2017).
17. Sanghai, Z.A. et al. Modular assembly of the nucleolar pre-60S ribosomal subunit. *Nature* **556**, 126-129 (2018).
18. Micic, J. et al. Coupling of 5S RNP rotation with maturation of functional centers during large ribosomal subunit assembly. *Nat Commun* **11**, 3751 (2020).
19. Pelletier, J., Thomas, G. & Volarevic, S. Ribosome biogenesis in cancer: new players and therapeutic avenues. *Nat Rev Cancer* **18**, 51-63 (2018).
20. Nyhus, C., Pihl, M., Hyttel, P. & Hall, V.J. Evidence for nucleolar dysfunction in Alzheimer's disease. *Rev Neurosci* **30**, 685-700 (2019).
21. Farley-Barnes, K.I., Ogawa, L.M. & Baserga, S.J. Ribosomopathies: Old Concepts, New Controversies. *Trends Genet* **35**, 754-767 (2019).
22. Mills, E.W. & Green, R. Ribosomopathies: There's strength in numbers. *Science* **358**(2017).
23. Parker, M.D., Collins, J.C., Korona, B., Ghalei, H. & Karbstein, K. A kinase-dependent checkpoint prevents escape of immature ribosomes into the translating pool. *PLoS Biol* **17**, e3000329 (2019).
24. Liang, X. et al. Structural snapshots of human pre-60S ribosomal particles before and after nuclear export. *Nat Commun* **11**, 3542 (2020).
25. Hernandez, J. et al. Tumor suppressor properties of the splicing regulatory factor RBM10. *RNA Biol* **13**, 466-72 (2016).
26. Inoue, A. et al. RBM10 regulates alternative splicing. *FEBS Lett* **588**, 942-7 (2014).
27. Rodor, J., FitzPatrick, D.R., Eyras, E. & Caceres, J.F. The RNA-binding landscape of RBM10 and its role in alternative splicing regulation in models of mouse early development. *RNA Biol* **14**, 45-57 (2017).
28. Sutherland, L.C. et al. Splicing arrays reveal novel RBM10 targets, including SMN2 pre-mRNA. *BMC Mol Biol* **18**, 19 (2017).
29. Bechara, E.G., Sebestyen, E., Bernardis, I., Eyras, E. & Valcarcel, J. RBM5, 6, and 10 differentially regulate NUMB alternative splicing to control cancer cell proliferation. *Mol Cell* **52**, 720-33 (2013).
30. Dieterich, D.C., Link, A.J., Graumann, J., Tirrell, D.A. & Schuman, E.M. Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc Natl Acad Sci U S A* **103**, 9482-7 (2006).
31. Khitun, A. & Slavoff, S.A. Proteomic Detection and Validation of Translated Small Open Reading Frames. *Curr Protoc Chem Biol* **11**, e77 (2019).
32. Wang, T., Liu, Y., Liu, Q., Cummins, S. & Zhao, M. Integrative proteomic analysis reveals potential high-frequency alternative open reading frame-encoded peptides in human colorectal cancer. *Life Sci* **215**, 182-189 (2018).
33. Kamiyama, D. et al. Versatile protein tagging in cells with split fluorescent protein. *Nat Commun* **7**, 11046 (2016).
34. Cao, X. & Slavoff, S.A. Non-AUG start codons: Expanding and regulating the small and alternative ORFeome. *Exp Cell Res* **391**, 111973 (2020).
35. Slavoff, S.A. et al. Peptidomic discovery of short open reading frame-encoded peptides in human cells. *Nat Chem Biol* **9**, 59-64 (2013).
36. Tan, L.Y. et al. Generation of functionally distinct isoforms of PTBP3 by alternative splicing and translation initiation. *Nucleic Acids Res* **43**, 5586-600 (2015).

37. Saghatelian, A. & Couso, J.P. Discovery and characterization of smORF-encoded bioactive polypeptides. *Nat Chem Biol* **11**, 909-16 (2015).
38. Cox, J. et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**, 2513-26 (2014).
39. Liu, Y., Deisenroth, C. & Zhang, Y. RP-MDM2-p53 Pathway: Linking Ribosomal Biogenesis and Tumor Surveillance. *Trends Cancer* **2**, 191-204 (2016).
40. Sondalle, S.B., Longerich, S., Ogawa, L.M., Sung, P. & Baserga, S.J. Fanconi anemia protein FANCI functions in ribosome biogenesis. *Proc Natl Acad Sci U S A* **116**, 2561-2570 (2019).
41. Iyer-Bierhoff, A. & Grummt, I. Stop-and-Go: Dynamics of Nucleolar Transcription During the Cell Cycle. *Epigenet Insights* **12**, 2516865719849090 (2019).
42. Andrews, J.O. et al. qSR: a quantitative super-resolution analysis tool reveals the cell-cycle dependent organization of RNA Polymerase I in live human cells. *Sci Rep* **8**, 7424 (2018).
43. Gorski, S.A., Snyder, S.K., John, S., Grummt, I. & Misteli, T. Modulation of RNA polymerase assembly dynamics in transcriptional regulation. *Mol Cell* **30**, 486-97 (2008).
44. Hernandez-Verdun, D. Assembly and disassembly of the nucleolus during the cell cycle. *Nucleus* **2**, 189-94 (2011).
45. Bassler, J. & Hurt, E. Eukaryotic Ribosome Assembly. *Annu Rev Biochem* **88**, 281-306 (2019).
46. Wild, T. et al. A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. *PLoS Biol* **8**, e1000522 (2010).
47. Chu, Q. et al. Identification of Microprotein-Protein Interactions via APEX Tagging. *Biochemistry* **56**, 3299-3306 (2017).
48. Nguyen le, X.T., Raval, A., Garcia, J.S. & Mitchell, B.S. Regulation of ribosomal gene expression in cancer. *J Cell Physiol* **230**, 1181-8 (2015).
49. Xing, Y.H. et al. SLERT Regulates DDX21 Rings Associated with Pol I Transcription. *Cell* **169**, 664-678 e16 (2017).
50. Ogawa, L.M. et al. Increased numbers of nucleoli in a genome-wide RNAi screen reveal proteins that link the cell cycle to RNA polymerase I transcription. *Mol Biol Cell* **32**, 956-973 (2021).
51. Klemke, M., Kehlenbach, R.H. & Huttner, W.B. Two overlapping reading frames in a single exon encode interacting proteins--a novel way of gene usage. *EMBO J* **20**, 3849-60 (2001).
52. Inoue, A. RBM10: Structure, functions, and associated diseases. *Gene* **783**, 145463 (2021).
53. Niceta, M. et al. TARP syndrome: Long-term survival, anatomic patterns of congenital heart defects, differential diagnosis and pathogenetic considerations. *Eur J Med Genet* **62**, 103534 (2019).
54. Imielinski, M. et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**, 1107-20 (2012).
55. Gregory, B. et al. The small and large ribosomal subunits depend on each other for stability and accumulation. *Life Sci Alliance* **2**(2019).
56. Schmidt, E.K., Clavarino, G., Ceppi, M. & Pierre, P. SUnSET, a nonradioactive method to monitor protein synthesis. *Nat Methods* **6**, 275-7 (2009).
57. Farley-Barnes, K.I. et al. Diverse Regulators of Human Ribosome Biogenesis Discovered by Changes in Nucleolar Number. *Cell Rep* **22**, 1923-1934 (2018).

58. Cao, X. et al. Histone H4K20 Demethylation by Two hHR23 Proteins. *Cell Rep* **30**, 4152-4164 e6 (2020).