Nascent alt-protein chemoproteomics reveals a repressor of ribosome biogenesis

Xiongwen Cao[1,2], Alexandra Khitun[1,2], Cecelia M. Harold[3], Carson J. Bryant[4], Shu-Jian Zheng[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

Many unannotated microproteins and alternative proteins (alt-proteins) have recently been found to be co-encoded with canonical proteins, but few of their functions are known. Motivated by the hypothesis that alt-proteins undergoing active or stress-induced synthesis could play important cellular roles, here, we developed a chemoproteomic pipeline to identify nascent alt-proteins in human cells. We identified 22 actively translated unannotated alt-proteins, one of which is upregulated after DNA damage stress. We further defined MINAS-60 (*MI*croprotein that *N*egatively regulates *AS*sembly of the pre-*60*S ribosomal subunit), a nucleolar localized alt-protein co-encoded with human RBM10.

Depletion of MINAS-60 increases the amount of the mature 60S ribosomal subunit, consequently upregulating global protein synthesis and cell proliferation by repressing late-stage pre-60S assembly and export of the 60S ribosome subunit to the cytoplasm. Together, these results implicate MINAS-60 as a repressor of ribosome biogenesis, and demonstrate that chemoproteomics can enable generation of functional hypotheses for uncharacterized alt-proteins.

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 well-defined, including alt-FUS, which cooperates with FUS in formation of cytotoxic aggregates[7], and alt-RPL36, which 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 human smORFs and alt-ORFs[3,11], but the overwhelming majority of smORFs and alt-ORFs remain uncharacterized, in part because their short lengths and, in some cases, limited conservation render homology-based annotation challenging[12]. Furthermore, most alt-ORFs remain unstudied because

it is challenging to separate their functions from the canonical protein CDS in which they are nested. Because strategies to identify alt-proteins that participate in biological processes are currently in the developmental stages, it remains unclear whether alt-proteins are broadly functional. We hypothesize that alt-proteins with properties (e.g., chemical reactivity, regulated expression) similar to those of canonical proteins are likely to play important cellular roles, and that chemoproteomic strategies targeted to those properties can be leveraged to identify functional alt-proteins. Providing precedent, a chemoproteomic profiling study of microproteins containing reactive cysteine residues, which is a feature of the active site of many enzymes, revealed 16 nucleophilic microproteins, though their cellular roles were not explored in that study[13]. In this work, we test the hypothesis that alt-proteins undergoing active or stress-induced synthesis are likely to be functional, and we develop a chemoproteomic approach to identify them.

Using this method, we identify 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[14-17]. 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 complexed 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)[18,19], revealing a sequential maturation process involving quality control checkpoints for 5S ribonucleoprotein particle incorporation and rotation[20], 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,

92 which associate to form translation-competent ribosomes. Dysregulated
93 ribosome biogenesis has been linked to numerous human disorders, including
94 cancer[21], Alzheimer's disease[22], and congenital disorders termed
95 ribosomopathies[23,24].
96
97 In this study, we developed a chemoproteomic pipeline to identify nascent alt-
98 proteins, which modifies a powerful previously reported strategy for metabolic
99 unnatural amino acid incorporation to be amenable to microprotein enrichment
100 and profiling, and identified 22 unannotated alt-proteins in a human cell line. We
101 confirmed the translation of six selected alt-proteins, and functionally defined one
102 alt-ORF nested in the human RBM10 CDS. We name the encoded alt-protein
103 MINAS-60, or *MI*croprotein that *N*egatively regulates *AS*sembly of the pre-*60S*
104 ribosomal subunit. We show that MINAS-60 localizes to the nucleolus,
105 associates with LSU assembly factors GTPBP4 and MRTO4, and co-fractionates
106 with pre-60S complexes in nuclear extracts. Finally, we engineer MINAS-60
107 knockdown and rescue cell lines to demonstrate that loss of MINAS-60 increases
108 cytoplasmic 60S ribosome subunit levels, global protein synthesis, and cell
109 proliferation. This is independent of the function of the canonical RBM10 protein,
110 which has been previously shown to inhibit cell proliferation through its role in
111 alternative pre-mRNA splicing[25-29]. These results implicate MINAS-60 as a rare
112 negative regulator of late nuclear steps in LSU biogenesis, and demonstrate that
113 chemoproteomic profiling can prioritize alt-ORFs for functional study.
114
115 **Results**
116 **Chemoproteomic profiling of nascent alt-proteins**
117 Motivated by the hypothesis that alt-proteins undergoing active synthesis or
118 stress-induced synthesis could play important cellular roles, we set out to
119 develop a chemoproteomic approach to identify newly translated alt-proteins
120 (Figure 1a). We leveraged bio-orthogonal non-canonical amino acid tagging
121 (BONCAT), an *in vivo* labeling strategy to identify nascent proteins[30]. BONCAT
122 utilizes the methionine analogue azidohomoalanine (AHA), which can be

123 metabolically incorporated into all newly synthesized proteins by the endogenous
124 protein translation machinery. Previous BONCAT workflows included procedures,
125 such as column-based biotin removal, that were likely to eliminate small proteins.
126 We modified the protocol to capture these small proteins, including microproteins
127 and alt-proteins. Labeled (and unlabeled) small proteins were selectively
128 enriched from whole proteome extracts with a C8 column following a previously
129 reported strategy[31] (Figure 1b), and AHA-containing small proteins were captured
130 with click chemistry directly on dibenzocyclooctyne (DBCO) magnetic beads. On-
131 bead digest was followed with our previously reported liquid
132 chromatography/tandem mass spectrometry-based method for identification of
133 unannotated microproteins and alt-proteins[32]. Using this strategy, we profiled
134 translation in HEK 293T cells under basal growth conditions (DMSO treatment),
135 oxidative stress (sodium arsenite treatment), DNA damage stress (etoposide
136 treatment) and unfolded protein response stress (DTT treatment). In total, we
137 identified 22 unannotated alt-proteins (Figure 1c, Supplementary Data 1), nine of
138 which were specifically detected under stress conditions, including alt-CNPY2,
139 which was specifically detected after etoposide treatment (Figure 1d,
140 Supplementary Data 1).
141
142 To confirm the translation and examine subcellular localization of six selected alt-
143 proteins (Supplementary Figures 1a-f, Supplementary Table 1), the cDNA
144 sequence comprising the 5'UTR of the encoding transcript through the stop
145 codon of the putative alt-ORF was cloned into a mammalian expression vector
146 with a FLAG-HA epitope tag on the C-terminus of the alt-protein, followed by
147 transfection and immunostaining. As shown in Figure 1e, over-expressed alt-
148 DRAP1 (previously validated[12] and included as a positive control), alt-PRR3, alt-
149 PRH1, alt-CNPY2 and alt-CACTIN were nucleocytoplasmic. Over-expressed
150 MINAS-60 co-localized with a nucleolar protein, fibrillarin (Figure 1f). Western
151 blotting further confirmed the translation of the six alt-proteins (Figure 1g).
152 MINAS-60 and alt-CNPY2 produced two immunoreactive bands, which is due to
153 multiple in-frame start codon for MINAS-60 (vide infra), and may be caused by

154 multiple start sites, phosphorylation or other post-translational modification for alt-
155 CNPY2, analogous to alt-RPL36[8].
156
157 We hypothesized that the alt-proteins detected specifically under stress
158 conditions are induced in response to cellular stress. To determine whether the
159 expression of alt-CNPY2 is induced by DNA damage, we treated HEK 293T
160 over-expressing alt-CNPY2 with etoposide, followed by western blotting. As
161 shown in Figure 1h, alt-CNPY2 is upregulated up to two fold upon etoposide
162 treatment in both dose- and time-dependent manner. However, the mRNA level
163 of *CNPY2* did not change (Supplementary Figure 1g), indicating that upregulation
164 of alt-CNPY2 is likely due to increased translation or decreased proteolysis, but
165 not increased transcription. Taken together, these results suggest that our
166 chemoproteomic pipeline is able to detect nascent or stress-induced alt-proteins.
167
168 **MINAS-60 is cell-cycle regulated and conserved in mouse**
169 We selected MINAS-60 from the BONCAT-detected microprotein dataset for
170 further study because it localizes to the nucleolus, and only one nucleolar
171 microprotein has previously been identified[33]; furthermore, MINAS-60 is nested
172 within the RBM10 CDS, and therefore probing its cellular and molecular roles
173 could shed light on the poorly characterized class of overlapping alt-proteins
174 (Figure 2a, Supplementary Data 2). A different MINAS-60 tryptic peptide was
175 previously detected in human colorectal cancer samples, supporting expression
176 of MINAS-60 in human tissue, but the alt-protein was not defined or
177 characterized in that study[34]. To validate expression of endogenous MINAS-60
178 from the *RBM10* genomic locus, we generated two independent Cas9-directed
179 knock-in (KI) HEK 293T cell lines with a 3×GFP11-FLAG-HA tag appended to the
180 3' end of MINAS-60 alt-ORF[35], followed with immunostaining. As shown in Figure
181 2b, endogenously expressed MINAS-60 co-localizes with fibrillarin, consistent
182 with the over-expression results, suggesting MINAS-60 likely functions in the
183 nucleolus.
184

185 The nucleolus is the site of ribosome biogenesis in eukaryotic cells. Ribosome
186 biogenesis starts with the transcription of rDNA, which oscillates during the cell
187 cycle, nearly ceasing during M phase, increasing during G1 phase, and
188 maximizing during S and G2 phases in human cells[36-39]. If MINAS-60 regulates
189 ribosome biogenesis, we hypothesized that the expression of nucleolar MINAS-
190 60 would be correlated with ribosome biogenesis during the cell cycle.
191 Immunostaining of synchronized MINAS-60 KI HEK 293T cells revealed that
192 nucleolar MINAS-60 staining intensity increased at early S phase, peaking at the
193 2 h time point. MINAS-60 expression then decreased by late S phase, and was
194 very low during G2/M phase. At G1 phase, the MINAS-60 staining intensity again
195 increased (Supplementary Figure 2). These results were confirmed with western
196 blotting (Figure 2c). MINAS-60 expression is therefore coordinated with ribosome
197 biogenesis activity during the cell cycle.
198
199 We then identified the start codon(s) that initiate MINAS-60 translation. Alt-ORFs
200 have been reported to initiate at upstream, non-AUG start codons[8,12,40] and
201 internal, AUG start codons[7,41]. We tested two upstream in-frame non-AUG start
202 codons, A_{383}TC and A_{386}GG (numbered relative to the first nucleotide of the
203 cDNA), as well as seven internal AUG start codons (numbered ATG1 – ATG7)
204 (Supplementary Figure 3a). The ~20 kDa MINAS-60 isoform likely initiates at
205 ATG1 and the ~15 kDa MINAS-60 isoform likely initiates at ATG6 or 7, because
206 the indicated species is abrogated only when the corresponding start codon is
207 deleted or mutated (Supplementary Figures b-c). These results were further
208 confirmed by over-expressing truncated MINAS-60 coding sequences starting
209 from ATG1 or ATG6 and comparing their products' sizes with the wild-type
210 construct (Supplementary Figure 3d). The MINAS-60 smORF is therefore entirely
211 contained within the RBM10 coding sequence, and the MINAS-60 ATG1 start
212 codon is only 7 nucleotides downstream of the RBM10 start codon (Figure 2a).
213 We therefore speculate that MINAS-60 translation may initiate within *RBM10*
214 transcript variant 1 via leaky scanning, and that the additional MINAS-60 isoform

215 diversity observed in KI cells could be generated by alternative splicing,
216 analogous to PTBP3[41].
217
218 ClustalW alignment of hypothetical MINAS-60 homologs from mouse, cattle, and
219 monkey revealed significant sequence similarity, suggesting MINAS-60 is
220 conserved (Figure 2d). To determine whether mouse MINAS-60 is also
221 translated, a mammalian expression vector contains the 5'UTR of mouse *RBM10*
222 transcript variant 1 through the stop codon of the putative mouse MINAS-60
223 homolog was transfected into 3T3 cells, followed by immunostaining. As shown
224 in Figure 2e, over-expressed mouse MINAS-60 localized to the nucleolus,
225 comparable to human MINAS-60. Taken together, these results indicate that
226 MINAS-60 is endogenously expressed, cell-cycle regulated and conserved from
227 humans to mouse.
228
229 **MINAS-60 is associated with nucleolar LSU assembly factors**
230 Because many SEPs characterized to date bind to and regulate other proteins[42],
231 we performed a co-immunoprecipitation (co-IP) with the nuclear lysates from
232 MINAS-60-FLAG KI cells, and HEK 293T cells as a control. Two major bands
233 were specifically present in the KI co-IP after SDS-PAGE that, upon analysis via
234 label-free quantitative proteomics[43], yielded 17 proteins enriched >30-fold over
235 control. GO analysis of these 17 hits shows that the top 2 enriched biological
236 processes are ribosome biogenesis and ribosomal large subunit (LSU)
237 biogenesis (Supplementary Figures 4a-c, and Supplementary Data 3). To obtain
238 a more comprehensive picture of MINAS-60-associated proteins, we performed
239 co-IPs and analyzed the entire molecular weight range with quantitative
240 proteomics; while abundant ribosomal proteins limited depth of detection of other
241 proteins, we observed enrichment of four LSU assembly factors: GTPBP4,
242 MRTO4, BRIX1 and NOP2 (Figure 3a and Supplementary Data 4). Co-IP
243 followed by western blotting confirmed the association of these four factors with
244 MINAS-60 (Figure 3b). The association of MINAS-60 with GTPBP4 and MRTO4
245 did not depend on RNA, because these associations largely survived RNaseA

246 treatment. In contrast, co-purification of MINAS-60 with BRIX1 and NOP2 was
247 severely diminished after treatment with RNaseA, suggesting that their
248 association with MINAS-60 was likely indirect and RNA-dependent (Figure 3b).
249 We therefore hypothesized that MINAS-60 associates with nucleolar, late pre-
250 60S particles containing GTPBP4 and MRTO4 to regulate ribosome
251 biogenesis[18,44].
252
253 To determine whether MINAS-60 associates with high molecular weight pre-60S
254 particles containing GTPBP4 and MRTO4, we performed sucrose gradient
255 fractionation of nuclear extracts of HEK 293T cells stably expressing epitope-
256 tagged MINAS-60, followed by western blotting analysis. As shown in Figure 3c,
257 a subpopulation of over-expressed MINAS-60 co-sedimented with ribosome
258 assembly factors in high-molecular weight fractions coincident with pre-ribosomal
259 particles. Combined with our co-IP and immunofluorescence data, these results
260 are consistent with a role for MINAS-60 in pre-ribosomal particles.
261
262 **MINAS-60 inhibits protein synthesis and cell proliferation**
263 We hypothesized that MINAS-60 could regulate ribosome biogenesis via its
264 association with GTPBP4 and MRTO4. Ribosome biogenesis is required for
265 protein synthesis, which promotes cell growth and proliferation. As a result,
266 ribosome biogenesis is commonly upregulated in cancer cells[21,45]. We reasoned
267 that, if MINAS-60 regulates ribosome biogenesis, its absence should result in
268 changes to protein synthesis and cell proliferation. In order to test this
269 hypothesis, we required a system to query the function of MINAS-60 independent
270 of RBM10, despite their co-encoding on the same transcript. To this end, we
271 knocked down (KD) *RBM10* in HEK 293T cells with two different shRNAs, which
272 silence the entire mRNA and, therefore, both proteins (RBM10 and MINAS-60).
273 To deconvolute phenotypic effects specific to MINAS-60 in the KD, as well as to
274 exclude off-target effects of shRNA, we generated rescue cell lines stably
275 expressing MINAS-60 (Rescue_MINAS-60) or RBM10 (Rescue_RBM10) on the
276 KD background. qRT-PCR and western blotting analysis revealed that *RBM10*

277 mRNA is efficiently silenced by both shRNAs, and the KD cells successfully re-
278 express MINAS-60 or RBM10 after rescue (Supplementary Figure 5). Noting that
279 the RBM10 rescue construct could be subject to leaky translation to produce both
280 RBM10 and MINAS-60, we also rescued the KD with an RBM10 construct
281 bearing an A_{398}TG to TAA mutation that eliminates the first MINAS-60 start
282 codon while preserving RBM10 translation (Supplementary Figure 5).
283
284 To test the effect of MINAS-60 expression on cellular protein synthesis, we
285 labeled nascent peptides in the control, RBM10 KD, Rescue_MINAS-60,
286 Rescue_RBM10 and Rescue_RBM10(A_{398}TG-TAA) cell lines with puromycin
287 followed by anti-puromycin western blotting[46]. As shown in Figure 3d and
288 Supplementary Figure 6a, depletion of the entire *RBM10* mRNA led to a
289 significant increase in global protein synthesis, and this increase was rescued by
290 reintroduction of MINAS-60. Partial rescue by RBM10 reintroduction was also
291 observed, but was not present in the RBM10(A_{398}TG-TAA) rescue cells. Similar
292 results were observed for a second shRNA targeting RBM10 (Supplementary
293 Figures 6b-c). Taken together, these results indicate that MINAS-60 plays an
294 unusual role in ribosome biogenesis to downregulate global protein synthesis. At
295 the same time, RBM10 is unlikely to play a role in this process. We speculate
296 that the partial rescue observed with RBM10 could have been due to leaky
297 expression of MINAS-60 from the internal alt-ORF in the RBM10 construct.
298
299 Because protein synthesis, cell growth and proliferation are linked[16,21], we asked
300 whether MINAS-60 regulates cell proliferation. As shown in Figure 3e, *RBM10*
301 depletion led to a significant increase in cell proliferation, consistent with a
302 published report using HeLa cells[29]. Remarkably, this increase was rescued by
303 reintroduction of MINAS-60 alone, and partially rescued by RBM10
304 reintroduction. Similar results were observed for a second shRNA targeting
305 RBM10 (Supplementary Figure 6d). These results are consistent with a model in
306 which MINAS-60 inhibits ribosome biogenesis, subsequently downregulating
307 cellular protein synthesis and cell proliferation.

308

309 **MINAS-60 inhibits the export of pre-60S ribosome subunits**

310 Eukaryotic ribosome biogenesis can be divided into sequential processes[47]: pre-
311 rRNA transcription; chemical modification and processing of the pre-rRNA, both
312 of which occur in the nucleolus; folding, assembly and maturation of the pre-
313 ribosomal subunits in the nucleolus and nucleus; and export and final maturation
314 of ribosome subunits in the cytoplasm. Because MINAS-60 co-purified with
315 proteins involved in various steps of LSU biogenesis, we wished to determine the
316 step in this process that it regulates.

317

318 To determine whether MINAS-60 controls pre-rRNA transcription, we performed
319 qRT-PCR targeting the primary pre-rRNA (47S/45S/30S) using previously
320 published primers[46] in lysates from control, RBM10 KD, Rescue_MINAS-60 and
321 Rescue_RBM10 cell lines. No significant differences in primary pre-rRNA were
322 observed, suggesting that MINAS-60 does not regulate pre-rRNA transcription
323 (Supplementary Figures 7a-b). To determine whether MINAS-60 regulates the
324 processing of LSU pre-rRNA, we performed northern blotting analysis with the
325 control, RBM10 KD, Rescue_MINAS-60 and Rescue_RBM10 cell lines using a
326 probe complementary to ITS2, which detects all LSU pre-rRNA processing
327 products, including 41S, 32S and 12S pre-rRNAs[46]. As shown in Supplementary
328 Figure 7c, no significant differences were observed between these cell lines,
329 suggesting that MINAS-60 does not regulate the processing of LSU pre-rRNA.

330

331 We therefore examined the role of MINAS-60 in LSU assembly and export. To
332 determine whether MINAS-60 regulates nucleocytoplasmic export of the pre-60S
333 ribosomal subunit, we quantified the ratio of nuclear vs. cytoplasmic RPL29-
334 GFP[48] stably expressed in the control, RBM10 KD, Rescue_MINAS-60 and
335 Rescue_RBM10 cell lines. As shown in Figure 4a and Supplementary Figure 8a,
336 depletion of RBM10 decreased the ratio of nuclear to cytoplasmic RPL29-GFP,
337 which can be rescued by reintroduction of MINAS-60, and partially rescued by
338 RBM10 reintroduction, likely due to leaky MINAS-60 translation. Similar results

339 were observed for a second shRNA targeting RBM10 (Figure 4a, right, and
340 Supplementary Figures 8b-c). These observations are consistent with the
341 inhibition of cytoplasmic export of pre-60S subunits by MINAS-60. As a control,
342 to determine whether MINAS-60 regulates nucleocytoplasmic export of the pre-
343 40S ribosomal subunit, we quantified the ratio of nuclear vs. cytoplasmic RPS2-
344 GFP[48] stably expressed in the control, RBM10 KD, Rescue_MINAS-60 and
345 Rescue_RBM10 cell lines. As shown in Supplementary Figure 9, no significant
346 changes were observed between these cell lines, suggesting MINAS-60 does not
347 regulate the assembly or export of 40S ribosomal subunits.
348
349 The observation of upregulated pre-60S export in the absence of MINAS-60
350 predicted that the same conditions should lead to an increase in mature
351 cytoplasmic 60S ribosomal subunits. To test this, we performed cytoplasmic
352 polysome profiling. As shown in Figures 4b-c, knockdown of RBM10 increased
353 the ratio of cytoplasmic 60S/40S ribosome subunits, and this increase can be
354 rescued by reintroduction of MINAS-60, and partially rescued by RBM10. Similar
355 results were observed for a second shRNA targeting RBM10 (Figure 4c, right
356 and Supplementary Figure 10). Taken together, these results suggest that
357 MINAS-60 specifically decreases mature large ribosomal subunits by negatively
358 regulating the assembly or export of pre-60S particles.
359
360 **MINAS-60 inhibits the late-stage assembly of pre-60S**
361 Lastly, we speculated that MINAS-60 functions as a checkpoint inhibitor in pre-
362 60S assembly prior to export from the nucleus, which would suggest that the
363 protein composition of LSU precursors should change in cells lacking MINAS-60.
364 MRTO4 and BRIX1 are LSU biogenesis factors present in multiple intermediate-
365 to-late, or early-to-intermediate, pre-60S particles, respectively, and we
366 examined their interactomes as a readout of changes in pre-60S protein
367 composition in the presence or absence of MINAS-60 (Figure 5a)[18]. We stably
368 expressed MRTO4-FLAG or BRIX1-FLAG, in control or RBM10 KD HEK 293T
369 cells to enable affinity purification of pre-ribosomal particles, followed by

370 quantitative proteomics and western blotting. We detected statistically significant
371 increases in several late LSU assembly factors co-purified with MRTO4 in
372 RBM10 KD cells (Figure 5b, Supplementary Data 5). These increases were
373 further confirmed by western blotting (Figure 5c). Similar results were observed
374 for BRIX1 co-IP, though fold changes were smaller (Figures 5d-f, Supplementary
375 Data 6). These results suggest that remodeling of pre-60S particles toward more
376 mature stages occurs in the absence of MINAS-60, consistent with MINAS-60
377 acting as an inhibitor for LSU assembly and export.
378

379 **Discussion**

380 In this work, we developed a bio-orthogonal strategy for the direct detection of
381 unannotated nascent or stress-induced alt-proteins. BONCAT has been
382 powerfully applied in prior studies to examine nascent protein synthesis,
383 particularly in neurons[30], but standard BONCAT workflows, like many proteomics
384 protocols, include column or gel resolution steps that de-enrich small proteins
385 and peptides, and are therefore refractory to detection of microproteins and alt-
386 proteins. For the first time in this work, we followed azidohomoalanine labeling
387 with small protein enrichment using a previously reported C8 column strategy[31],
388 coupled with direct on-bead capture to eliminate the need for column-based
389 removal of excess biotin probe molecules, which also remove small proteins.
390 Interfacing this BONCAT-based chemoproteomic pipeline with our platform
391 technology for unannotated microprotein and alt-protein detection, we identified
392 22 alt-proteins undergoing active synthesis in HEK 293T cells. For six selected
393 alt-proteins, we verified their translation, one of which, alt-CNPY2, is likely post-
394 transcriptionally upregulated by DNA damage stress. We furthermore propose
395 that BONCAT-mediated detection of MINAS-60 reflects the subpopulation of cells
396 that actively synthesize this protein during S phase of the cell cycle, consistent
397 with its coordinated synthesis with the ribosome biogenesis machinery. Taken
398 together, these results show that our method is able to detect unannotated alt-
399 proteins expressed during the cell cycle and cellular stress, and that these
400 proteins may play important roles in the cell. In the future, an important next step

401 will be to characterize the biological role of alt-CNPY2, and to expand this
402 strategy to additional cellular stress conditions.
403
404 We then conducted a focused functional study of nucleolar MINAS-60, loss of
405 which promotes the late-stage assembly of pre-60S and the export of pre-60S
406 particles into cytoplasm, leading to increases in cytoplasmic 60S ribosomal
407 subunits, global protein synthesis, and cell proliferation (Figure 6). The MINAS-
408 60 alt-protein is therefore a repressor of LSU biogenesis. In contrast to smORF-
409 encoded proteins, few alt-ORFs nested within canonical protein CDS have been
410 defined in molecular detail. This work, combined with previous literature[7,49],
411 expands the recent finding that a single human transcript can encode
412 overlapping, sequence-independent, yet functionally related proteins.
413 Importantly, the *RBM10* gene plays important roles in human physiology and
414 disease. RBM10 (RNA Binding Motif 10) is an RNA binding protein that regulates
415 alternative pre-mRNA splicing[29,50]. Null mutations in the *RBM10* gene are found
416 in patients with TARP syndrome[51], an X-linked inherited pathology associated
417 with malformation of multiple organs and significant early-life mortality. The
418 *RBM10* gene was also found to be among the most frequently mutated genes in
419 lung adenocarcinoma samples, and RBM10 inhibits cancer cell proliferation[52].
420 The finding that *RBM10* dually encodes MINAS-60, which also downregulates
421 cell proliferation via repression of ribosome biogenesis, opens the question of
422 whether this alt-protein contributes to *RBM10* mutation associated disease
423 phenotypes, analogous to alt-FUS, which forms pathogenic cytoplasmic
424 aggregates similar to those caused by the co-encoded FUS protein[3].
425
426 Several studies have suggested that as many as 50% of multicistronic human
427 genes encode proteins that directly interact to form complexes[9,53]. In contrast, the
428 alt-protein alt-RPL36 can regulate the same pathway (translation) as the
429 canonical protein co-encoded with it (RPL36), without directly interacting with
430 each other[8]. In this work, we show that MINAS-60 and RBM10 exhibit
431 comparable antiproliferative cellular effects but act via independent pathways:

432 ribosome biogenesis vs. alternative pre-mRNA splicing, respectively. While it
433 remains possible that RBM10 regulates splicing of some nucleolar proteins,
434 RBM10 reintroduction did not rescue the protein translation defect observed in
435 cells lacking *RBM10* expression in this study. Based on these results we
436 conclude that RBM10 does not significantly contribute to ribosome biogenesis.
437 We therefore speculate that selective pressure for co-encoding multiple
438 functional proteins in multicistronic human genes may act at the level of cellular
439 fitness, and direct interaction between alt-proteins and canonical proteins is only
440 one mechanism among many by which the fitness effects of co-encoded proteins
441 may be optimized. Another possible mechanism is that, since alt-ORFs may
442 represent protogenes[54], *de novo* acquisition of a novel alt-ORF within a protein
443 coding gene might lead to positive selection if both the alt-ORF and canonical
444 protein coding sequence increase cellular fitness under the same conditions,
445 without necessarily acting on the same pathway or process. This broadens the
446 current models of multicistronic human gene function[9].
447
448 To our knowledge, MINAS-60 is only the second human microprotein or alt-
449 protein validated to localize to the nucleolus to date, and the first reported to
450 regulate ribosome biogenesis. A previous study identified nucleolar microprotein
451 C11orf98, which interacts with nucleolar proteins NPM1 and NCL, but no role for
452 this factor in ribosome biogenesis has been investigated[33]. In this work, we
453 reported MINAS-60 as the first protein that specifically inhibits LSU biogenesis,
454 and we suspect that it functions as a mammal-specific checkpoint to ensure
455 correct pre-60S assembly prior to nuclear export. Supporting the hypothesis that
456 MINAS-60 has arisen in higher eukaryotes, the *RBM10* gene is not conserved in
457 yeast, though we cannot yet rigorously conclude that a functional homolog of
458 MINAS-60 is absent in lower eukaryotes. Direct experimental evidence will be
459 required to test the hypothesis that MINAS-60 could represent a control point for
460 ribosome biogenesis that is unique to mammalian cells.
461

462 Interestingly, another mammal-specific, putative LSU biogenesis inhibitor has
463 recently been identified. A structural study revealed that an unidentified protein
464 (protein X) is positioned to block the incorporation of nuclear export factor NMD3
465 into the late pre-60S particle, hypothetically suppressing pre-60S export into the
466 cytoplasm[44]. Protein X binds in the immature peptidyl transfer center and directly
467 contacts GTPBP4 as well as helix 89 of the rRNA. Importantly, protein X was
468 specifically observed in cryo-electron microscopy structures of human pre-60S
469 particles and has not been detected in structures of yeast LSU precursors[44].
470 Based on the observations that 1) both MINAS-60 and protein X likely function as
471 repressors of LSU biogenesis at the late assembly and export stage; and 2) both
472 MINAS-60 and protein X associate with GTPBP4, we speculate that MINAS-60
473 functions by a similar mechanism to protein X. It is possible that protein X and
474 MINAS-60 are the same protein, and that protein X could not be identified due to
475 its absence from the UniProt database; however, efforts to model MINAS-60
476 sequences into the reported structures remain inconclusive to date. It is therefore
477 also possible that multiple previously unknown mammalian proteins regulate LSU
478 biogenesis, suggesting that regulation of ribosome biogenesis may be more
479 complex in human cells than in yeast.
480
481 Why, then, might MINAS-60 have arisen in mammalian cells? In order to achieve
482 a balance between cellular growth requirements and energy-intensive ribosome
483 production[15], and to ensure that improperly assembled ribosome subunits do not
484 lead to mistranslation[23], ribosome biogenesis needs to be precisely monitored in
485 cells. For example, the Rio1-Nob1-Pno1 network establishes a checkpoint to
486 safeguard against the release of immature 40S subunits into translating
487 ribosomes[55]. However, while hundreds of proteins required for ribosome
488 biogenesis in human cells have now been identified[56], few repressors of
489 ribosome biogenesis have been reported, and the ones that have been identified
490 predominately act on pre-ribosomal RNA transcription[57-59]. LSU biogenesis
491 involves multiple composition and conformational changes during generation of
492 the peptidyl transferase center (PTC). Our data lead us to speculate that MINAS-

493 60 and/or Protein X may act as checkpoints for this process to safeguard cellular
494 energy expenditure and/or faithful PTC formation. This hypothesis could be
495 tested by examining the effects of MINAS-60 on cellular fitness under nutrient
496 stress, as well as on the fidelity and accuracy of global protein translation.
497
498 In conclusion, this study, combined with previous literature[13], demonstrates the
499 power of chemoproteomics to reveal alt-proteins that are regulated and/or
500 functional in important cellular processes including DNA damage and ribosome
501 biogenesis, and reveals an entirely new regulatory node of the ribosome
502 biogenesis pathway, meriting further development of chemical tools to enrich and
503 identify functional alt-proteins.
504
505
506

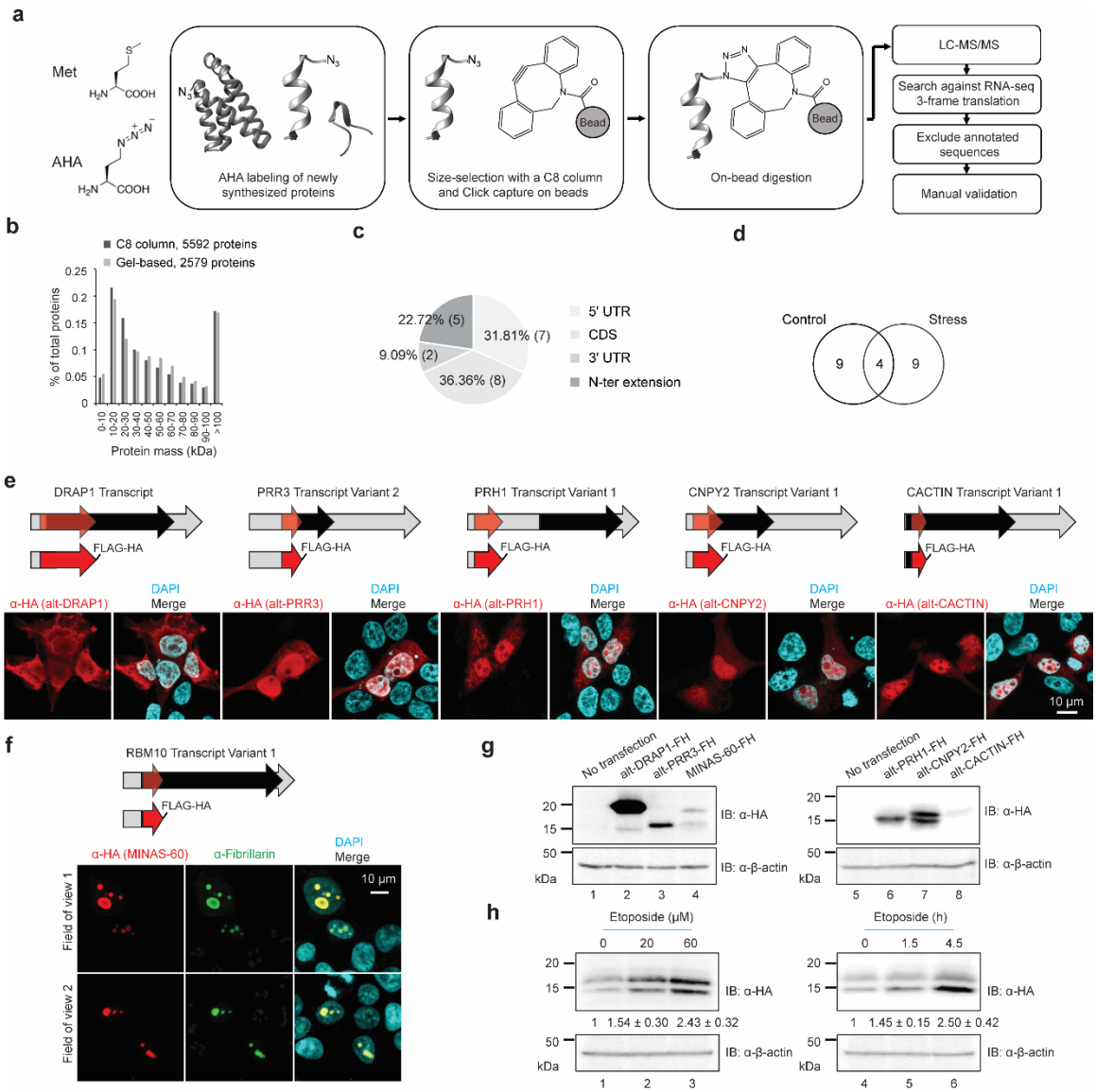


Figure 1. BONCAT-based chemoproteomic identification of newly synthesized alt-proteins. a Schematic workflow for BONCAT-based chemoproteomic analysis of nascent alt-proteins. **b** Size distribution of canonical, annotated proteins detected using C8 column-based selection (black) and our previously reported[60] gel-based selection (gray). **c** Distribution of locations of identified alt-proteins relative to the annotated coding sequence (CDS). **d** Venn diagram of alt-proteins identified under control or stress conditions. **e, f** Top: a schematic representation of human *DRAP1* transcript, *PRR3* transcript variant 2, *PRH1* transcript variant 1, *CNPY2* transcript variant 1, *CACTIN* transcript variant 1, or *RBM10* transcript variant 1 (**c**); light gray arrow, 5' and 3' untranslated

518 regions (UTR); red, alternative open reading frame (alt-ORF) coding sequence;
519 black, annotated coding sequence. Middle: a schematic representation of the
520 expression constructs containing the complete 5'UTR and the alt-ORF of the
521 transcript indicated above, with a dual FLAG and HA tag appended to the C-
522 terminus of the alt-protein. Bottom: HEK 293T cells transfected with the
523 expression construct indicated (middle) were immunostained with anti-HA (red),
524 DAPI (cyan), and anti-fibrillarin (green, **c**). Scale bar, 10 μm. Data are
525 representative of three biological replicates. **g** HEK 293T cells transfected with
526 the expression construct (middle) were immunoblotted (IB) with antibodies
527 indicated to the right, with untransfected (no transfection) HEK 293T cells as a
528 control. Data are representative of three biological replicates. **h** HEK 293T cells
529 transfected with the alt-CNPY2 expression construct were treated with increasing
530 amounts of etoposide or vehicle for 2 h (left), or with 60 μM etoposide for
531 different times or vehicle (right), followed by western blotting with antibodies
532 indicated to the right. Quantitative analysis (N = 3) of the western blot signal of
533 alt-CNPY2-FLAG-HA are indicated at the bottom.
534

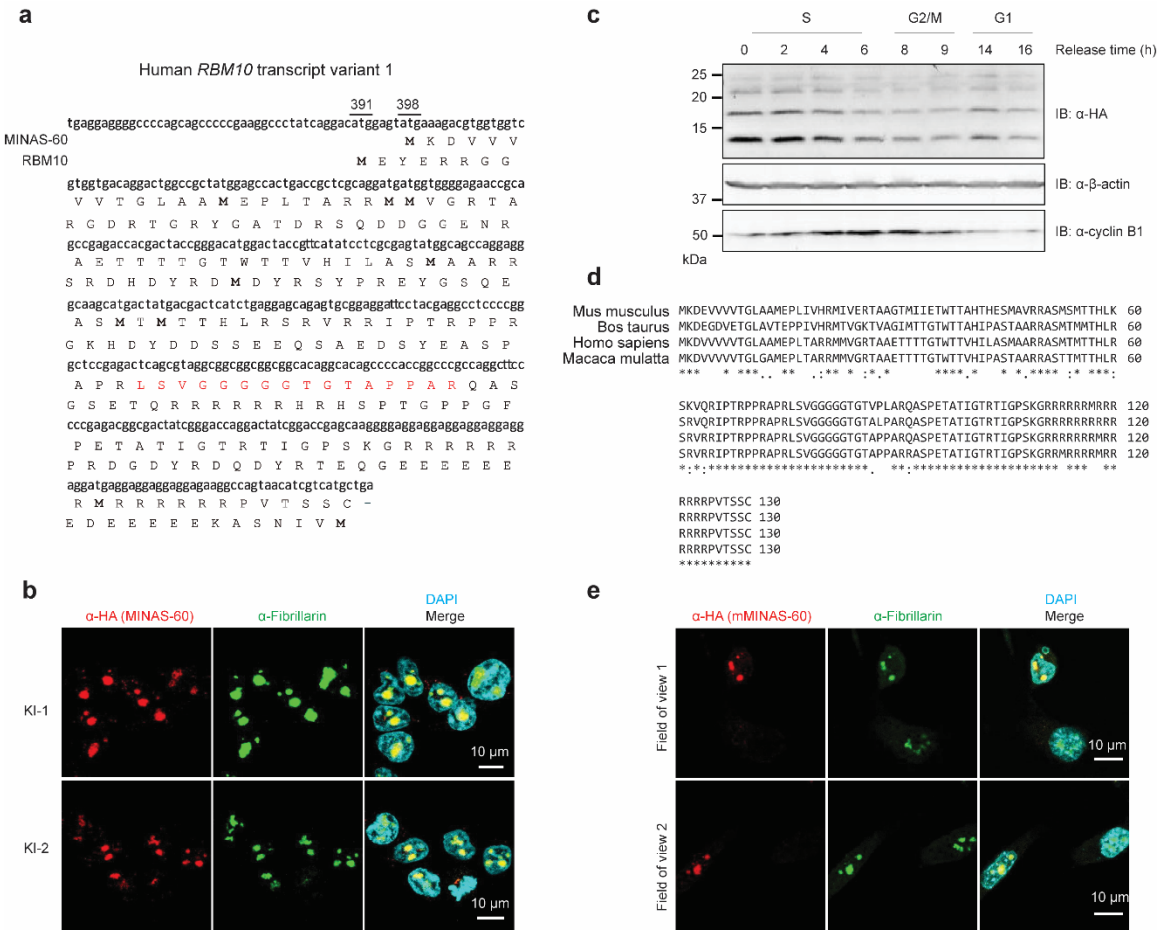
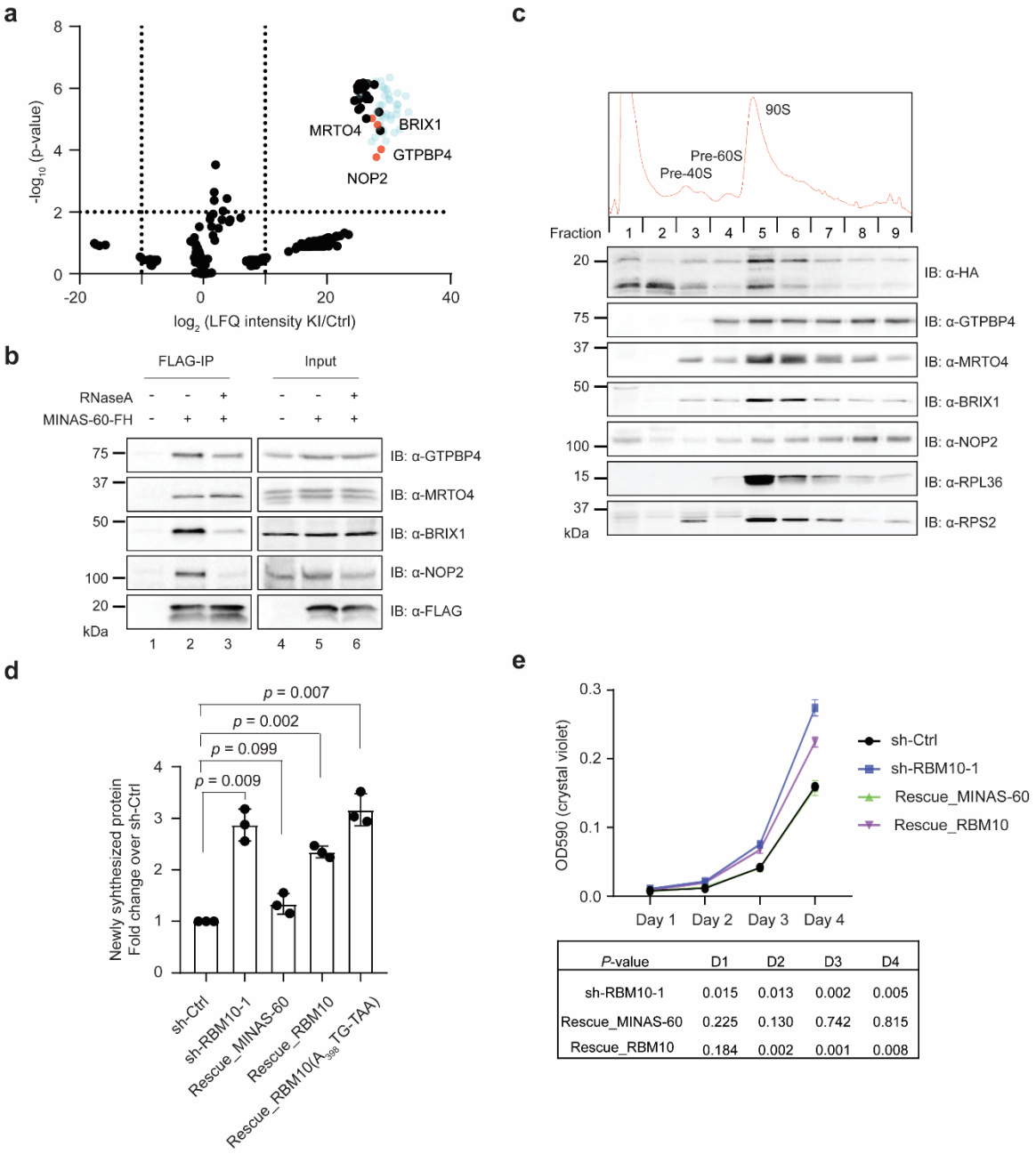


Figure 2. MINAS-60 is endogenously expressed, cell-cycle regulated and conserved in mouse. a Shown is the cDNA sequence of human *RBM10* transcript variant 1 with the protein sequences of MINAS-60 and RBM10 indicated below. The start codons of RBM10 (391) and MINAS-60 (398) are numbered above the cDNA sequence. Highlighted in red is the tryptic peptide of MINAS-60 detected by LC-MS/MS in this study. **b** Immunostaining of two MINAS-60 knock-in (KI) HEK 293T cell lines with anti-HA (red), anti-fibrillarin (green), and DAPI (cyan). Scale bar, 10 μm. Data are representative of three biological replicates. **c** 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. Data are representative of three biological replicates. **d** RBM10 mRNAs from *Mus musculus* (NM_145627.3), *Bos taurus* (NM_001206693.1), and *Macaca mulatta* (XM_015127275.2) were obtained from NCBI nucleotide database, then translated in the +1, +2 and +3 frames using the

550 Expasy translate tool. Cognate start codons in-frame with sequences
551 homologous to human MINAS-60 were identified in the 5'UTR of each transcript
552 in order to predict the full-length sequence of hypothetical MINAS-60 homologs.
553 Shown is the ClustalW2 alignment of predicted MINAS-60 protein sequences. **e**
554 3T3 cells transfected with the mouse MINAS-60 (mMINAS-60) expression
555 construct were immunostained with anti-HA (red), anti-fibrillarin (green), and
556 DAPI (cyan). Scale bar, 10 μm. Data are representative of three biological
557 replicates.
558



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

Figure 3. MINAS-60 associates with nucleolar LSU assembly factors, downregulates global protein synthesis and cell proliferation. 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 4. **b** HEK 293T cells were transfected with MINAS-60-FLAG-HA (MINAS-60-FH,

567 lanes 2 and 3) or vehicle (lane 1), and immunoprecipitation (FLAG-IP) was
568 performed in absence (lanes 1 and 2) or presence (lane 3) of RNaseA, followed
569 with immunoblotting (IB) with antibodies indicated on the right. Cell lysates (4%)
570 before IP (input, lanes 4-6) were used as loading controls. **c** Top: Sucrose-
571 gradient sedimentation analysis of nuclear lysates containing ribosome precursor
572 complexes (pre-40S, pre-60S and 90S pre-ribosome) from HEK 293T cells stably
573 expressing MINAS-60-FLAG-HA. Bottom: Western blot analysis of fractions
574 numbered at the top with antibodies indicated on the right. **d** ImageJ was used to
575 quantify the relative puromycin incorporation for cells indicated at the bottom
576 relative to sh-Ctrl from three biological replicates. Data represent mean values ±
577 s.e.m., and significance was evaluated with two-tailed *t*-test. **e** Growth curve of
578 control (sh-Ctrl), *RBM10* knockdown (sh-RBM10-1), rescue with MINAS-60
579 (Rescue_MINAS-60) and rescue with RBM10 (Rescue_RBM10) HEK 293T cells
580 at the indicated number of days (N = 3). Data represent mean values ± s.e.m.,
581 and significance was evaluated with two-tailed *t*-test and shown below.
582

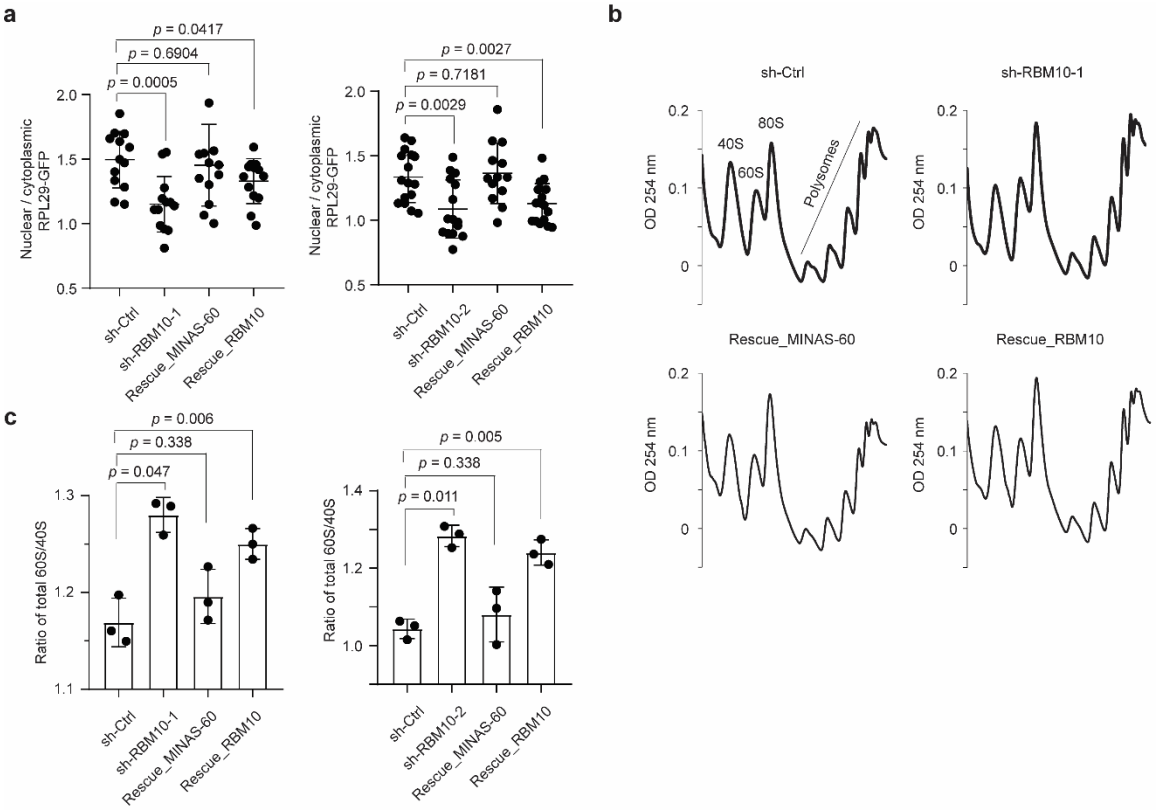


Figure 4. MINAS-60 inhibits LSU export. a Quantitation of the ratio of RPL29-GFP intensity in the nucleus vs. cytoplasm in control (sh-Ctrl), *RBM10* knockdown with one of the two shRNAs (sh-RBM10-1 (left), sh-RBM10-2 (right)), rescue with MINAS-60 (Rescue_MINAS-60), or rescue with RBM10 (Rescue_RBM10) HEK 293T cells stably expressing RPL29-GFP. 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. **b** 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. Data are representative of three biological replicates. **c** Quantitation of the ratio of cytoplasmic 60S to 40S subunits in the cell lines indicated below after sucrose gradient fractionation. The area under each peak was measured using ImageJ following a previously published method[61]. Data represent mean values ± s.e.m., and significance was evaluated with two-tailed *t*-test.

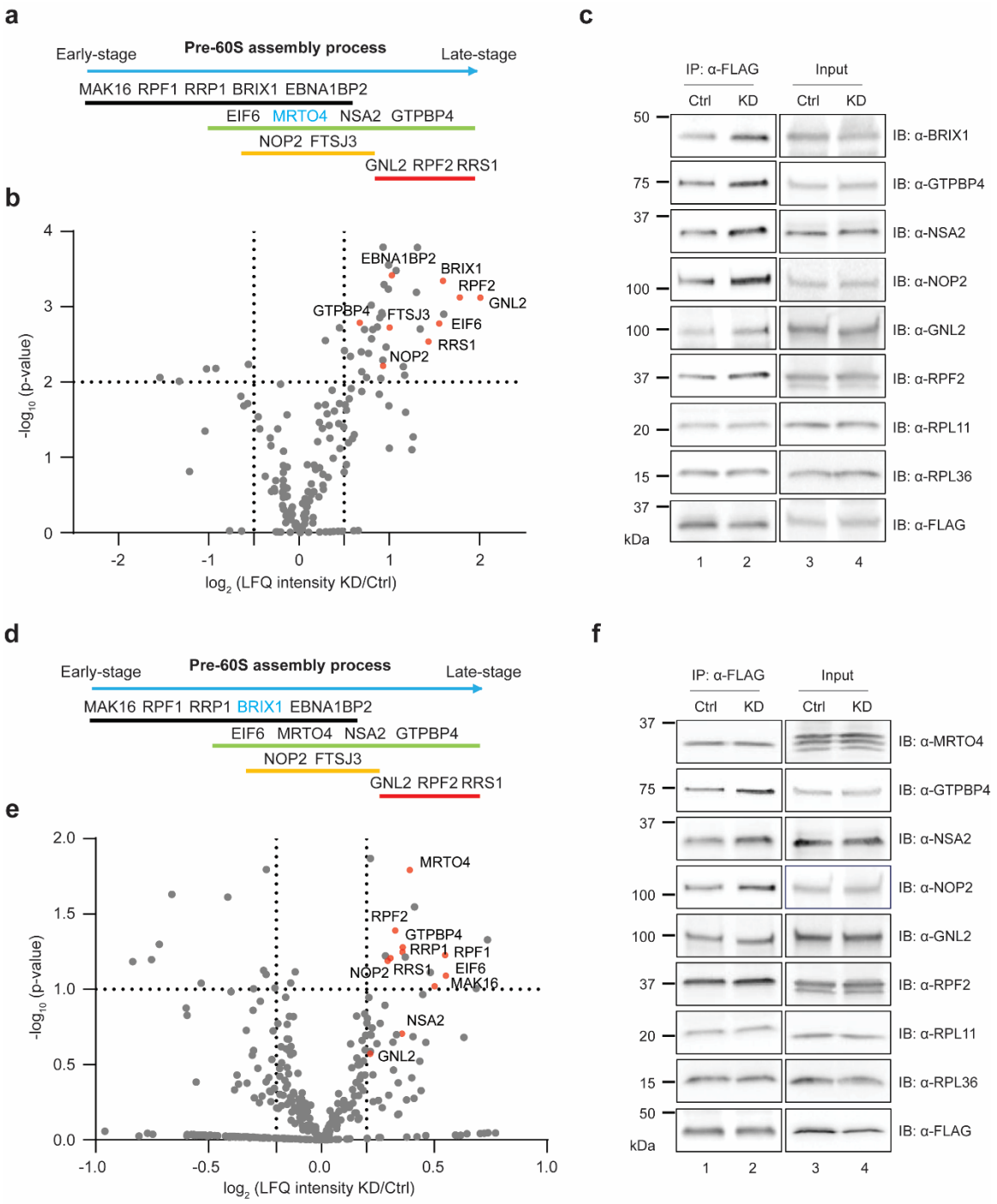


Figure 5. *RBM10* silencing promotes pre-60S assembly. a, d Schematic representation of pre-60S assembly factors associated with different states based on a pre-60S structure report[18]. The bait protein MRTO4 (**a**) and BRIX1 (**d**) is indicated in blue. **b, e** Volcano plot of quantitative proteomics (*N* = 4 (**b**), *N* = 5 (**e**)) of MRTO4-FLAG (**b**) or BRIX1-FLAG (**e**) pulldown from HEK 293T cells

606 stably expressing the bait protein and control shRNA (Ctrl), or the bait protein
607 and *RBM10* shRNA (KD), to quantify relative changes in the bait protein co-IP of
608 LSU assembly factors in *RBM10* KD over control HEK 293T cells. Increased
609 assembly factors are indicated in red and gene names are labeled. For complete
610 quantitative proteomics results, see Supplementary Datas 5 and 6. **c, f** MRTO4
611 FLAG-IP (**c**) or BRIX1-FLAG-IP (**f**) and western blotting with antibodies indicated
612 on the right using the two cell lines described above. Cell lysates (4%) before IP
613 (input) were used as the loading control. Data are representative of three
614 biological replicates.
615

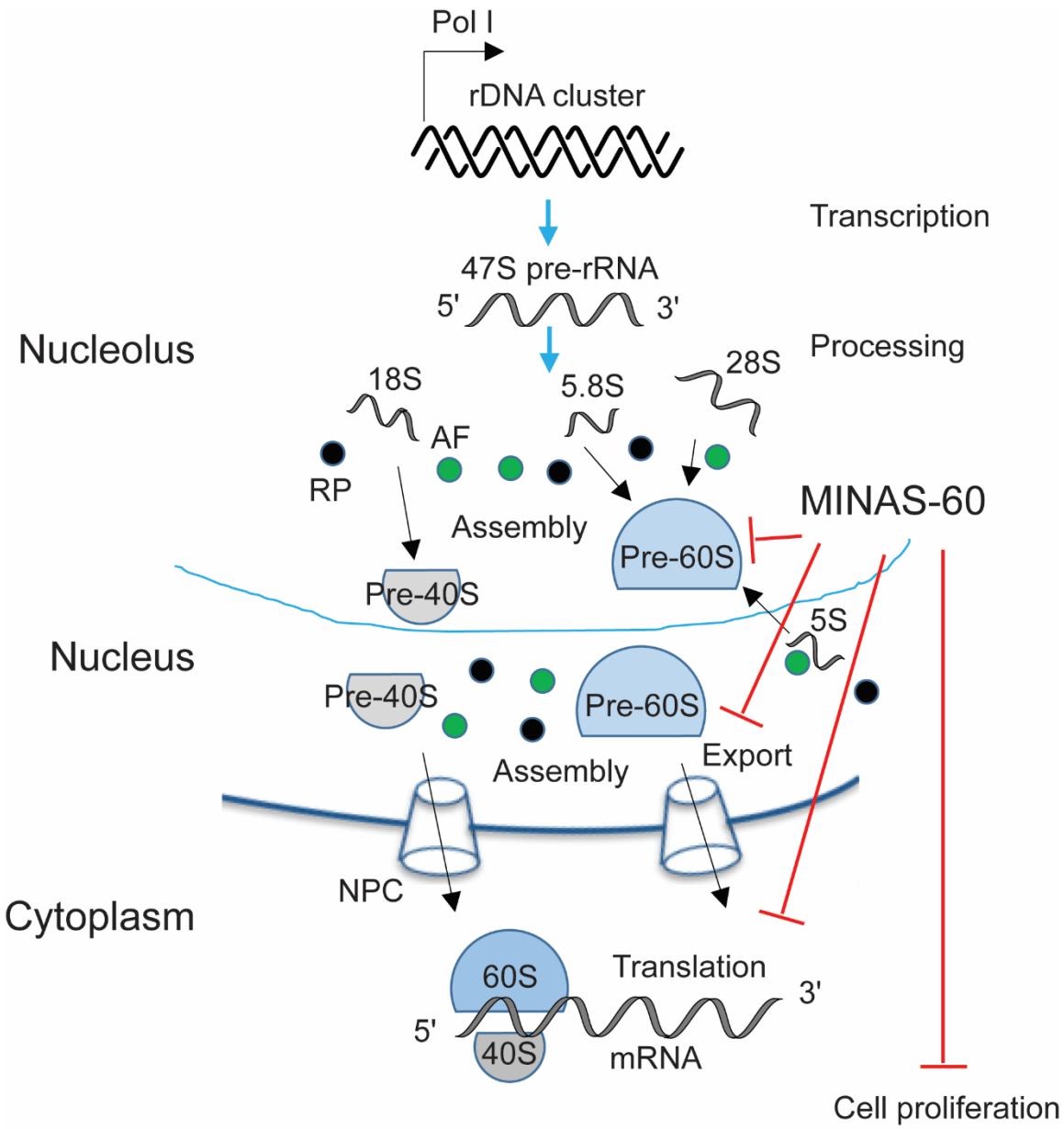


Figure 6. Model of MINAS-60 regulatory pathway. MINAS-60 localizes to the nucleolus, where it associates with multiple pre-60S assembly factors, including GTPBP4 and MRTO4, to inhibit the late-stage pre-60S assembly and the export of pre-60S into cytoplasm, consequently downregulating global protein synthesis and cell proliferation.

623 **Online methods**

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

627 **Antibodies.** Primary antibodies for western blotting include the following: anti-
628 FLAG (Sigma, F3165 or Cell Signaling, 14793); anti-HA (Invitrogen, 71-5500);
629 anti-β-actin (Invitrogen, BA3R); anti-GTPBP4 (Abclonal, A4565); anti-MRTO4
630 (ThermoFisher, 20194-1-AP); anti-NSA2 (Abclonal, A14475); anti-NOP2 (Cell
631 Signaling, 25017); anti-BRIX1 (Abclonal, A14481); anti-RPF2 (Abclonal,
632 A17224); anti-GNL2 (Abclonal, A13191); anti-RPL11 (Cell Signaling, 18163);
633 anti-RPL36 (Bethyl Laboratories, A305065A-M); anti-RPS2 (Invitrogen, PA5-
634 30160); anti-RBM10 (Abcam, ab72423); anti-puromycin (Kerafast, EQ0001); anti-
635 GFP (Abcam, ab183734); anti-cyclin B1 (Cell Signaling, 4138).
636 Immunoprecipitation was performed with anti-FLAG M2 affinity gel (Sigma,
637 A2220). Secondary antibodies for western blotting are goat anti-rabbit IgG
638 horseradish peroxidase conjugate (Rockland, 611-1302) and goat anti-mouse
639 IgG horseradish peroxidase conjugate (Rockland, 610-1319-0500). Primary
640 antibodies for immunostaining are rabbit anti-HA (Invitrogen, 71-5500) and
641 mouse anti-Fibrillarin (abcam, ab4566). Secondary antibodies for
642 immunostaining are goat anti-rabbit IgG Alexa fluor 568 (Invitrogen, A11011) and
643 goat anti-mouse IgG Alexa fluor 647 (Invitrogen, A21235).
644

645 **Cloning and genetic constructs.** A DNA sequence comprising the full 5'UTR of
646 human *DRAP1* transcript, *PRR3* transcript variant 2, *PRH1* transcript variant 1,
647 *CNPY2* transcript variant 1, *CACTIN* transcript variant 1, or *RBM10* transcript
648 variant 1 through the stop codon of the relative alt-protein was amplified by PCR
649 with a dual FLAG and HA epitope tag appended to the 3' end of the alt-protein
650 coding sequence from an in-house library of reverse-transcribed HEK 293T
651 cDNAs, then subcloned into pcDNA3. Deletion or mutation constructs of MINAS-
652 60 bearing a dual FLAG and HA tag were generated by ligating PCR products

653 into BamHI and EcoRI cloning sites in the pcDNA3 vector. For generation of HEK
654 293T cells stably expressing MINAS-60, a dual FLAG and HA tag were
655 appended to the 3' end of MINAS-60 by PCR, and the dually tagged coding
656 sequence was then cloned into pLJM1. The cDNA clone expressing RPS2 was
657 purchased from Addgene (a gift from Thomas Tuschl), and the coding sequences
658 of MRTO4, BRIX1 and RPL29 were amplified by PCR from an in-house HEK
659 293T cDNA library, then subcloned into pJLM1 for producing lentivirus. RNAi
660 constructs were made by synthesizing oligonucleotides encoding a 21 bp short
661 hairpin RNA that targets RBM10 (shRNA1, CTTCGCCTTCGTCGAGTTTAG;
662 shRNA2, TCCAACGTGCGCGTCATAAAG), then subcloned into pLKO.1. The
663 empty pLKO.1 vector control was purchased from Sigma (SHC001). qPCR
664 primer sequences are provided in Supplementary Table 2.
665
666 **Cell culture, lentivirus production and stable cell line generation.** HEK 293T
667 cells were purchased from ATCC and early-passage stocks were established in
668 order to ensure cell line identity. Cells were maintained up to only 10 passages.
669 HEK 293T cells were cultured as previously described[8]. To produce lentivirus
670 and generate stable cell lines, HEK 293T cells were co-transfected using
671 polyethyleneimine (Polysciences, 23966) with expression construct in pLJM1,
672 along with pMD2.G and psPAX2, and growth media were replaced after 7-8 h. 48
673 h post-transfection, media containing viruses was harvested, filtered through a
674 0.45-μm filter, and infection was performed by mixing with two volumes of fresh
675 media containing suspended HEK 293T cells. 24 h post-infection, the growth
676 media was replaced. 48 h post-infection, stably expressing cells were selected
677 with 4 μg/mL puromycin for 2 days. Early stocks of stable cell lines were
678 established after selection. Stable cell lines were released from puromycin for 2
679 days prior to use in experiments.
680
681 **Immunostaining and live-cell imaging.** HEK 293T cells were plated on glass
682 coverslips and transfected the next day. Forty-eight hours later, the cells were
683 fixed in 10% formalin for 15 min at room temperature (RT), permeabilized with

684 PBS containing 0.2% (v/v) TritonX-100, then incubated with primary antibodies
685 for 18 h at 4℃. After washing with PBS, the cells were incubated with secondary
686 antibodies and DAPI for 1 h at RT, washed with PBS and mounted with Mowiol
687 (Sigma, 81381) before viewing.
688
689 HEK 293T cells stably expressing RPL29-GFP or RPS2-GFP were grown to 70%
690 confluency on coverslips in 12-well plates. Coverslips were inverted and imaged
691 in pre-warmed DMEM with 10% FBS, 1% penicillin-streptomycin in MatTek
692 imaging dishes. Confocal imaging was performed on a Leica SP8 LS confocal
693 microscope with 63× oil immersion objective with atmosphere-controlled stage at
694 37℃. Nuclear/cytoplasmic ratios of RPL29 and RPS2 were measured using the
695 ImageJ Intensity Ratio Nuclei Cytoplasm Tool (RRID:SCR_018573;
696 https://github.com/MontpellierRessourcesImagerie/imagej_macros_and_scripts/w
697 iki/Intensity-Ratio-Nuclei-Cytoplasm-Tool).
698
699 **Immunoprecipitation and proteomics.** Control HEK 293T cells or MINAS-60 KI
700 cells were grown to 80-90% confluency in 15 cm dishes. Cells were harvested
701 and suspended in 1 mL nuclear isolation buffer (10 mM HEPES-KOH pH 7.4,
702 100 mM KCl, 5 mM $MgCl_2$ with 0.5% NP40 and Roche Complete protease
703 inhibitor cocktail tablets (Roche, 11873580001)), and incubated on ice for 10 min,
704 followed by centrifugation at 3,000 g, 4℃, 3 min. The nuclear pellets were
705 suspended in 1 mL lysis buffer (Tris-buffered saline (TBS) with 1% Triton X-100
706 and Roche Complete protease inhibitor cocktail tablets), followed with sonication
707 and immunoprecipitation as previously described[8]. After the final wash, elution
708 was in 40 µL of 3× FLAG peptide (Sigma, F4799), at a final concentration of 100
709 µg/mL in lysis buffer at 4℃ for 1 h. The eluted proteins were subjected to SDS-
710 PAGE separation prior to LC-MS/MS analysis.
711
712 Gel slices, containing either resolved protein bands or entire lanes, were
713 digested with trypsin at 37℃ for 14-16 h. The resulting peptide mixtures were
714 extracted from the gel, dried, subjected to ethyl acetate extraction to remove

715 residual detergent, de-salted with peptide cleanup C18 spin column (Agilent
716 Technologies, 5188-2750), then re-suspended in 35 µL 0.1% formic acid (FA),
717 followed by centrifugation at 21,130 g, 4℃, 30 min. A 5 µL aliquot of each
718 sample was injected onto a pre-packed column attached to a nanoAcquity UPLC
719 (Waters) in-line with a Thermo Scientific™ Q Exactive™ Plus Hybrid
720 QuadrupoleOrbitrap™ mass spectrometer (Thermo Scientific) and a 130-min
721 gradient was used to further separate the peptide mixtures as follows (solvent A:
722 0.1% FA; solvent B: acetonitrile (ACN) with 0.1% FA): Isocratic flow was
723 maintained at 0.1 µL/min at 1% B for 40 min, followed by linear gradients from
724 1% B to 6% B over 2 min, 6% B to 24% B over 48 min, 24% B to 48% B over 5
725 min, 48% B to 80% B over 5 min. Isocratic flow at 80% B was maintained for 5
726 min, followed by a gradient from 80% B to 1% B over 5 min, and isocratic flow at
727 1% B was maintained for 10 min. The full MS was collected over the mass range
728 of 300-1,700 m/z with a resolution of 70,000 and the automatic gain control
729 (AGC) target was set as 3 x 10^6. MS/MS data was collected using a top 10 high-
730 collisional energy dissociation method in data-dependent mode with a normalized
731 collision energy of 27.0 eV and a 1.6 m/z isolation window. MS/MS resolution
732 was 17,500 and dynamic exclusion was 90 seconds.
733
734 For identification of alt- and microproteins, ProteoWizard MS Convert was used
735 for peak picking and files were analyzed using Mascot. Oxidation of methionine
736 and N-terminal acetylation were set as variable modifications, and a previously
737 reported[12] three-frame translation of assembled transcripts from HEK 293T
738 mRNA-seq was used as the database. For co-IP proteomics searches and
739 quantitative analysis, files were analyzed using MaxQuant, oxidation of
740 methionine and N-terminal acetylation were set as variable modifications, and
741 human UniProt plus MINAS-60 was used as the database for searching. For all
742 analysis, a mass deviation of 20 p.p.m. was set for MS1 peaks, and 0.6 Da was
743 set as maximum allowed MS/MS peaks with a maximum of two missed
744 cleavages. Maximum false discovery rates (FDR) were set to 1% both on peptide
745 and protein levels. Minimum required peptide length was five amino acids.

746 Protein quantitation was accomplished by calculating the LFQ intensity ratio of KI
747 or KD pulldown to negative control samples using MaxQuant (version 1.6.8.0)
748 with standard parameters.
749
750 **Dibenzocyclooctyne (DBCO) bead construction.** NHS-activated beads (Pierce,
751 88826) were conjugated to dibenzocyclooctyne-amine (Sigma, 761540) by 90 min
752 incubation at room temperature with rotation in a saturated 100 mM sodium
753 bicarbonate (pH 8.0) solution. The beads were then washed and blocked for 2 h
754 with 3 M ethanolamine according to the manufacturer's instructions. After blocking,
755 the beads were re-suspended in PBS prior to immediate protein conjugation.
756
757 **BONCAT (bio-orthogonal non-canonical amino acid tagging).** HEK 293T
758 cells were grown to 80-90% confluency in 15 cm dishes, treated with the
759 methionine aminopeptidase inhibitor TNP470 (50 nM) for 2 h, then immersed in
760 methionine-free DMEM (Corning, 17-204-CI) for 30 min before addition of 4 mM
761 AHA (Click Chemistry tools, 1066-1000) in methionine-free DMEM with 10%
762 FBS. For stress conditions, cells were treated with 4 mM AHA with simultaneous
763 exposure to 500 μM sodium arsenite, 20 μM etoposide (Sigma, E1383), or 1 mM
764 DTT. After a 2 h incubation, the cells were washed twice with cold PBS,
765 harvested, and flash frozen until further processing.
766
767 **DBCO bead enrichment of AHA-labeled proteins and ERLIC (electrostatic**
768 **repulsion hydrophilic interaction chromatography) fractionation.** AHA-
769 labeled cells were lysed by boiling in 50 mM HCl with 0.01% 2-mercaptoethanol
770 and 0.05% TritonX-100 for 10 min. The cells were then pelleted at 21,100 g, 4℃
771 for 30 min, and passed through a 5 μm filter before size selection with a C8
772 column (Agilent Technologies, 12102100) essentially as previously reported[32].
773 The C8 column was pre-conditioned with 1 bed volume of methanol and 2 bed
774 volumes of 0.25 M triethylammonium formate (TEAF, pH 3.0). Then up to 2 bed
775 volumes of cell lysate were loaded on the column, followed by twice washes with
776 2 bed volumes of TEAF and elution with 2 bed volumes of ACN:TEAF (1:3). The

777 sample was then dried and reconstituted in 300 µL PBS. Cell lysates were
778 incubated with DBCO beads for 1 h at RT. The beads were washed twice with 1
779 mL RIPA buffer, once with 1 M KCl, once with 0.1 M sodium carbonate, once
780 with 2 M urea, twice with RIPA buffer, and finally 6 times with PBS. Proteins
781 covalently conjugated to the beads were then subjected to reduction, alkylation
782 and on-bead trypsin digestion according to standard protocols[62].
783
784 Before LC-MS/MS, the digested peptides were fractionated using ERLIC on an
785 Agilent 1100 HPLC. Peptides were re-suspended in 55 µL of 85% ACN/0.1% FA
786 and 50 µL was loaded onto a polyWAX LP column (150×1.0 mm; 5 µm 300 Å;
787 PolyLC). Samples were run on an 80 min gradient protocol as follows (Solvent A:
788 80% ACN 0.1% FA; Solvent B: 30% ACN 0.1% FA): Isocratic flow was
789 maintained at 100% A at a flow rate of 0.3 mL/min for 5 min, followed by a 17 min
790 linear gradient to 8% B, and a 25 min linear gradient to 45% B. Finally, a 10 min
791 gradient to 100% B was followed by a 5 min hold at 100% B before a 10 min
792 linear gradient back to 100% A, followed by an 8 min hold at 100% A. The
793 digested peptides were separated into 12-15 fractions which were dried and re-
794 suspended in 7 µL of 3:8 70% FA : 0.1% TFA before LC-MS/MS analysis.
795
796 **Generation of knock-in cell lines.** MINAS-60 3xGFP11-FLAG-HA KI HEK 293T
797 cells were generated using CRISPR-Cas9. Guide RNAs (gRNAs) were designed
798 with the guide design tool from the Zhang lab (crispr.mit.edu) to target the
799 RBM10 genomic region gRNA1, 5'- TGTCGGCCAGGATTCCTACG-3'; gRNA2,
800 5'- CCCGATAGTCGCCGTCTCGG-3'. Double-stranded DNA oligonucleotides
801 corresponding to the gRNAs were inserted into pSpCas9(BB)-2A-GFP vector
802 (Addgene, as a gift from F. Zhang, MIT, Cambridge, MA). A donor plasmid
803 containing 300 bp homology left-arm and 300 bp homology right-arm sequence
804 around the stop codon of MINAS-60, which are separated with 3xGFP11-FLAG-
805 HA tag and BamHI / NotI restriction sites was synthesized by GenScript, and a
806 DNA sequence containing pGK promoter and hygromycin resistance gene were
807 subcloned into the donor plasmid using the BamH1 and NotI restriction sites. An

808 equal mixture of the gRNA and donor plasmids were transfected into HEK 293T
809 cells using polyethyleneimine, and hygromycin selection was performed 2 days
810 post-transfection. MINAS-60-3xGFP11-FLAG-HA KI cells were confirmed by
811 genomic DNA PCR and sequencing.
812
813 **Puromycin incorporation.** SUnSET was used to measure protein synthesis[63].
814 Briefly, HEK 293T cells were grown to 80-90% confluency in 6 well plates, then
815 growth media was replaced with media containing 1 μM puromycin and cells
816 were incubated at 37℃ in a humidified atmosphere with 5% CO_2 for 1 h
817 according to a previously published protocol[56]. The cells were then washed once
818 with PBS, harvested and analyzed with western blotting.
819
820 **Sucrose gradient profiling.** HEK 293T cells were seeded in 15 cm dishes at
821 2.5×10^7 cells per dish and cultured 24 h. Cells were then treated with 100 μg/mL
822 cycloheximide for 5 min, washed with cold PBS containing 100 μg/mL
823 cycloheximide twice, harvested and flash frozen until further processing.
824
825 For cytoplasmic polysome profiling, after thawing on ice, cells were lysed in
826 polysome lysis buffer (20 mM HEPES-KOH pH 7.4, 100 mM KCl, 5 mM $MgCl_2$,
827 100 μg/mL cycloheximide, 1 mM DTT with 1% TritonX-100, Roche Complete
828 protease inhibitor cocktail tablets and Ribonuclease Inhibitors (Promega N2511)),
829 and incubated on ice for 10 min, followed by centrifugation at 21,130 g, 4℃, 10
830 min. The supernatants were normalized according to absorbance (A260) and
831 layered onto 12 mL 10-50% sucrose gradients (20 mM HEPES-KOH pH 7.4, 100
832 mM KCl, 5 mM $MgCl_2$, 100 μg/mL cycloheximide, 1 mM DTT with Roche
833 Complete protease inhibitor cocktail tablets and Ribonuclease Inhibitors),
834 followed with centrifugation in an SW-41Ti rotor at 252,878 g, 4℃, 3 h, then
835 sampled using a Biocomp Gradient profiler (model 251) with constant monitoring
836 of optical density at 254 nm using standard parameters. Data analysis was
837 performed with Excel.
838

839 For sucrose gradient sedimentation analysis of nuclear lysates, thawed cells
840 were suspended in 1 mL nuclear isolation buffer (10 mM HEPES-KOH pH 7.4,
841 100 mM KCl, 5 mM MgCl2 with 0.5% NP40, Roche Complete protease inhibitor
842 cocktail tablets and Ribonuclease Inhibitors), and incubated on ice for 10 min,
843 followed by centrifugation at 3,000 g, 4℃, 3 min. The nuclear pellets were
844 suspended in 1 mL nuclear lysis buffer (20 mM HEPES-KOH pH 7.4, 300 mM
845 KCl, 5 mM MgCl2, 1 mM DTT with 1% TritonX-100, Roche Complete protease
846 inhibitor cocktail tablets and Ribonuclease Inhibitors), followed with sonication
847 (50% intensity, 5 s pulse with 25 s rest, 5×, MICROSON XL 2000), and
848 centrifugation at 21,130 g, 4℃, 10 min. The supernatants were normalized by
849 A260 and layered onto 12 mL 10-50% sucrose gradients, centrifuged and
850 sampled as described above.
851
852 **Crystal violet staining and cell cycle synchronization.** 4×10^4 HEK 293T cells
853 were seeded in 12-well plates in triplicate, and fixed in 10% formalin for 15 min at
854 RT every 24 h. After washing with ddH_2O, cells were stained with 0.1% crystal
855 violet in methanol for 30 min at RT in the dark, followed with three ddH_2O
856 washes and dried. The cells were then immersed in 1 mL 10% acetic acid with
857 shaking for 20 min. 20 μL of the solution was combined with 80 μL ddH_2O in a
858 96-well plate, and the optical density at 590 nm was monitored with Synergy™
859 HT.
860
861 For cell cycle synchronization, 1.5×10^5 HEK 293T cells were seeded in 12-well
862 plates and cultured overnight. Cells were then treated with 2 mM thymidine for 16
863 h, followed with two PBS washes, and incubated with fresh media for 9 h before
864 the second 2 mM thymidine block for 14 h following a previously published
865 protocol[64]. Cells were washed with PBS, then incubated with fresh media to
866 release from the G1/S boundary, and fixed in 10% formalin for immunostaining or
867 harvested for western blotting.
868
869 **Northern blot**

870 Total RNA was extracted from HEK 293T cells using TRIzol reagent. To
871 determine changes in levels of LSU pre-rRNA intermediates, 3 µg of total RNA
872 was run on a 1% agarose/1.25% formaldehyde gel in a 1.5 M tricine/ 1.5 M
873 triethanolamine buffer, followed by an overnight transfer to a Hybond XL nylon
874 membrane (GE Healthcare, RPN 303S) in 10× saline-sodium citrate transfer
875 buffer after a brief 15 min soak in a 0.5 M sodium hydroxide solution. Membranes
876 were then exposed to UV (254 nm) to immobilize the RNA, followed by
877 incubation with denatured yeast tRNA for 1 h at 42℃, and hybridized overnight at
878 37℃ with 5' end radiolabeled oligonucleotide probe (P4 5'-
879 CGGGAACTCGGCCCGAGCCGGCTCTCTCTTTCCCTCTCCG-3') in a solution
880 of 7.5× Denhardt's solution, 5× sodium chloride-sodium phosphate-EDTA buffer
881 with 0.1% SDS, as previously reported[59]. Membranes were also hybridized with a
882 7SL probe (7SL 5'-TGCTCCGTTTCCGACCTGGGCCGGTTCACCCCTCCTT-3')
883 as a loading control.
884
885 **Data availability**
886 Proteomic data were deposited under accession PXD026880. During review,
887 they can be accessed with username reviewer_pxd026880@ebi.ac.uk and
888 password 7dkNTlia.
889
890 **Acknowledgements**
891 We thank Franziska Bleichert and all members of the Slavoff and Baserga labs
892 for helpful conversations. This work was supported by a Searle Scholars
893 Program Award, an Odyssey Award from the Richard and Susan Smith Family
894 Foundation, and start-up funds from Yale University West Campus (to S. A. S.).
895 X.C. was supported in part by a Rudolph J. Anderson postdoctoral fellowship
896 from Yale University. A.K. was in part supported by an NIH Predoctoral Training
897 Grant (5T32GM06754 3-12). S.J.B, C.J.B. and C.M.H. were supported by R35
898 GM131687. C.M.H. was supported by an NSF GFRP.
899
900 **Author contributions**

X.C., A.K., C.M.H., C.J.B., and S.Z. and 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. Slavoff, S.A. et al. Peptidomic discovery of short open reading frame-encoded peptides in human cells. *Nat Chem Biol* **9**, 59-64 (2013).
13. Schwaid, A.G. et al. Chemoproteomic discovery of cysteine-containing human short open reading frames. *J Am Chem Soc* **135**, 16750-3 (2013).
14. Klinge, S. & Woolford, J.L., Jr. Ribosome assembly coming into focus. *Nat Rev Mol Cell Biol* **20**, 116-131 (2019).
15. Pena, C., Hurt, E. & Panse, V.G. Eukaryotic ribosome assembly, transport and quality control. *Nat Struct Mol Biol* **24**, 689-699 (2017).
16. De Keersmaecker, K., Sulima, S.O. & Dinman, J.D. Ribosomopathies and the paradox of cellular hypo- to hyperproliferation. *Blood* **125**, 1377-82 (2015).
17. Nerurkar, P. et al. Eukaryotic Ribosome Assembly and Nuclear Export. *Int Rev Cell Mol Biol* **319**, 107-40 (2015).
18. Kater, L. et al. Visualizing the Assembly Pathway of Nucleolar Pre-60S Ribosomes. *Cell* **171**, 1599-1610 e14 (2017).

945 19. Sanghai, Z.A. et al. Modular assembly of the nucleolar pre-60S ribosomal subunit.
946 *Nature* **556**, 126-129 (2018).
947 20. Micic, J. et al. Coupling of 5S RNP rotation with maturation of functional centers during
948 large ribosomal subunit assembly. *Nat Commun* **11**, 3751 (2020).
949 21. Pelletier, J., Thomas, G. & Volarevic, S. Ribosome biogenesis in cancer: new players and
950 therapeutic avenues. *Nat Rev Cancer* **18**, 51-63 (2018).
951 22. Nyhus, C., Pihl, M., Hyttel, P. & Hall, V.J. Evidence for nucleolar dysfunction in
952 Alzheimer's disease. *Rev Neurosci* **30**, 685-700 (2019).
953 23. Farley-Barnes, K.I., Ogawa, L.M. & Baserga, S.J. Ribosomopathies: Old Concepts, New
954 Controversies. *Trends Genet* **35**, 754-767 (2019).
955 24. Mills, E.W. & Green, R. Ribosomopathies: There's strength in numbers. *Science*
956 **358**(2017).
957 25. Hernandez, J. et al. Tumor suppressor properties of the splicing regulatory factor
958 RBM10. *RNA Biol* **13**, 466-72 (2016).
959 26. Inoue, A. et al. RBM10 regulates alternative splicing. *FEBS Lett* **588**, 942-7 (2014).
960 27. Rodor, J., FitzPatrick, D.R., Eyras, E. & Caceres, J.F. The RNA-binding landscape of RBM10
961 and its role in alternative splicing regulation in models of mouse early development.
962 *RNA Biol* **14**, 45-57 (2017).
963 28. Sutherland, L.C. et al. Splicing arrays reveal novel RBM10 targets, including SMN2 pre-
964 mRNA. *BMC Mol Biol* **18**, 19 (2017).
965 29. Bechara, E.G., Sebestyen, E., Bernardis, I., Eyras, E. & Valcarcel, J. RBM5, 6, and 10
966 differentially regulate NUMB alternative splicing to control cancer cell proliferation. *Mol*
967 *Cell* **52**, 720-33 (2013).
968 30. Dieterich, D.C., Link, A.J., Graumann, J., Tirrell, D.A. & Schuman, E.M. Selective
969 identification of newly synthesized proteins in mammalian cells using bioorthogonal
970 noncanonical amino acid tagging (BONCAT). *Proc Natl Acad Sci U S A* **103**, 9482-7 (2006).
971 31. Ma, J. et al. Improved Identification and Analysis of Small Open Reading Frame Encoded
972 Polypeptides. *Anal Chem* **88**, 3967-75 (2016).
973 32. Khitun, A. & Slavoff, S.A. Proteomic Detection and Validation of Translated Small Open
974 Reading Frames. *Curr Protoc Chem Biol* **11**, e77 (2019).
975 33. Chu, Q. et al. Identification of Microprotein-Protein Interactions via APEX Tagging.
976 *Biochemistry* **56**, 3299-3306 (2017).
977 34. Wang, T., Liu, Y., Liu, Q., Cummins, S. & Zhao, M. Integrative proteomic analysis reveals
978 potential high-frequency alternative open reading frame-encoded peptides in human
979 colorectal cancer. *Life Sci* **215**, 182-189 (2018).
980 35. Kamiyama, D. et al. Versatile protein tagging in cells with split fluorescent protein. *Nat*
981 *Commun* **7**, 11046 (2016).
982 36. Iyer-Bierhoff, A. & Grummt, I. Stop-and-Go: Dynamics of Nucleolar Transcription During
983 the Cell Cycle. *Epigenet Insights* **12**, 2516865719849090 (2019).
984 37. Andrews, J.O. et al. qSR: a quantitative super-resolution analysis tool reveals the cell-
985 cycle dependent organization of RNA Polymerase I in live human cells. *Sci Rep* **8**, 7424
986 (2018).
987 38. Gorski, S.A., Snyder, S.K., John, S., Grummt, I. & Misteli, T. Modulation of RNA
988 polymerase assembly dynamics in transcriptional regulation. *Mol Cell* **30**, 486-97 (2008).
989 39. Hernandez-Verdun, D. Assembly and disassembly of the nucleolus during the cell cycle.
990 *Nucleus* **2**, 189-94 (2011).
991 40. Cao, X. & Slavoff, S.A. Non-AUG start codons: Expanding and regulating the small and
992 alternative ORFeome. *Exp Cell Res* **391**, 111973 (2020).

993 41. Tan, L.Y. et al. Generation of functionally distinct isoforms of PTBP3 by alternative
994 splicing and translation initiation. *Nucleic Acids Res* **43**, 5586-600 (2015).
995 42. Saghatelian, A. & Couso, J.P. Discovery and characterization of smORF-encoded
996 bioactive polypeptides. *Nat Chem Biol* **11**, 909-16 (2015).
997 43. Cox, J. et al. Accurate proteome-wide label-free quantification by delayed normalization
998 and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**, 2513-26
999 (2014).
1000 44. Liang, X. et al. Structural snapshots of human pre-60S ribosomal particles before and
1001 after nuclear export. *Nat Commun* **11**, 3542 (2020).
1002 45. Liu, Y., Deisenroth, C. & Zhang, Y. RP-MDM2-p53 Pathway: Linking Ribosomal Biogenesis
1003 and Tumor Surveillance. *Trends Cancer* **2**, 191-204 (2016).
1004 46. Sondalle, S.B., Longerich, S., Ogawa, L.M., Sung, P. & Baserga, S.J. Fanconi anemia
1005 protein FANCI functions in ribosome biogenesis. *Proc Natl Acad Sci U S A* **116**, 2561-2570
1006 (2019).
1007 47. Bassler, J. & Hurt, E. Eukaryotic Ribosome Assembly. *Annu Rev Biochem* **88**, 281-306
1008 (2019).
1009 48. Wild, T. et al. A protein inventory of human ribosome biogenesis reveals an essential
1010 function of exportin 5 in 60S subunit export. *PLoS Biol* **8**, e1000522 (2010).
1011 49. Klemke, M., Kehlenbach, R.H. & Huttner, W.B. Two overlapping reading frames in a
1012 single exon encode interacting proteins--a novel way of gene usage. *EMBO J* **20**, 3849-60
1013 (2001).
1014 50. Inoue, A. RBM10: Structure, functions, and associated diseases. *Gene* **783**, 145463
1015 (2021).
1016 51. Niceta, M. et al. TARP syndrome: Long-term survival, anatomic patterns of congenital
1017 heart defects, differential diagnosis and pathogenetic considerations. *Eur J Med Genet*
1018 **62**, 103534 (2019).
1019 52. Imielinski, M. et al. Mapping the hallmarks of lung adenocarcinoma with massively
1020 parallel sequencing. *Cell* **150**, 1107-20 (2012).
1021 53. Gagnon, M. et al. Potentiation of B2 receptor signaling by AltB2R, a newly identified
1022 alternative protein encoded in the human bradykinin B2 receptor gene. *J Biol Chem* **296**,
1023 100329 (2021).
1024 54. Carvunis, A.R. et al. Proto-genes and de novo gene birth. *Nature* **487**, 370-4 (2012).
1025 55. Parker, M.D., Collins, J.C., Korona, B., Ghalei, H. & Karbstein, K. A kinase-dependent
1026 checkpoint prevents escape of immature ribosomes into the translating pool. *PLoS Biol*
1027 **17**, e3000329 (2019).
1028 56. Farley-Barnes, K.I. et al. Diverse Regulators of Human Ribosome Biogenesis Discovered
1029 by Changes in Nucleolar Number. *Cell Rep* **22**, 1923-1934 (2018).
1030 57. Nguyen le, X.T., Raval, A., Garcia, J.S. & Mitchell, B.S. Regulation of ribosomal gene
1031 expression in cancer. *J Cell Physiol* **230**, 1181-8 (2015).
1032 58. Xing, Y.H. et al. SLERT Regulates DDX21 Rings Associated with Pol I Transcription. *Cell*
1033 **169**, 664-678 e16 (2017).
1034 59. Ogawa, L.M. et al. Increased numbers of nucleoli in a genome-wide RNAi screen reveal
1035 proteins that link the cell cycle to RNA polymerase I transcription. *Mol Biol Cell* **32**, 956-
1036 973 (2021).
1037 60. Cao, X. et al. Comparative Proteomic Profiling of Unannotated Microproteins and
1038 Alternative Proteins in Human Cell Lines. *J Proteome Res* **19**, 3418-3426 (2020).
1039 61. Gregory, B. et al. The small and large ribosomal subunits depend on each other for
1040 stability and accumulation. *Life Sci Alliance* **2**(2019).

1041 62. Weerapana, E., Speers, A.E. & Cravatt, B.F. Tandem orthogonal proteolysis-activity-
1042 based protein profiling (TOP-ABPP)--a general method for mapping sites of probe
1043 modification in proteomes. *Nat Protoc* **2**, 1414-25 (2007).
1044 63. Schmidt, E.K., Clavarino, G., Ceppi, M. & Pierre, P. SUnSET, a nonradioactive method to
1045 monitor protein synthesis. *Nat Methods* **6**, 275-7 (2009).
1046 64. Cao, X. et al. Histone H4K20 Demethylation by Two hHR23 Proteins. *Cell Rep* **30**, 4152-
1047 4164 e6 (2020).

1048