

1 **Live-cell mapping of organelle-associated RNAs via proximity biotinylation** 2 **combined with protein-RNA crosslinking**

3 **Authors/Affiliations**

4 Pornchai Kaewsapsak^{1,5}, David M. Shechner^{2-3,5}, William Mallard²⁻³, John L. Rinn²⁻³,
5 and Alice Y. Ting^{1,4,6*}

6 ¹Department of Chemistry, Massachusetts Institute of Technology, Cambridge,
7 Massachusetts, USA;

8 ²Department of Stem Cell and Regenerative Biology, and Molecular and Cellular
9 Biology, Harvard University, Cambridge, Massachusetts, USA

10 ³Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge,
11 Massachusetts, USA

12 ⁴Departments of Genetics, Biology, and Chemistry, Stanford University, Stanford, CA,
13 USA

14 ⁵Co-first authors

15 ⁶Lead Contact

16 *Correspondence: ayting@stanford.edu

17

18 **Keywords**

19 RNA localization, Subcellular Transcriptomics, Sequencing Technologies,
20 Peroxidase, APEX2, Horseradish Peroxidase, FISH, RNA-Seq, Mammalian Cell RNA
21 Analysis

22

23 **Abstract**

24 The spatial organization of RNA within cells is a crucial factor in a wide range of
25 biological functions, spanning all kingdoms of life. However, a general
26 understanding of RNA localization has been hindered by a lack of simple, high-
27 throughput methods for mapping the transcriptomes of subcellular compartments.
28 Here, we develop such a method, termed APEX-RIP, which combines peroxidase-
29 catalyzed, spatially restricted in situ protein biotinylation with RNA-protein
30 chemical crosslinking. We demonstrate that, using a single protocol, APEX-RIP can
31 isolate RNAs from a variety of subcellular compartments, including the
32 mitochondrial matrix, nucleus, bulk cytosol, and endoplasmic reticulum (ER), with
33 higher specificity and coverage than do conventional approaches. We furthermore
34 identify candidate RNAs localized to mitochondria-ER junctions and nuclear lamina,
35 two compartments that are recalcitrant to classical biochemical purification. Since
36 APEX-RIP is simple, versatile, and does not require special instrumentation, we
37 envision its broad application in a variety of biological contexts.

38

39 **Introduction**

40 Spatial compartmentalization of RNA is central to many biological processes across
41 all kingdoms of life, and enables diverse regulatory schemes that exploit both coding
42 and noncoding functions of the transcriptome. For example, the localization and
43 spatially restricted translation of mRNA plays a fundamental role in asymmetric cell
44 division in bacteria and yeast, body-pattern formation in *Drosophila* and *Xenopus*,
45 signaling at mammalian neuronal synapses (Jung et al. 2014), and a wide variety of

46 other biological contexts. In another example, the localization of noncoding RNAs
47 (ncRNAs) can play an architectural role in the assembly of subcellular structures,
48 including short-range chromatin loops, higher-order chromatin domains, and large
49 sub-nuclear structures like nucleoli and Barr bodies (Rinn and Guttman 2014;
50 Engreitz, Ollikainen, and Guttman 2016). However, despite these examples, our
51 general understanding of the breadth and biological significance of RNA subcellular
52 localization remains incomplete.

53 Techniques that elucidate the subcellular localization of RNAs are therefore
54 critical for advancing our understanding of RNA biology. Classically, such techniques
55 rely either on imaging or biochemical approaches. Imaging methods—such as
56 Fluorescence *In Situ* Hybridization (FISH) and artificial RNA reporter schemes—are
57 powerful tools for elucidating the positions of a small number of target RNAs at low-
58 to-moderate throughput (Wilk et al. 2016; K. H. Chen et al. 2015; Paige, Wu, and
59 Jaffrey 2011; Hocine et al. 2013; Nelles et al. 2016). Alternatively, unbiased
60 approaches for RNA discovery couple biochemical manipulations to deep
61 sequencing. For example, the RNA partners of proteins with characteristic
62 subcellular localization can be identified through a variety of techniques that couple
63 protein immunoprecipitation to RNA-seq (Ule et al. 2003; Christopher Gilbert et al.
64 2004). Such methods have revealed many localized mRNAs, in addition to novel
65 non-coding RNAs involved in RNA splicing (Chi et al. 2009) and RNAi (Motamedi et
66 al. 2004). On a broader scale, a deep sampling of RNAs residing within a cellular
67 compartment—for example, an intact organelle of interest, or partitions along a
68 sucrose gradient, can be identified by coupling subcellular fractionation to RNA-Seq
69 (“Fractionation-Seq”) (Sterne-Weiler et al. 2013; Mercer et al. 2011).

70 However, a technological gap exists among these current methods for
71 studying RNA localization. Imaging approaches are of limited throughput, and may
72 require specialized reagents, constructs, or microscopes that are only accessible to a
73 handful of laboratories (Wilk et al. 2016; K. H. Chen et al. 2015; Paige, Wu, and
74 Jaffrey 2011; Hocine et al. 2013; Nelles et al. 2016). The efficacy of
75 immunoprecipitation-based approaches is highly sensitive to the antibodies and
76 enrichment protocols used (Hendrickson et al. 2016) and captures only RNAs that
77 are directly complexed with each target protein. Fractionation-Seq is applicable only
78 to organelles and subcellular fractions that can be purified, and is frequently
79 complicated by contaminants (false positives) and loss of material (false negatives)
80 Therefore, a new technology is needed for unbiased and large-scale discovery and
81 characterization of RNA *neighborhoods*, with high spatial specificity, and within
82 cellular structures that cannot be enriched by biochemical fractionation.

83 Here we introduce such a technology—termed APEX-RIP— that enables
84 unbiased discovery of endogenous RNAs in specific cellular locales. APEX-RIP
85 merges two existing technologies: APEX (engineered ascorbate peroxidase)-
86 catalyzed proximity biotinylation of endogenous proteins (Rhee et al. 2013), and
87 RNA ImmunoPrecipitation (RIP) (Christopher Gilbert et al. 2004). We demonstrate
88 that APEX-RIP is able to enrich endogenous RNAs in membrane-enclosed cellular
89 organelles, such as the mitochondrion and nucleus, and in membrane-abutting
90 cellular regions such as the cytosolic face of the endoplasmic reticulum. The
91 specificity and coverage of this approach are much higher than those obtained by

92 traditional Fractionation-Seq. Moreover, by applying APEX-RIP to multiple
93 mammalian organelles, we have generated high quality datasets of
94 compartmentalized RNAs that should serve as valuable resources for testing and
95 generating novel hypotheses pertinent to RNA biology.

97 **Development of APEX-RIP method and application to mitochondria**

98 APEX is an engineered peroxidase that can be targeted by genetic fusion to various
99 subcellular regions of interest (Rhee et al. 2013) (Figure 1A). Upon addition of its
100 substrates, biotin-phenol (BP) and hydrogen peroxide (H_2O_2), to live cells, APEX
101 catalyzes the formation of biotin-phenoxy radicals that then diffuse outward and
102 covalently biotinylate nearby endogenous proteins. More distal proteins are not
103 significantly labeled because the biotin-phenoxy radical has a half-life of less than 1
104 millisecond (Wishart and Madhava Rao 2010). Previous work has shown that APEX-
105 catalyzed proximity biotinylation, coupled to streptavidin enrichment and mass
106 spectrometry, can generate proteomic maps of the mitochondrial matrix,
107 intermembrane space, outer membrane, and nucleoid, each with <5 nm spatial
108 specificity (Rhee et al. 2013; Hung et al. 2014; Hung et al. 2017; Han et al. 2017).

109 Because most cellular RNAs exist in close proximity to proteins, we reasoned
110 that APEX-tagged subcellular proteomes could also provide access to the nearby
111 RNA content, if proteins and RNA could be crosslinked together *in situ*, immediately
112 before or after APEX labeling. As our first target organelle, we selected the
113 mitochondrion because its RNA content--derived from both the mitochondrial
114 genome and from imported, nuclear-encoded RNAs--has been extensively
115 characterized by a wide array of complementary methods (Mercer et al. 2011; Alán
116 et al. 2010; Piechota et al. 2006; Ro et al. 2013), hence providing a “gold-standard”
117 to which we can compare our results. The mitochondrial matrix was also the first
118 mammalian compartment mapped by APEX proteomics methodology (Rhee et al.
119 2013). As a RNA-protein chemical crosslinker, we opted for mild formaldehyde
120 treatment, which covalently captures most protein-protein and protein-nucleic acid
121 interactions, and can be achieved with minimal disruption of native interactions in
122 live cells. It is for these reasons that formaldehyde is used for several RIP (Chris
123 Gilbert and Svejstrup 2006) technologies for identifying the RNA partners of specific
124 proteins of interest, including our own “fRIP-Seq” protocol (Hendrickson et al. 2016).

125 Since it was unclear *a priori* whether APEX-catalyzed biotinylation should
126 precede or follow the formaldehyde crosslinking step, we explored both schemes in
127 parallel (Figure S1A; see methods). Each protocol, applied to HEK 293T cells
128 expressing mitochondrially-localized APEX2 (“mito-APEX2”, Figures 1B-C), resulted
129 in clear enrichment of fifteen mitochondrial-encoded RNAs—relative to the
130 cytosolic marker *GAPDH*—as gauged by RT-qPCR (average of 49.3 ± 3.5 and
131 60.9 ± 4.1 -fold enrichment, respectively, Figure S1A). Assuming that fixing cells prior
132 to biotinylation would better capture transient or weak RNA-protein interactions,
133 we selected the crosslinking-then-BP protocol for RNA-Seq analysis. While this
134 confirmed that mitochondrial mRNAs were enriched, a sizeable “shoulder” of
135 conspicuous off-target RNAs were also unexpectedly enriched (Figure S1B). Thus,
136 we re-examined our labeling and crosslinking protocols, using a sampling of these
137 off-target RNA markers (e.g., the abundant nuclear RNA *XIST*, and cytosol-localized

138 RNAs *HOOK2* and *MAN2C1*). This more comprehensive analysis revealed that APEX
139 labeling followed by crosslinking provides superior specificity (Figure S1C). We
140 hypothesize that the mild formaldehyde treatment compromises membrane
141 integrity (Fox et al. 1985), allowing BP radicals to escape to adjoining compartments
142 when APEX labeling is performed after formaldehyde treatment.

143 We used the optimized APEX followed by crosslinking protocol to map
144 mitochondrial RNAs in mito-APEX2-expressing HEK 293T cells (Figure 1D, Table 1,
145 tab 2). Gene-level analysis, comparing RNA counts before and after streptavidin
146 enrichment, revealed that all 13 mRNAs encoded by the mitochondrial genome were
147 highly enriched (greater than 3.5 fold) in three independent replicates (Figures 1D
148 and S1E, Table 1 tab 1). Enrichment was absent in negative controls with H₂O₂
149 omitted (Figure S1F). Read density plots mapped to the mitochondrial genome
150 demonstrated that most of our captured RNAs correspond to fully-processed
151 transcripts, including mRNAs, interstitial tRNAs, and the D-loop leader sequence
152 from which mitochondrial transcription initiates (Figure 1E). Intriguingly, mito-
153 mRNA read densities appeared to correlate with previous measures of mRNA half-
154 life (Nagao, Hino-Shigi, and Suzuki 2008). For example, mRNAs encoding MTCO1-3
155 have longer half-lives, and more reads from APEX-RIP, than mRNAs encoding
156 MTND1-2.

157

158 **APEX-RIP mapping of the nuclear-cytoplasmic RNA distribution**

159 Having established that APEX-RIP is both specific and sensitive in the
160 mitochondrion, we next turned our attention to a more challenging compartment:
161 the mammalian nucleus. The nucleus is more complex and has a less defined
162 transcriptome than the mitochondrial matrix, but previous Fractionation-Seq
163 datasets, including by ENCODE (Dunham et al. 2012) (Figures S2A–D), again provide
164 a reference list to which we can compare our results.

165 We generated HEK 293T cells that stably express APEX2 in the nucleus
166 (APEX-NLS) or in the cytosol (APEX-NES; NES is a nuclear export signal). The
167 specificity of *in situ* biotinylation by these constructs within each compartment was
168 confirmed by imaging (Figure 2A). Whole cell lysates prepared from each cell line
169 also produced distinct “fingerprints” of biotinylated proteins, as assayed by
170 streptavidin blotting (Figure S1D).

171 We performed APEX-RIP on both APEX-NLS and APEX-NES cells, using the
172 biotinylation-first/crosslinking-second protocol established above, with an
173 additional one-minute radical-quenching step in between the APEX and crosslinking
174 steps (Figure S3A; see *methods*). Encouragingly, “gold standard” nuclear and
175 cytosolic RNAs (defined from the ENCODE data as the top 1000 RNAs in each
176 compartment; see Table 2 tab 4) were enriched from the corresponding cell lines as
177 predicted (Figure 2B *histograms* and Figures S2E–F). Moreover, when directly
178 comparing the fold-enrichments from each compartment to one another, it was
179 apparent that APEX-NLS had effectively enriched known nuclear-localized RNAs,
180 while APEX-NES had enriched known cytosol-localized RNAs (Figure 2B *scatter plot*,
181 Table 2 tab 3). We calculated for each RNA a “nuclear preference score,” defined as
182 the minimum geometric distance of each point to the line $y=x$ (corresponding to the
183 set of genes which are not preferentially enriched from either compartment).

184 Receiver Operator Characteristic (ROC) analysis of these nuclear preference scores
185 was used to filter the data and obtain final transcript lists of 5,467 nuclear RNAs and
186 10,130 cytosolic RNAs from living HEK 293T cells (Table 2 tabs 1 and 2). The false
187 discovery rates of these two lists are <0.6% and <0.4%, respectively.

188 When plotted by nuclear preference score, the human transcriptome
189 displayed an overall bimodal distribution, wherein the majority of species were
190 cytoplasmic, appended by a smaller right-shifted populace of predominantly nuclear
191 RNAs (Figure 2C, left). As might be predicted (Derrien et al. 2012), many of this latter
192 group were lncRNAs, which clearly showed preferential nuclear localization (Figure
193 2C, middle). Most mRNAs appeared to be cytosolic in our data (Figure 2C, right).
194 Notably, we also observed sizeable populaces of RNAs exhibiting noncanonical
195 nuclear–cytoplasmic partitioning (Figure 2D). 3323 mRNAs—including *C1orf63*, for
196 example (Figure 2D)—appeared preferentially nuclear. Many of these species have
197 been proposed to play a role in dampening gene expression noise (Bahar Halpern et
198 al. 2015). Conversely, 234 lncRNAs appeared preferentially cytoplasmic; these
199 include the known cytoplasmic lncRNA *SNHG5*, a modulator of staufen-mediated
200 decay that influences colorectal tumor growth (Derrien et al. 2012; Damas et al.
201 2016) (Figure 2D).

202 Our APEX-RIP nuclear and cytosolic RNA lists provide an opportunity for a
203 head-to-head comparison with the traditional Fractionation-Seq method for
204 mapping subcellular RNA localization. ROC analysis of the ENCODE Fractionation-
205 Seq data yielded a list of 3,056 RNAs enriched by nuclear fractionation (Table 2 tab
206 5). Of these RNAs, 81% (2469) were also enriched in our APEX-RIP nuclear dataset,
207 implying general agreement between the two technologies (Figure 2E). Notably,
208 APEX-RIP also enriched nearly 3000 additional transcripts. These may be nuclear-
209 localized RNAs that were opaque to the ENCODE protocol, or contaminants enriched
210 by APEX-RIP. To address this possibility, we examined each dataset for conspicuous
211 non-nuclear contaminants: RNAs that are known to be localized at the Endoplasmic
212 Reticulum (Jan, Williams, and Weissman 2014). Satisfyingly, the APEX-RIP nuclear
213 dataset, though larger, contained fewer ER contaminants than did the analogous
214 fractionation-based dataset, implying that APEX-RIP produces higher specificity
215 than Fractionation-Seq (Figure 2F, left).

216 To compare the coverage/sensitivity of each method (sometimes termed
217 recall), we examined the enrichment in each dataset of lncRNAs, which are thought
218 to be predominantly nuclear (Derrien et al. 2012). We assembled a list of 827
219 annotated lncRNAs (GENCODE hg19) with average FPKM pre-enrichment greater
220 than 1.0 (Table 2 tab 7). Of these lncRNAs, 71.7 % are enriched in our APEX-RIP-
221 derived nuclear dataset, while nuclear Fractionation-Seq from the same cell line
222 enriched only 43.4 % (Figure 2E, right). We conclude that APEX-RIP can be superior
223 to Fractionation-Seq in terms of *both* specificity and coverage, for analysis of
224 endogenous RNA subcellular localization.

225

226 **Enrichment of RNAs proximal to the ER membrane**

227 The above efforts establish that APEX-RIP can enrich RNAs in membrane-enclosed
228 cellular compartments. We next sought to address whether the technique could
229 successfully capture the transcriptomes of “open” subcellular regions. Previous

230 proteomic work has shown that APEX tagging exhibits sufficient spatial specificity
231 for such open compartments, since this technology has produced highly specific
232 proteomic maps of, for example, the mammalian neuronal synaptic cleft(Loh et al.
233 2016), outer mitochondrial membrane(Hung et al. 2017), mitochondrial
234 nucleoid(Han et al. 2017), and G-protein coupled receptor interaction
235 network(Lobingier et al. 2017; Paek et al. 2017). We were unsure, however, if the
236 additional formaldehyde crosslinking step would preserve or blur the estimated <10
237 nanometer spatial resolution of APEX labeling(Rhee et al. 2013).

238 As a test case for the generality of APEX-RIP at such open compartments, we
239 selected the Endoplasmic Reticulum (ER). The ER is an appealing target for several
240 reasons. First, it is host to a known set of characteristic RNAs that we can use as
241 positive controls—the so-called “secretome”—which comprises mRNAs encoding
242 secreted, glycosylated, and/or transmembrane proteins, that are translated by
243 ribosomes on the rough ER. Second, the ER provides the opportunity to compare the
244 efficacy of APEX-RIP to alternative approaches, since RNAs in this subcellular locale
245 have been previously characterized both by Fractionation-Seq, and by a newer
246 methodology termed proximity-dependent ribosome profiling(Jan, Williams, and
247 Weissman 2014; Williams, Jan, and Weissman 2014). This latter technique maps
248 active protein translation at the ER membrane by combining ribosome
249 profiling(Ingolia et al. 2009) with proximity-restricted sequence-specific
250 biotinylation, using an ER-targeted biotin ligase and ribosomes that are tagged with
251 the peptide substrate (AviTag) of that ligase.

252 Since it was initially unclear which face of the ER membrane (cytosolic or
253 luminal) would be most amenable to the APEX-RIP method, we generated fusion
254 constructs that localized the peroxidase catalytic center to each (Figures 3A–B).
255 ERM-APEX2 targets APEX2 to the ER cytosolic surface via a 27-amino acid fragment
256 derived from the native ER membrane (ERM) protein cytochrome P450 C1. HRP-
257 KDEL targets horseradish peroxidase (HRP) to the ER lumen via an N-terminal ER-
258 targeting signal and a C-terminal KDEL ER-retention motif(Martell et al. 2012). We
259 have shown that HRP catalyzes the same proximity-dependent biotinylation
260 chemistry as APEX2(Loh et al. 2016), but has higher specific activity than APEX2 in
261 the ER lumen(Lam et al. 2014). We generated HEK 293T cells stably expressing
262 ERM-APEX2 and HRP-KDEL, and confirmed by microscopy and streptavidin blotting
263 that each produced the expected labeling patterns (Figures 3C and D). Next, we
264 compared the efficacy of each construct for target RNA isolation, using the
265 biotinylation-first/crosslinking-second APEX-RIP protocol, and analyzing our
266 results via RT-qPCR analysis of established secretome and non-secretome mRNAs²⁴.
267 Parallel experiments with APEX2-NES cells served as negative controls (Figure 3E).

268 Intriguingly, while APEX-RIP from HRP-KDEL cells efficiently enriched target
269 secretome mRNAs relative to non-target controls (average fold enrichment = 19.5,
270 paired t-test p-value = 0.00009), parallel experiments in ERM-APEX2 cells exhibited
271 only modest, qualitative enrichment of target species (average fold enrichment =
272 1.49, paired t-test p-value = 0.0515). Indeed, results from ERM-APEX2 cells were
273 nearly indistinguishable from those acquired from APEX2-NES control cells (paired
274 t-test p-value = 0.830, Figure 3E, right). This is surprising, since proteomic

275 experiments in HEK 293T cells expressing the identical ERM-APEX2 construct
276 yielded highly specific enrichment of ER-localized proteins (Hung et al. 2017).
277 Our data strongly imply that APEX-RIP does *not* have the same spatial
278 specificity as peroxidase-catalyzed proteomic labeling, and may be limited by
279 perturbations induced by formaldehyde crosslinking. However, we were highly
280 encouraged by the data obtained with the HRP-KDEL construct. We hypothesize that
281 APEX-RIP with this construct is effective because formaldehyde crosslinking
282 physically couples RNAs on the cytosolic face of the ER to protein complexes that are
283 biotinylated within the ER lumen, thereby allowing target RNAs to be enriched by
284 streptavidin (Figure 3A). Furthermore, we observed that the target specificity of this
285 approach could be greatly improved by addition of a one-minute radical-quenching
286 step in between the biotinylation and crosslinking steps in our protocol (Figure
287 S3A). We surmise that this additional step prevents residual peroxidase-generated
288 radicals from leaking into adjoining compartments when the integrity of the ER
289 membrane is compromised during formaldehyde treatment.

290 Using this improved protocol, we performed APEX-RIP on HRP-KDEL cells
291 (Table 3 tab 2). Gene-level analysis, comparing RNA counts before and after
292 streptavidin pulldown, revealed a distinct population of substantially enriched RNAs
293 (Figures 3F and S3B). Encouragingly, the majority (63.4%) of secretome mRNAs
294 (defined by ER proximal RNAs (Jan, Williams, and Weissman 2014) and Phobius
295 predicted mRNAs with exclusion of nuclear encoded mitochondrial mRNAs, see
296 *methods*) resided in this set, while most (97.1%) mRNAs in a test set of known non-
297 secreted genes were not enriched, thus demonstrating the ability of APEX-RIP to
298 isolate ER-associated transcripts from the larger population of cellular RNAs (Figure
299 3F). Using histogram and ROC analysis, we determined the optimal log₂ FKPM
300 significance threshold cutoff for each experimental replicate (Figure S3C; see
301 *methods*), obtaining a final list of 2970 ERM-associated RNAs that were
302 independently enriched in multiple experiments (Table 3 tab 1). This dataset
303 exhibited 94% specificity, based on previous secretory annotation as defined by
304 GOCC, SignalP, TMHMM, or Phobius (Ashburner et al. 2000; Petersen et al. 2011;
305 Krogh et al. 2001; Käll, Krogh, and Sonnhammer 2004). Figure 3H shows that we
306 also de-enriched mRNAs lacking such signals. Coverage was likewise exceptional
307 (97%), as gauged by the recall of 71 literature-curated well-established ER resident
308 proteins' mRNAs (Table 3 tab 5; Figure 3I, see *methods*).

309 We next compared the ERM APEX-RIP dataset to analogous results obtained
310 by subcellular biochemical fractionation (Reid and Nicchitta 2012), and by
311 proximity-dependent ribosome profiling (Jan, Williams, and Weissman 2014) (Table
312 3, tabs 3 and 4, respectively). Encouragingly, APEX-RIP captures the majority of
313 RNAs enriched by each of these alternative techniques (70% and 93%, respectively,
314 Figure 3J), implying broad agreement between the different methodologies. To
315 examine this further, we quantified the specificity and coverage of each approach, as
316 above (see *methods*). Specificity analysis demonstrated that APEX-RIP and ribosome
317 profiling exhibited similarly high specificity (94% and 98%, respectively). However,
318 Fractionation-Seq was substantially noisier, such that only 90% of enriched mRNAs
319 bore a secretory annotation (Figure 3H); the remaining 10% comprised sizeable
320 populations of conspicuous contaminants (Figure S3E). The coverage of ER-localized

321 mRNAs retrieved by APEX-RIP (97%) was also considerably higher than those
322 retrieved by both Fractionation-Seq and ribosome profiling (73% and 77%,
323 respectively, [Figure 3J](#)). We attribute the enhanced coverage of APEX-RIP to its
324 higher sensitivity, since this method appears better suited for capturing RNAs with
325 lower abundances than do the alternative approaches ([Figure S3 F-G](#)). Such higher
326 sensitivity may also explain why the set of RNAs enriched by APEX-RIP is so much
327 larger than those obtained by fractionation and ribosome-profiling ([Figure 3H](#)).
328 Excitingly, this further underscores the ability of APEX-RIP to recover RNAs that are
329 opaque to other methods. While the vast majority (88.7%) of our enriched RNAs are
330 mRNAs, we also enrich hundreds of noncoding RNA species—including antisense
331 RNAs and lincRNAs ([Figure 3G](#)). These RNAs are not translated, and thus cannot be
332 detected by ribosome profiling, and tend to be lowly expressed, making them
333 difficult targets for either ribosome profiling or Fractionation-Seq.

334 In summary, APEX-RIP is superior to existing methods for mapping
335 endogenous RNAs proximal to the ER membrane, and may be extensible to other
336 membrane-abutting subcellular regions as well.

337

338 **Hypotheses from ER and nuclear APEX-RIP datasets**

339 We wondered if the highly specific and comprehensive RNA subcellular localization
340 datasets produced by APEX-RIP could be mined for new biological hypotheses. We
341 first observed that, of the 2635 mRNAs in our ERM dataset, 141 code for
342 mitochondrial proteins. It is thought that that the bulk of the nuclear-encoded
343 mitochondrial proteome is translated within the bulk cytosol, or in proximity to
344 mitochondria themselves(Lesnik, Golani-Armon, and Arava 2015), raising the
345 possibility that the translation or subsequent processing of these 141 protein
346 products require machinery localized to the ER. Additionally, these mRNAs may be
347 translated at mitochondria-ER contact sites, some of which have been observed to
348 contain ribosomes(Csordás et al. 2006). To gain initial insight into these unusual
349 RNAs, we analyzed these 141 genes to see whether, relative to total pool of mRNAs
350 encoding mitochondrially-localized proteins, they were enriched in particular
351 properties ([Table 4 tab 1](#)). Intriguingly, 57.1% of these mRNAs code for
352 transmembrane proteins (as predicted by TMHMM), compared to only 20.4% for all
353 mitochondrial protein mRNAs ([Figure 4A](#)). Mitochondrial subcompartment analysis
354 showed that the ER-proximal population is enriched for proteins destined for the
355 inner mitochondrial membrane, and is depleted for resident matrix proteins,
356 compared to the total mitochondrial proteome ([Figure 4B](#)). Interestingly, proximity-
357 dependent ribosome profiling by Williams et al. in yeast using biotin ligase targeted
358 to the outer mitochondrial membrane(Williams, Jan, and Weissman 2014) also
359 showed enrichment of mRNAs encoding proteins destined for the inner
360 mitochondrial membrane. Perhaps a subset of inner mitochondrial membrane-
361 destined proteins are locally translated at mitochondria-ER contact sites.

362 Next, we tested whether new insights could be gained by examining RNAs
363 that APEX-RIP had enriched from more than one subcellular compartment. Because
364 the ER lumen is contiguous with that of the nuclear envelope, we hypothesized that
365 the HRP-KDEL APEX-RIP experiment, in addition to enriching RNAs proximal to the
366 ER, might also enrich RNAs proximal to the nuclear membrane. This region within

367 the nucleus, termed the nuclear lamina, is widely thought to play a critical role in
368 gene repression (Kind and van Steensel 2010), and in shaping the global three-
369 dimensional architecture of chromatin (C.-K. Chen et al. 2016). However, no
370 exclusively laminar-resident RNAs have yet been identified. We hypothesized that
371 we might identify such long-sought lamina RNAs by intersecting our APEX-RIP
372 nuclear and ERM RNA lists (Figure 4C). Encouragingly, we observed 673 such RNAs
373 in the intersection list, 34 of which are long noncoding RNAs (Figure 4D; Table 4 tab
374 2). This small list is a compelling starting point for exploration of regulatory RNAs
375 that may reside in the nuclear lamina.

376

377 Discussion

378 Methods for mapping RNA subcellular localization are constrained by the limits of
379 their spatiotemporal precision, the diversity of RNA species that they can
380 simultaneously analyze, their generality across cell types and compartments, and
381 their ease of use. We believe that APEX-RIP is superior to existing imaging- and
382 sequencing-based techniques with regard to many of these factors.

383 Compared to imaging-based technologies, APEX-RIP offers superior target
384 throughput, ease of use, and temporal control. For example, although modern
385 variants of FISH can achieve extremely high spatial precision—even enabling the
386 visualization of individual RNA molecules (Batish, Raj, and Tyagi 2011)—this
387 technique requires the synthesis and testing of customized fluorescent probes for
388 each transcript under enquiry, a cumbersome process that limits throughput (Cabili
389 et al. 2015). A highly multiplexed FISH variant, merFISH, substantially boosts
390 throughput—enabling thousands of transcripts to be simultaneously visualized—
391 but requires complex protocols for probe set design and imaging (K. H. Chen et al.
392 2015). An alternate approach, FISSEQ, achieves similar target depth without the
393 need for gene-specific probes, but instead relies on customized instrumentation and
394 a rococo process of *in situ* sequencing and imaging (Lee et al. 2014). Notably, without
395 incorporating additional stains or markers, these imaging-based approaches
396 provide little information regarding the local environment (*i.e.*, proximal cellular
397 compartments or features) near each RNA target. Furthermore, these techniques
398 fundamentally lack temporal precision: each requires extensively fixing and
399 permeabilizing cells prior to data collection, during which time diffusion or the loss
400 of cellular integrity can perturb endogenous RNA localization. This latter issue can
401 be circumvented through a variety of live-cell imaging technologies, but these
402 require the implementation of customized reagents that limit throughput, and may
403 even distort the localization of the RNA targets under enquiry (Paige, Wu, and Jaffrey
404 2011; Hocine et al. 2013; Nelles et al. 2016). By contrast, APEX-RIP is not
405 encumbered by any of these constraints. It does not require the development of
406 target-specific expression constructs or probes; nor does it rely on specialized
407 instrumentation. The ensemble of RNA targets analyzed (and, for that matter, the
408 array of RNA classes analyzed) is theoretically limited only by the library synthesis
409 and sequencing protocols employed. Moreover, since APEX-RIP captures only RNAs
410 proximal to a specific subcellular compartment, and does so during a one-minute
411 reaction, the technique offers both greater information content and higher temporal
412 resolution than do its imaging-based alternatives.

413 Compared to fractionation-based technologies, APEX-RIP offers superior
414 accuracy, ease of use, and general versatility. As illustrated in the nucleus and ER,
415 our technique outperforms conventional Fractionation-Seq with regard to both
416 target specificity and recall, apparently circumventing the dual issues of target loss
417 and off-target contamination that can plague such approaches (Figures 2E-F). We
418 ascribe this performance boost to two principal factors. First, the high
419 spatiotemporal precision afforded by *in situ* biotinylation (Rhee et al. 2013) allows
420 us to efficiently isolate target material from contaminants that might be difficult to
421 remove by classical fractionation, thereby improving specificity. Second, covalently
422 coupling target RNAs to affinity-tagged proteins allows us to recover low-abundance
423 or weakly affiliated transcripts that might otherwise be lost during biochemical
424 enrichment, thereby improving target recall. Perhaps more importantly, however,
425 we have achieved these results in a variety of subcellular compartments using a
426 common protocol, thus obviating the need to develop customized purification
427 schemes for each compartment. This generality should enable APEX-RIP to access
428 “unpurifiable” subcellular compartments for which such purification schemes would
429 be impossible. While a related technology, proximity-dependent ribosome profiling,
430 exhibits similar versatility within diverse subcellular milieus (Jan, Williams, and
431 Weissman 2014), this approach is limited to mRNAs actively undergoing translation.
432 It also requires biotin starvation prior to tagging, which is toxic to mammalian cells.
433 As we have demonstrated, APEX-RIP can map diverse classes of noncoding RNA and
434 quiescent mRNA (Figure 3G), and toxic protocols starving cells of essential nutrients
435 for hours are not required.

436 The APEX-RIP methodology does have notable limitations. Cells to be
437 analyzed must be transfected with a recombinant construct, in contrast to FISH and
438 Fractionation-Seq, which can be performed on native tissues. APEX-RIP also gives
439 poor spatial specificity in membrane-free subcellular regions.

440 The APEX peroxidase used here has also previously been used to generate
441 contrast for electron microscopy in fixed cells (Martell et al. 2012; Lam et al. 2014),
442 and for spatially-resolved proteomic mapping in living cells (Rhee et al. 2013; Hung
443 et al. 2014; Loh et al. 2016; Han et al. 2017; Hung et al. 2017; Mick et al. 2015). This
444 study extends APEX to a new class of applications and to a new biopolymer. In
445 principle, it should be possible to utilize a single APEX-expressing cell line to
446 characterize a target subcellular compartment by electron microscopic, proteomic,
447 and transcriptomic means. Related methods for proteomic mapping, such as
448 BioID (Roux et al. 2012), lack this versatility, because the underlying chemistry is not
449 as flexible as the one-electron oxidation reaction catalyzed by APEX.

450 We anticipate that the initial subcellular transcriptomic map presented in
451 this work—probing the mitochondrial matrix, cytosol, nucleus, and ER membrane of
452 HEK293T cells—will serve as valuable resources for cell biologists. Analysis of these
453 data has already yielded potential insight into nuclear-retained mRNAs, cytosolic
454 lncRNAs, putative lamina-localized RNAs, and genes that may be translated locally at
455 mitochondria-endoplasmic reticulum junctions. Applying APEX-RIP at other
456 subcellular compartments will further expand the depth and breadth of this map.
457 Furthermore, given the high temporal resolution of APEX-RIP, we imagine that our
458 technology might enable profiling of subcellular RNA pools in response to acute

459 stimuli or drugs, or throughout stages of the cell cycle and development.
460 Collectively, such studies would yield an understanding into the biology of RNA
461 subcellular localization at unprecedented scale.

462

463 **Significance**

464 RNA subcellular localization is a critical factor that influences a wide array of
465 biological processes, ranging from *Drosophila* embryogenesis to mammalian
466 neuronal signaling. However, while this spatial layer of transcriptome regulation has
467 been characterized in a handful of contexts, a broader understanding of its overall
468 extent, the factors governing its establishment, and its impact on biological function,
469 remain inchoate. The limitations hindering this understanding have been largely
470 technical, since conventional methods—such as fluorescence *in situ* hybridization
471 (FISH) and Fractionation-Sequencing (“Frac-Seq”)—depend upon specialized
472 reagents and protocols that can limit throughput and general applicability. To
473 address this fundamental need, we have developed a new strategy—APEX-RIP—
474 which uses a simple toolkit and workflow to map the transcriptomes of discrete
475 subcellular compartments at high depth and spatiotemporal resolution. APEX-RIP
476 utilizes the engineered ascorbate peroxidase APEX to biotinylate proteins within a
477 target subcellular compartment in live cells; these affinity-tagged proteins are then
478 chemically crosslinked *in situ* to nearby RNAs. When applied to a variety of
479 membrane-enclosed and membrane-adjacent compartments, the APEX-RIP strategy
480 exhibited higher target specificity and coverage than do conventional fractionation-
481 sequencing-based approaches, at a depth far exceeding those attainable by imaging-
482 based methods. Furthermore, APEX-RIP can be applied to compartments that are
483 recalcitrant to conventional biochemical purification. Given the superior precision,
484 flexibility, and ease of this approach, we anticipate that APEX-RIP will provide a
485 powerful tool for dissecting RNA subcellular localization in a broad range of
486 biological contexts.

487

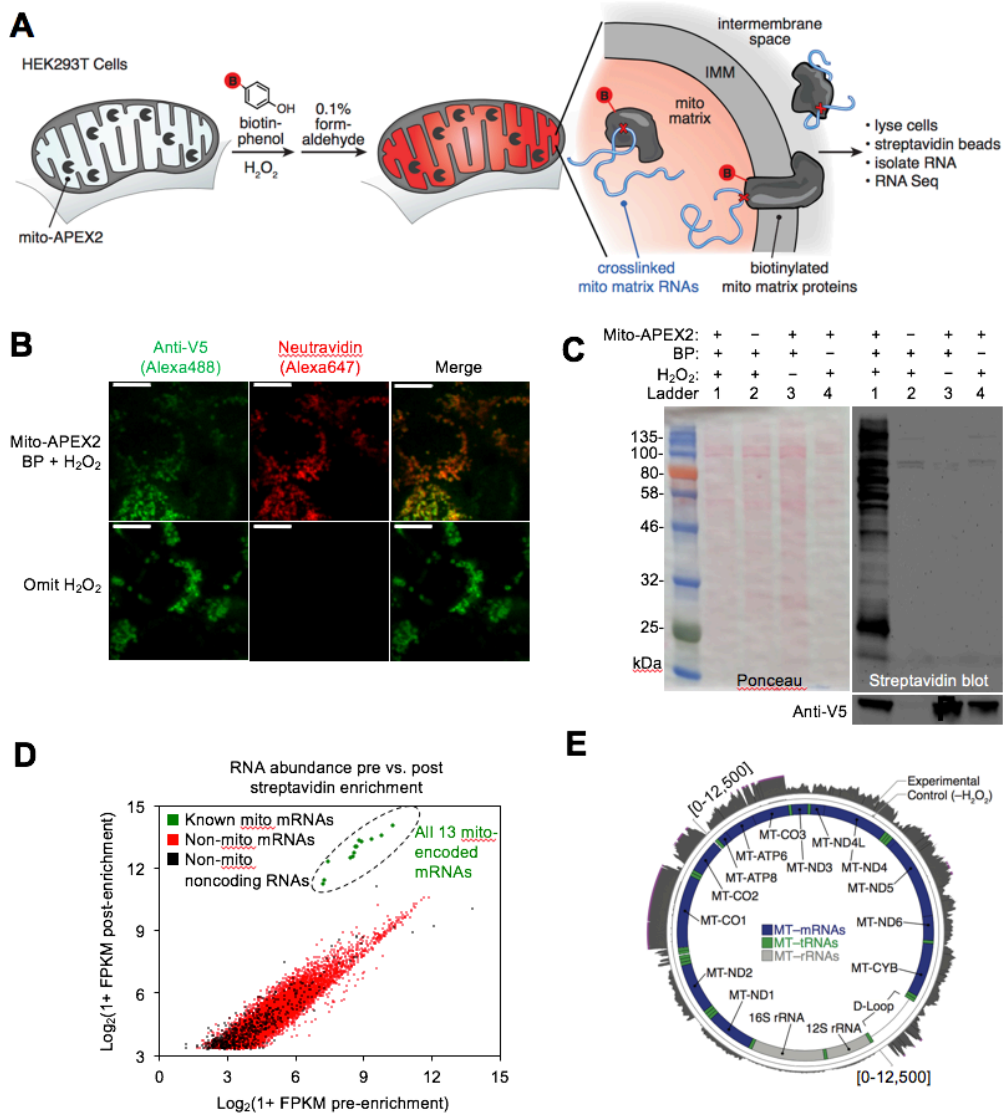
488 **Author contributions**

489 Conceptualization, PK, DMS, AYT; Methodology, PK, DMS, AYT; Validation, PK and
490 DMS; Formal Analysis, PK and WM; Investigation, PK and DMS; Data Curation, PK;
491 Writing – Original Draft, PK, DMS, and AYT; Writing – Review & Editing, PK, DMS,
492 JLR, and AYT; Visualization, PK, DMS, and AYT; Supervision, JLR and AYT; Project
493 Administration, AYT; Funding Acquisition, JLR and AYT.

494

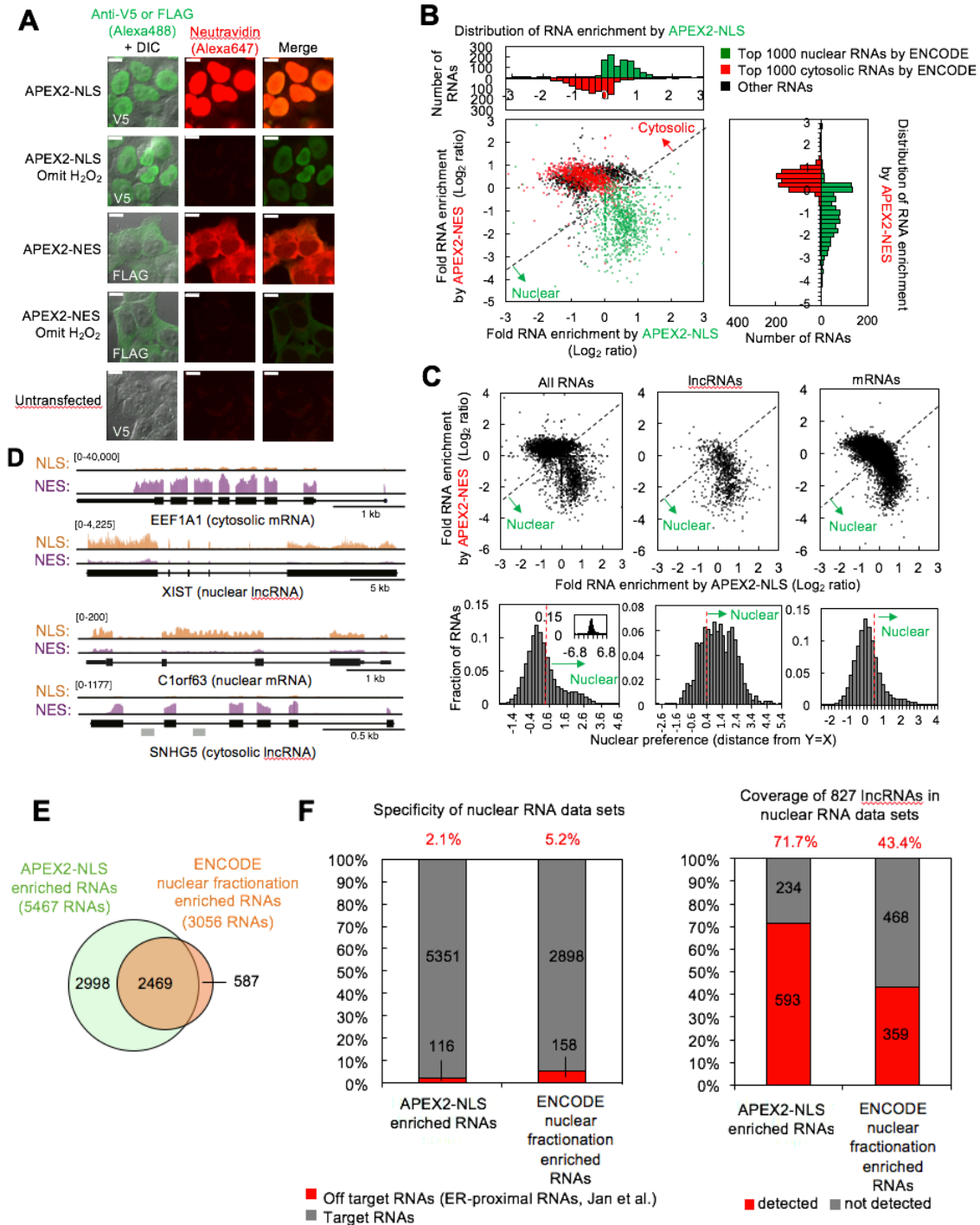
495 **Acknowledgements**

496 We thank members of the Ting laboratory, especially Jeffrey Martell for valuable
497 experimental advice and Ozan Aygun for the curated ER protein list. We thank the
498 Rinn laboratory, especially Chinmay Shukla, for valuable computational advice.
499 Funding was provided by the NIH (R01-CA186568 to A.Y.T. and U01 DA040612 to
500 J.L.R.) and Stanford (to A.Y.T).



501
502 **Figure 1. APEX-RIP in mitochondria.** (A) Overview of the APEX-RIP workflow. Cells
503 expressing APEX2 (grey 'pacmen') targeted to the compartment of interest (here, the
504 mitochondrial matrix.) are incubated with the APEX substrate biotin-phenol (BP; red B:
505 biotin). A one-minute pulse of H_2O_2 initiates biotinylation of proximal endogenous
506 proteins (Rhee et al. 2013), which are then covalently crosslinked to nearby RNAs by 0.1%
507 formaldehyde. Following cell lysis, biotinylated species are enriched by streptavidin
508 pulldown, and coeluting RNAs are analyzed by RT-qPCR or RNA-Seq. IMM: inner
509 mitochondrial membrane. (B) Imaging APEX2 biotinylation *in situ*. HEK 293T cells
510 expressing V5-tagged mito-APEX2 were biotinylated and fixed as described in (A) and
511 stained as indicated. The bottom row is a negative control in which H_2O_2 treatment was
512 omitted. Scale bars, 10 μ m. (C) Streptavidin blot analysis of whole cell lysates prepared as
513 described in (A). *In situ* biotinylation (lane 1) is ablated in the absence of the APEX2 protein,
514 H_2O_2 , or BP. Anti-V5 blot detects expression of mito-APEX2. (D-E) mito-APEX-RIP efficiently
515 recovers the mitochondrial transcriptome. (D) Gene-level RNA-Seq analysis of mito-APEX-
516 RIP enrichment for RNAs longer than 200nt. Data from one biological replicate are shown.
517 (E) Nucleotide-level RNA-Seq analysis of mito-APEX-RIP, mapped to the human
518 mitochondrial genome (innermost circle). Outermost circle: reads from the full APEX-RIP

519 protocol; middle circle: reads from the negative control. Note the enrichment of several
520 mitochondrially-encoded tRNAs and the D-loop leader transcript. Ribosomal RNAs were
521 removed during library preparation (*see methods*). See also: [Figure S1](#).



522

523

524

525

526 **Figure 2. APEX-RIP mapping of the nuclear-cytoplasmic RNA distribution. (A)**

527 Fluorescence imaging of nuclear and cytosol-targeted APEX2 fusion constructs. HEK 293T

528 cells expressing the indicated constructs (“NLS,” nuclear localization signal; “NES,” nuclear

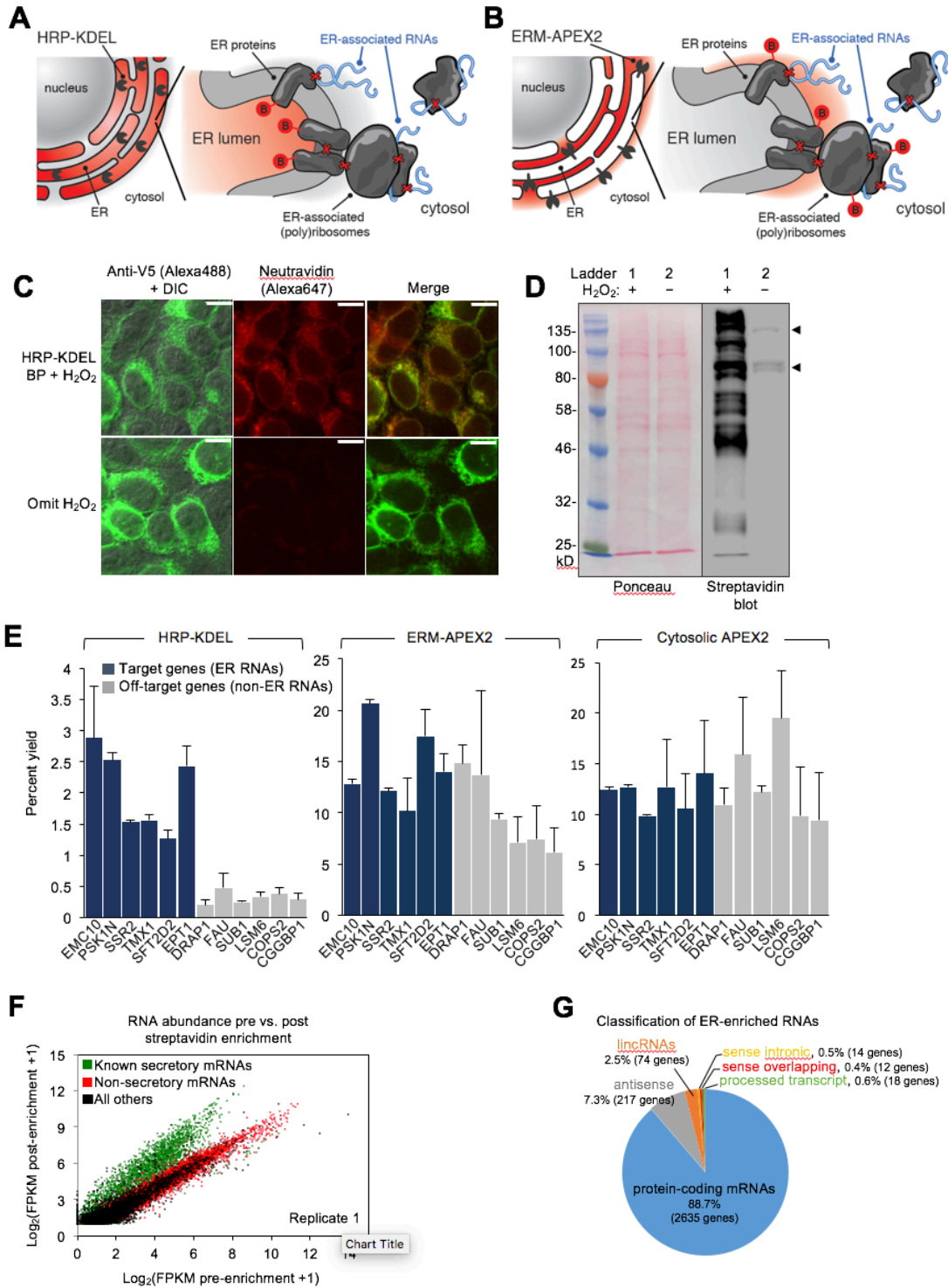
529 export signal) were labeled with biotin-phenol, crosslinked and stained as indicated. DIC,

530 Differential Interference Contrast. Scale bars, 10 μ m. (B) Combined analysis of APEX2-NLS

531 and NES experiments distinguishes nuclear and cytoplasmically localized RNAs. Fold

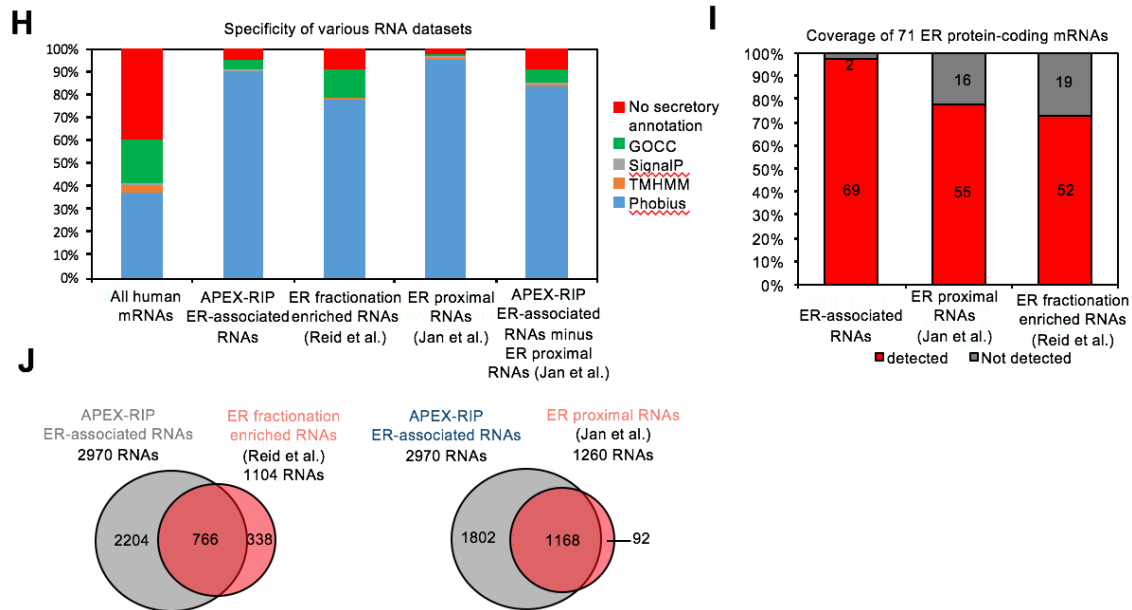
enrichment values were calculated relative to matched input samples; the median values of

532 three replicates are shown (see *methods*). The 1000 RNAs with the highest predicted
533 nuclear and cytosolic localization by ENCODE(Dunham et al. 2012) are colored green and
534 red, respectively (see *methods*). Histogram plots, summarizing the separation of these RNA
535 standards by each APEX2 construct, are projected along the axes of the scatter plot and use
536 the same scales. The black dotted line marks the cutoff between nuclear and cytosolic RNAs.
537 (C) APEX-RIP captures the established nuclear-cytoplasmic distribution of mRNAs and
538 lncRNAs. *Top*: APEX2-NLS versus APEX2-NES scatter plots, as in (B), for all RNAs (*left*),
539 lncRNAs (*middle*), and mRNAs (*right*). Data are the medians of three replicates. Dotted lines
540 mark the cutoff between nuclear and cytosolic RNAs, as in (B). *Bottom*: histogram plots of
541 nuclear preference scores (see *methods*) for each class of RNA. Dotted red lines: the ROC-
542 derived significance threshold (see *methods*). Inset: the complete distribution. (D) Read
543 density plots of RNAs with stereotypical and atypical localization. For each gene, a common
544 y-scale is used for all read tracks. SnoRNAs encoded in the *SNHG5* gene body are indicated
545 as gray rectangles. (E) Venn diagram comparing APEX-RIP and ENCODE nuclear RNA
546 datasets(Dunham et al. 2012). (F) Nuclear APEX-RIP is more specific and sensitive than is
547 biochemical fractionation. *Left*: Specificity of the APEX-RIP and ENCODE nuclear RNA
548 datasets(Dunham et al. 2012). Off-target RNAs were defined actively translated ER-
549 proximal mRNAs (Jan, Williams, and Weissman 2014). *Right*: Recall of nuclear standard
550 RNAs, defined as a set of 827 lncRNAs annotated by GENCODE hg19 with average pre-
551 enrichment FPKM ≥ 1.0 . See also: [Figure S2](#).

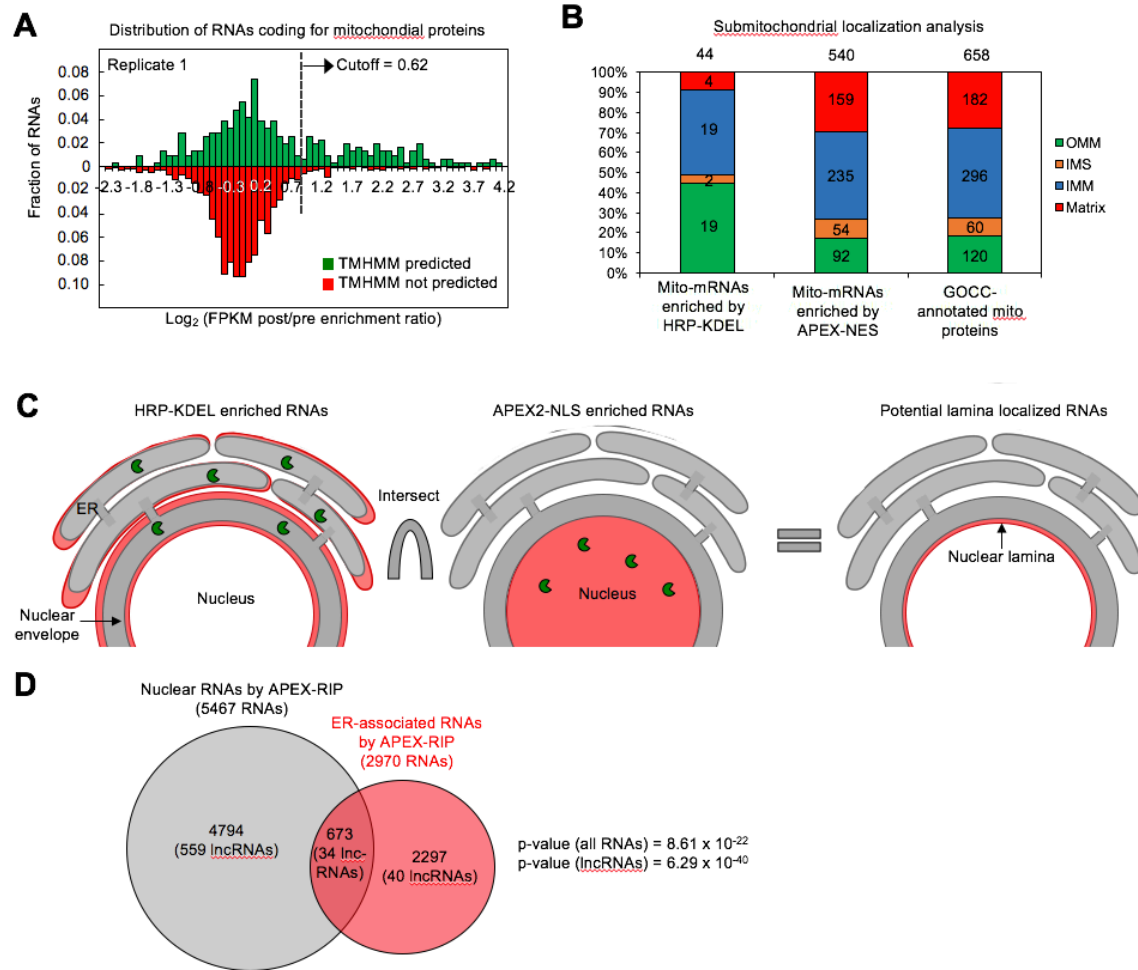


552

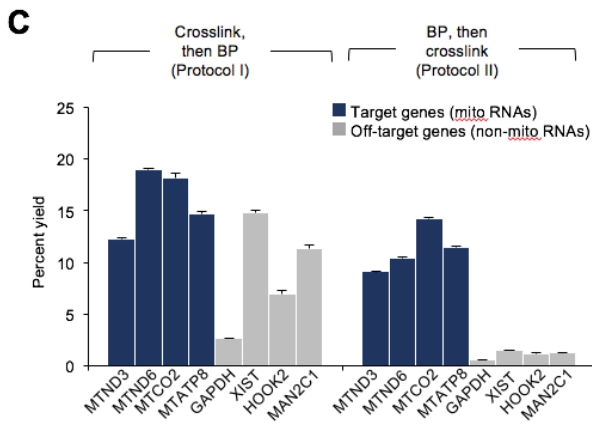
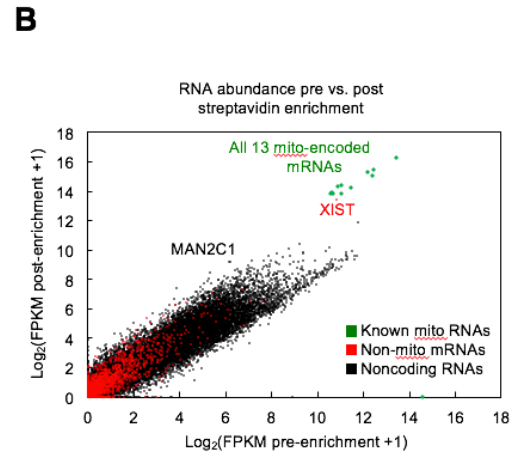
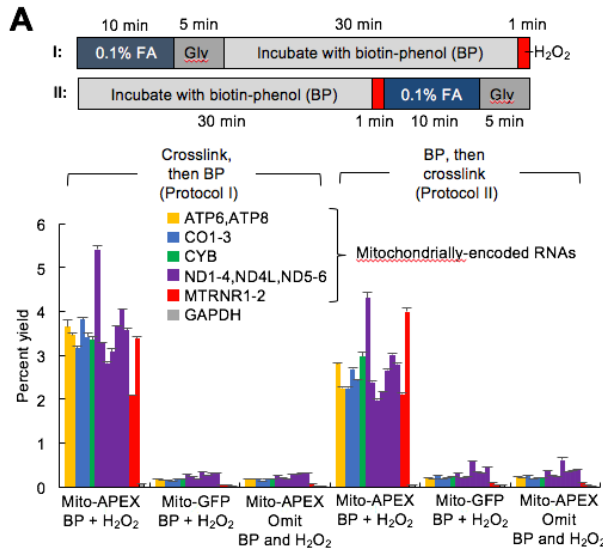
553



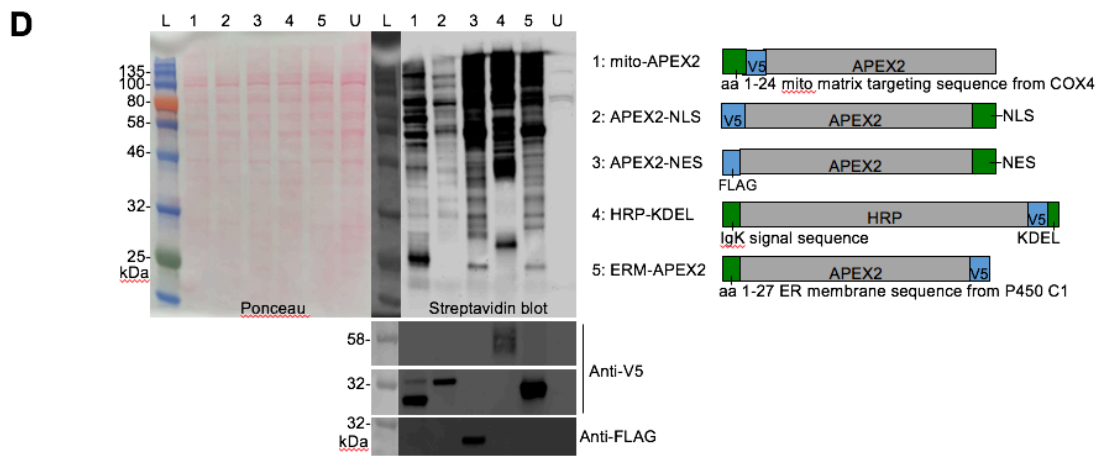
554
555 **Figure 3. APEX-RIP maps RNAs proximal to the Endoplasmic Reticulum. (A–B)**
556 Schematics summarizing alternate ER-targeting strategies. (A) HRP, targeted to the ER with
557 a KDEL sequence, biotinylates proteins within the ER lumen. Red B: biotin. Red Xs: chemical
558 crosslinks induced by 0.1% Formaldehyde treatment. (B) APEX2, displayed on the ER
559 membrane (ERM) by fusing it to the transmembrane segment of rabbit P450 C1, faces the
560 cytosol. (C) Imaging HRP-KDEL-catalyzed biotinylation. HEK293T cells stably expressing
561 HRP-KDEL were labeled with BP, fixed and imaged as in Figure 1B. DIC, Differential
562 Interference Contrast. Scale bars, 10 μ m. (D) Streptavidin blot detection of resident ER
563 proteins biotinylated by HRP-KDEL, as in Figure 1C. Arrowheads denote endogenously
564 biotinylated proteins (Chapman-Smith and Cronan 1999). (E) RT-qPCR analysis, comparing
565 specificities of the labeling schemes shown in (A–B). Target and off-target genes were
566 selected using previously-reported fold enrichments at the ER membrane (Jan, Williams,
567 and Weissman 2014). Data are the mean of three replicates, \pm one standard deviation. (F)
568 Scatter plot showing RNA abundance before and after streptavidin enrichment. Data shown
569 are for one replicate. Additional replicates in Figure S3B. (G) Classification of APEX-RIP
570 enriched, ER-associated RNAs. Collectively, non-coding RNAs constitute 11.3% of enriched
571 genes (335 of 2970 RNAs). (H) Specificity analysis for protein-coding mRNAs in our ER-
572 associated RNA list. 95% of the 2635 APEX-RIP ER-enriched mRNAs exhibit some form of
573 secretory annotation (as predicted by Phobius, TMHMM, SignalP, or GOCC, *see methods*),
574 whereas only 60.3% of all human mRNAs are similarly classified (*left*). RNAs in our dataset
575 that were not enriched by ribosome profiling (1802 RNAs) were also predominantly
576 secretory (90.9%). (I) Target recall of ER APEX-RIP exceeds those of proximity-restricted
577 ribosome profiling (Jan, Williams, and Weissman 2014) and biochemical fractionation (Reid
578 and Nicchitta 2012). See also: Table S5. (J) The gene set enriched by ER APEX-RIP largely
579 recapitulates those enriched by alternative methods. See also: Figure S3.



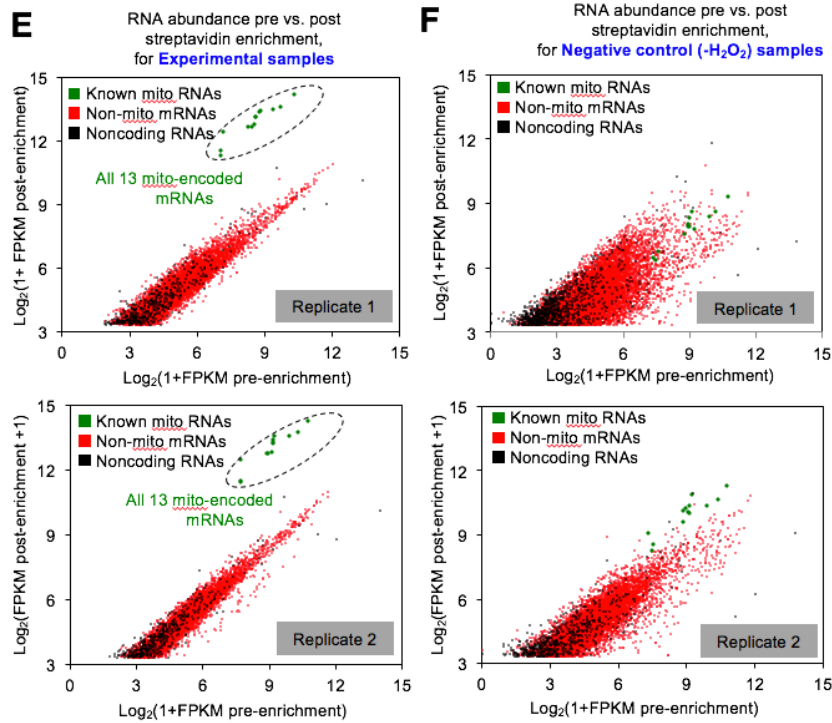
580
 581 **Figure 4. APEX-RIP reveals RNAs with potentially novel localization.** (A) Many
 582 mitochondrial transmembrane proteins appear to be translated at the ER. mRNAs encoding
 583 mitochondrial proteins defined by GOCC and MitoCarta 1.0 (Pagliarini et al. 2008; Ashburner
 584 et al. 2000), with predicted transmembrane helices (predicted by TMHMM (Krogh et al.
 585 2001); green distribution) are preferentially enriched by HRP-KDEL APEX-RIP, relative to
 586 mitochondrial mRNAs lacking transmembrane domains (red distribution) (B) Predicted
 587 localization of mitochondrial proteins encoded by mRNAs that were enriched in ER- and
 588 bulk cytosol-APEX-RIP experiments (left and middle, respectively), and of all GOCC-
 589 annotated mitochondrial proteins. OMM: Outer mitochondrial membrane. IMS:
 590 Intermembrane space. IMM: Inner mitochondrial membrane. (C) Computational scheme for
 591 identifying putative lamina-associated RNAs. Since HRP-KDEL-enriched RNAs (left)
 592 comprise both ER- and lamina-associated RNAs, candidate nuclear lamina-localized RNAs,
 593 were identified as the intersection (right) of RNAs enriched by both APEX2-NLS (middle)
 594 and HRP-KDEL. Red: enriched RNAs; green pacmen: is APEX2 or HRP peroxidases. (D) Venn
 595 diagram identifying putative lamina-associated RNAs, defined as the overlap between HRP-
 596 KDEL- and APEX2-NLS-enriched RNAs. See also: Table 4, tab 2. The significance of overlap
 597 between ER-associated RNAs and nuclear-enriched RNAs by APEX-RIP is measured by
 598 hypergeometric test using all type of RNAs or only lncRNAs as population, respectively.



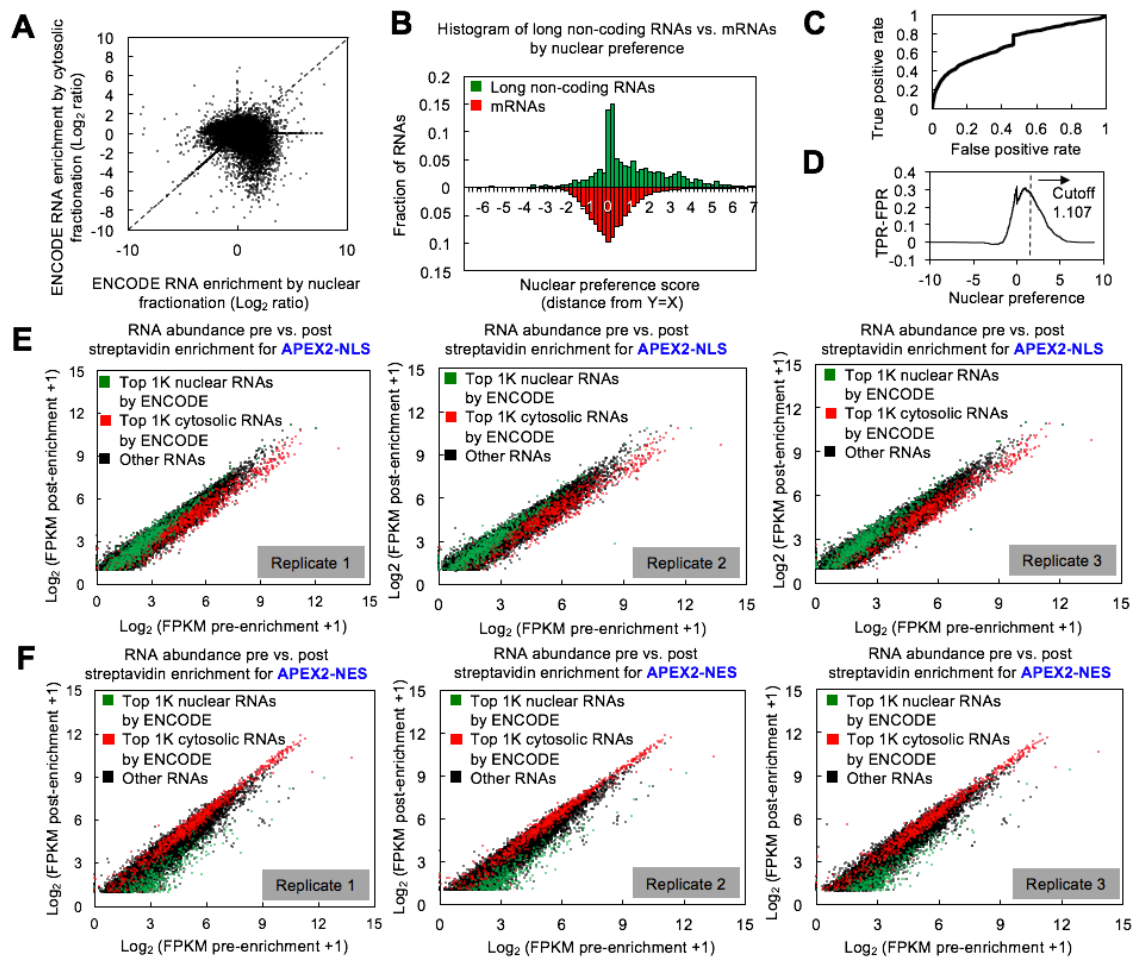
599



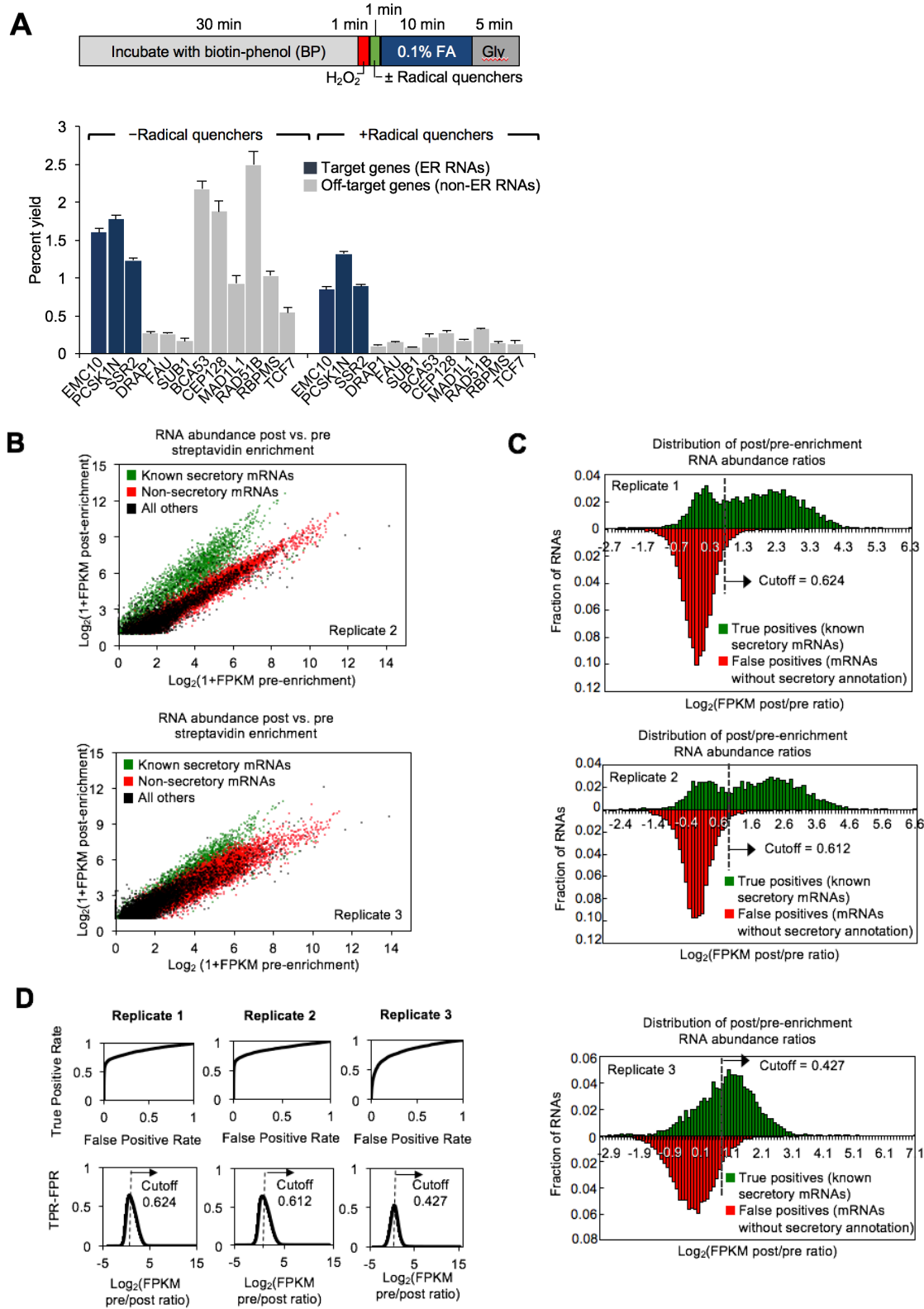
600



601
602 **Supporting Figure 1. Optimization of APEX-RIP protocol and additional mitochondrial**
603 **APEX-RIP data (related to Figure 1).** (A) *Top*: Alternate labeling and crosslinking
604 protocols. In protocol I., cells are crosslinked with formaldehyde (FA) and quenched with
605 Glycine (Gly) prior to the introduction of biotin-phenol (BP) and the initiation of APEX-
606 catalyzed biotinylation with H₂O₂. In Protocol II., live cells are incubated in BP, and *in situ*
607 biotinylation is initiated prior to FA crosslinking. In both cases, cell lysis, streptavidin
608 enrichment and RNA purification proceed as described (*see methods*). *Bottom*: qRT-PCR
609 analysis comparing Protocols I and II. Negative control experiments replace mito-APEX with
610 mito-GFP, omit BP or omit H₂O₂. All constructs were APEX1 derivatives, transiently
611 expressed in HEK 293T cells. Data are the means of three replicates ± one standard
612 deviation. (B) RNA-Seq analysis of RNAs enriched by protocol I. Although all 13
613 mitochondrially-encoded mRNAs (green) were enriched, these were accompanied by
614 substantial contaminating RNAs, including *XIST* and *MAN2C1*. (C) qRT-PCR analysis
615 comparing Protocols I. and II., including off-target controls designed using the results from
616 RNA-Seq. Note the superior enrichment obtained using Protocol II. Cells in this experiment
617 stably-expressed mito-APEX2. This protocol and cell line were used to collect all data in
618 (Figure 1). Data are the means of three replicates ± one standard deviation. (D)
619 Characterization of all APEX2 fusion constructs used in this work. HEK 293T cells stably
620 expressing the indicated constructs (*right*) were labeled and crosslinked via Protocol II.,
621 lysed and analyzed by SDS-PAGE, blotting with streptavidin-HRP, Anti-V5 and anti-FLAG. L:
622 ladder; U: untransfected HEK 293T cells. (E) Data from additional replicates of the
623 mitochondrial APEX-RIP experiment, depicted as in (Figure 1D). (F) In the absence of H₂O₂
624 treatment, mito-APEX-RIP fails to enrich mitochondrial target genes. Data from individual
625 replicates are shown.

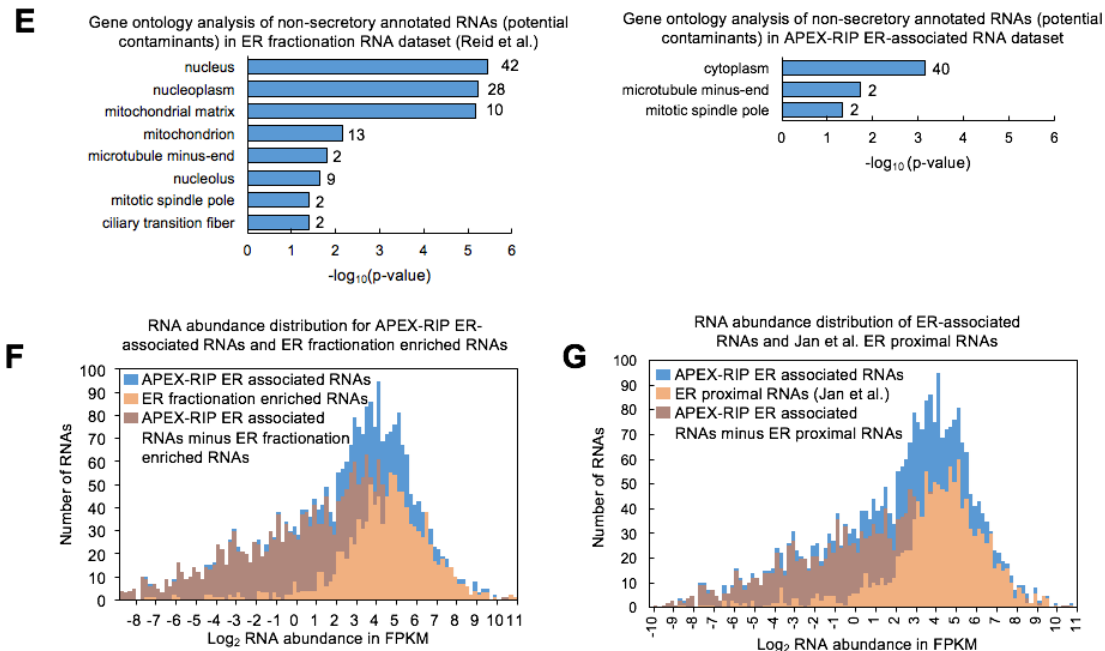


626
 627 **Supporting Figure 2. Identification of ENCODE-derived nuclear RNA standards;**
 628 **additional representations of the APEX-RIP nuclear—cytoplasmic data. (A)** human cell
 629 line (NHEK) fractionation data from ENCODE(Dunham et al. 2012), illustrating the relative
 630 enrichment of genes during nuclear and cytosolic fractionation. RNA enrichments are
 631 calculated as in (Figure 2B) (see methods). Displayed are the average values of three
 632 replicates, for all genes with $FPKM_{pre\ enrichment} \geq 1.0$. Dashed line denotes $y = x$ line. (B)
 633 Distribution of genes by nuclear preference score (see methods). lncRNAs, predicted to be
 634 mostly nuclear, are plotted in green, and mRNAs, predicted to be mostly cytosolic, are
 635 plotted in red. (C–D) Selection of nuclear-enriched RNA standards using ENCODE
 636 fractionation data. (C) ROC analysis applied to histogram in (B). For each nuclear preference
 637 cutoff value, the True Positive Rate (TPR)—defined as the fraction of lncRNAs above the
 638 cutoff—was plotted against the False Positive Rate (FPR)—defined as the fraction of mRNAs
 639 above the cutoff. (D). Using the output of ROC analysis, a nuclear preference score cut-off
 640 value was calculated, defined as the first local maximum in the graph of (TPR-FPR) versus
 641 Nuclear Preference Score. This value was applied to obtain a list of 3056 nuclear-enriched
 642 standard RNAs. (E) Data from individual replicates of the NLS-APEX-RIP experiment. Genes
 643 enriched by nuclear and cytoplasmic fractionation in the ENCODE dataset are colored green
 644 and red, respectively. (F) Data from individual replicates of the NES-APEX-RIP experiment,
 645 depicted as in (E).



646

647



648
649 **Supporting Figure 3. Further optimization of the APEX-RIP protocol; additional HRP-**
650 **KDEL data (Related to Figure 3).** (A) Addition of a radical quenching step between APEX2
651 labeling and formaldehyde crosslinking improves the specificity of RNA capture. *Top:*
652 schematic of the revised labeling–crosslinking workflow. *Bottom:* qRT-PCR analysis of the
653 ER-APEX-RIP experiment with and without this additional quenching step, as in (Figure
654 S1C). The radical quenchers used were Trolox and ascorbic acid. (B–D) Quality assessment
655 for individual replicates of the ER-APEX-RIP experiment. (B) Data for two additional
656 replicates of the HRP-KDEL APEX-RIP experiment, depicted as in (Figure 3F). (C)
657 Histograms showing the distribution of RNA enrichment values $-\log_2\left(\frac{[\text{FKPM}]_{\text{post enrichment}}}{[\text{FKPM}]_{\text{pre enrichment}}}\right)$ —for mRNAs encoding known secretory proteins (top
658 histograms, green), and mRNAs encoding non-secretory proteins (bottom histograms, red).
659 Significance cutoffs were determined using ROC analysis (*see below*). Known secretory and
660 non-secretory standard mRNAs were cataloged as described (*see methods*). (D)
661 Determination of significance thresholds ratio cut-offs by ROC analysis. Data were
662 processed as described in Figures S2D–E, using True Positive secretory and False Positive
663 non-secretory standard RNAs (*see methods*). Given these analyses, and those in (B–C), only
664 data from Replicates 1 and 2 were used to generate our final ER-associated RNA list. To do
665 this, the individual ROC-derived thresholds were used to calculate significantly enriched
666 genes from each dataset, and the final ER-associated RNA list was defined as the
667 intersection of these two lists. (E) Gene ontology (GO) analysis of mRNAs in ER datasets
668 lacking secretory annotation. *Left:* non-secretory mRNAs enriched by biochemical
669 fractionation(Reid and Nicchitta 2012) predominantly exhibit nuclear and mitochondrial
670 annotation. *Right:* non-secretory mRNAs enriched by ER-APEX-RIP have predominantly
671 cytosolic annotation. All terms shown have $p < 0.05$, as assessed using DAVID(Huang,
672 Sherman, and Lempicki 2009). (F–G) APEX-RIP recovers low-abundance targets with
673 greater efficiency than do conventional approaches. In each case, starting abundances are
674 defined by ENCODE. (F) Comparison of APEX-RIP to biochemical fractionation. (G)
675 Comparison of APEX-RIP to proximity-restricted ribosome profiling.
676
677
678

679 **Experimental Procedures**

680

681 **Plasmids and cloning**

682 The pCDNA3 mito-APEX plasmid was published previously (Rhee et al. 2013). The
683 Mito-APEX2 construct was cloned from this plasmid using a two-step protocol. First,
684 the A134P mutation(Lam et al. 2014) was introduced into the APEX gene itself,
685 using QuikChange mutagenesis (Agilent), and thereafter the APEX2 gene was moved
686 to the lentiviral vector pLX304 via Gateway cloning (Thermofisher), to generate
687 plasmid pLX304 mito-APEX2. Other APEX-fusion constructs (pLX304 APEX2-NLS,
688 pLX304 APEX2-NES, and plx304 ERM-APEX2) were cloned by Gibson assembly
689 (NEB), using PCR to add targeting sequences and Gibson Assembly homology arms
690 to the APEX2 gene, and joining the resulting insert into the pLX304 vector digested
691 by BstBI and NheI. For HRP-KDEL, HRP C previously published(Martell et al. 2016)
692 was used as a template to make HRP-KDEL-IRES-Puromycin PCR fragment. Then the
693 insert was cloned into PCDNA3 vector digested by NotI and XbaI. Targeting
694 sequences and restriction sites for all constructs are listed in (Table S1).

695

696 **Mammalian cell culture**

697 Human embryonic kidney (HEK) 293T cells were obtained from ATCC, and cultured
698 in growth media consisting of 1:1 DMEM:MEM (Cellgro), supplemented with 10%
699 Fetal Bovine Serum (FBS), 50 units/mL penicillin, and 50 µg/mL streptomycin, at 37
700 °C and under 5% CO₂. Cells were discarded at 25 passages, and were periodically
701 tested for Mycoplasma contamination using Universal Mycoplasma Detection kit
702 (ATCC). For fluorescence microscopy imaging experiments (Figures 1B, 2A and 3C),
703 cells were grown on 7×7-mm glass coverslips in 48-well plates. To improve cell
704 adherence, coverslips were pretreated with 50 µg/mL fibronectin (Millipore) for 20
705 min at 37 °C and washed once with Dulbecco's phosphate-buffered saline (DPBS),
706 pH 7.4. Cells used for generating lentivirus were grown on T25 plates, in MEM
707 supplemented as above, at 37 °C under 5% CO₂.

708

709 **Preparation of cell lines stably expressing APEX-fusion constructs**

710 To prepare lentivirus, one ~70% confluent T25 plate of HEK 293T cells, grown as
711 above, was co-transfected with 2.5 µg of APEX2 fusion plasmid, along with 0.25 µg
712 and 2.25 µg, respectively, of the lentivirus packaging plasmids VSV-G, and
713 dR8.91(Pagliarini et al. 2008). Transfection mixes used 10 µL Lipofectamine 2000
714 (Invitrogen) and were brought to a final volume of 2 mL with unsupplemented
715 MEM. The cells were transfected for 3 hours, after which media was replaced with 2
716 ml of fresh growth media with FBS. After 48 hours, the lentiviral supernatant was
717 collected by aspiration and filtered through a 0.45 µm syringe-mounted filter. This
718 filtered supernatant was immediately used to infect cells. HEK293T cells, grown in
719 6-well plates as described above, were infected at ~50% confluency, grown for 2
720 days, followed by selection in growth medium supplemented with 8 µg/mL
721 blasticidin for 7 days, before further analysis.

722

723 For the cells stably expressing HRP-KDEL, HEK293T cells at ~60%
confluency, grown in 6-well plates as described above, were transfected with the

724 mixture of 150 µg of plasmid and 10 µL Lipofectamine 2000 (Invitrogen) in
725 unsupplemented MEM for 3 hours, after which media was replaced with 2 ml of
726 fresh growth media with FBS. After 48 hours, the cells were trypsinized and
727 replated in T25 flask in growth medium supplemented with 1 µg/mL puromycin for
728 7 days, before further analysis.

729

730 ***In situ* biotinylation and crosslinking**

731 Stable-expression HEK 293T cells were grown to 90% confluency in 6-well plates, as
732 described above. For the crosslinking-then-BP biotinylation protocol ([Figure S1A](#),
733 [top](#)), cells were washed once with 5 mL PBS, and crosslinked in 5 mL 0.1% (v/v)
734 formaldehyde in PBS for 10 min at room temperature, with gentle agitation. The
735 crosslinking reaction was quenched by addition of glycine (1.2 M, in PBS) to final
736 concentration 125 mM, and gentle agitation for 5 minutes at room temperature.
737 Crosslinked cells were then washed three times with PBS and incubated with 500
738 µM biotin-phenol (BP) (Rhee et al. 2013) in PBS at room temperature, for 30 min.
739 Thereafter, H₂O₂ was added to a final concentration 1 mM, for 1 min. The liquid
740 phase was then removed by aspiration, and cells were washed twice with 2 mL
741 quenching solution (5 mM Trolox, 10 mM Ascorbate, 10 mM sodium azide, in PBS).
742 Crosslinked, labeled cells were collected by scraping, and pelleted by centrifugation,
743 and either processed immediately or flash frozen in liquid nitrogen and stored at -
744 80 °C before further analysis.

745 For the BP-then-crosslinking protocol ([Figure S1A, bottom](#)) used for mito-
746 APEX2 experiments ([Figure 1](#)), cell growth media was replaced with fresh media
747 supplemented with 500 µM BP. Cells were incubated in BP-supplemented media for
748 30 minutes at 37 °C, after which H₂O₂ was added to a final concentration of 1 mM.
749 After 1 min, the media was replaced with 5 mL crosslink-quench solution (0.1%
750 (v/v) formaldehyde, 10 mM ascorbate, and 5 mM Trolox, in PBS) for one minute, to
751 simultaneously quench the APEX2 BP labeling reaction and initiate formaldehyde
752 crosslinking. Thereafter, cells were washed and incubated in 5 mL of fresh crosslink-
753 quench for two additional 1-minute incubation steps, followed by a third, 8-minute
754 wash. Thereafter, crosslinking was terminated by the addition of Glycine, and cells
755 were harvested as described above.

756 The BP-quench-then-crosslinking protocol ([Figure S3A](#)) used for all other
757 subcellular compartments was identical to the BP-then-crosslinking protocol,
758 except that, following BP-labelling, and prior to the addition of crosslink-quench
759 solution, cells were incubated in 2 mL azide-free quenching solution (10 mM
760 ascorbate and 5 mM Trolox, in PBS) for one minute. Subsequently, cells were
761 subjected to only two (1 and 9 minute) treatments in crosslink-quench solution.
762 Thereafter, crosslinking was terminated by the addition of Glycine, and cells were
763 harvested as described above.

764

765 **Immunofluorescence staining and microscopy**

766 For immunofluorescence experiments ([Figures 1B, 2A, and 3C](#)), stable APEX- or
767 HRP-expressing cells were BP-labeled and crosslinked, as above, and subsequently
768 fixed with 4% (v/v) paraformaldehyde in PBS at room temperature for 10 min. Cells

769 were then washed with PBS three times and permeabilized with cold methanol at –
770 20 °C for 5 min. Cells were washed again three times with room-temperature PBS
771 and then incubated with primary antibodies in PBS-supplemented with 1% (w/v)
772 Bovine Serum Albumin (BSA)–for 1 h at room temperature. After washing three
773 times with PBS, cells were incubated with secondary antibodies and neutravidin-
774 AlexaFluor647 (1:1000 dilution) in BSA-supplemented PBS for 30 min. Cells were
775 then washed three times with PBS and imaged by confocal fluorescence microscopy,
776 or in PBS at 4 °C in light-tight containers prior to imaging. Primary and secondary
777 antibodies used were listed in [Table S2](#).

778 Fluorescence confocal microscopy was performed with a Zeiss AxioObserver
779 microscope with 63× oil-immersion objectives, outfitted with a Yokogawa spinning
780 disk confocal head, a Cascade II:512 camera, a Quad-band notch dichroic mirror
781 (405/488/568/647), 405 (diode), 491 (DPSS), 561 (DPSS) and 640 nm (diode)
782 lasers (all 50 mW). Alexa Fluor488 (491 laser excitation, 528/38 emission), Alexa
783 Fluor 568 (561 laser excitation, 617/73 emission), and AlexaFluor647 (640 laser
784 excitation, 700/75 emission) and differential interference contrast (DIC) images
785 were acquired through a 63x oil-immersion lens. Acquisition times ranged from 100
786 to 1,000 ms. For imaging quantitation and analysis, we used the SlideBook 6.0
787 software (Intelligent Imaging Innovations) to process and normalize the images.

788 The data in these figures ([Figure 1B, 2A, and 3C](#)) are representative of three
789 independent experiments with ≥ 5 fields of view each.

790

791 **Western and Streptavidin blotting**

792 For blotting experiments ([Figures 1C, 3D and S1D](#)), stable APEX- or HRP-expressing
793 cells were grown in 6-well plates. After labeling, the cells were harvested by
794 scraped, pelleted by centrifugation at 3,000×g for 10 min, and stored at –80 °C prior
795 to use. Thawed pellets were lysed by gentle pipetting in RIPA lysis buffer (50 mM
796 Tris, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100, 5 mM
797 EDTA), supplemented with 1× protease cocktail (Sigma Aldrich), 1 mM PMSF
798 (phenylmethylsulfonyl fluoride), for 5 min at 4 °C. Lysates were then clarified by
799 centrifugation at 15,000×g for 10 min at 4 °C before separation on homemade 8%
800 SDS-PAGE gels. Gels were transferred to nitrocellulose membranes, stained by
801 Ponceau S (0.1% (w/v) Ponceau S, 5% (v/v) acetic acid, in water) for 10 min at
802 room temperature, and imaged. The blots were then blocked with blocking buffer
803 (3% (w/v) BSA, 0.1% (v/v) Tween-20 in Tris-buffered saline) for 1 h at room
804 temperature, and incubated with primary antibodies in blocking buffer for 1 h more.
805 The dilutions of the antibodies are as followed: Mouse anti-V5 antibody (Life
806 Technologies) 1:1000 dilution and Mouse anti-FLAG antibody (Life Technologies)
807 1:800 dilution. Blots were rinsed four times for 5 min with wash buffer (0.1%
808 Tween-20 in Tris-buffered saline), and then immersed in blocking buffer
809 supplemented with Goat anti-Mouse IgG H + L-HRP Conjugate (1:3,000 dilution, Bio-
810 Rad), for 1 h at room temperature. Blots were rinsed four times for 5 min with wash
811 buffer, and developed with the Clarity reagent (Bio-Rad) and imaged on an Alpha
812 Innotech gel imaging system. Processing of streptavidin blots was similar.

813 Following Ponceau imaging, blots were blocked in blocking buffer for 30 min at
814 room temperature, immersed in blocking buffer supplemented with streptavidin-

815 HRP (1:3,000 dilution, ThermoFisher Scientific) at room temperature for 15 min,
816 rinsed with blocking buffer five times for 5 min each, developed and imaged using
817 the Clarity reagent and an Alpha Innotech gel imaging system.

818 The data in these experiments (Figures 1C, 3D and S1D) were also
819 reproduced for quality control prior to quantitative PCR and sequencing.

820

821 **Streptavidin bead enrichment of biotinylated material and RNA isolation**

822 Unless otherwise noted, all buffers used during RNA isolation were supplemented to
823 0.1 U/ μ L RNaseOUT (Thermo Fisher), 1 \times EDTA-free proteinase inhibitor cocktail
824 (Thermo Fisher) and 0.5 mM DTT, final. APEX- or HRP-expressing stable cells were
825 grown, labeled, crosslinked and harvested as described above. Labeled cell pellets
826 were lysed by incubation in 1 mL ice-cold RIPA buffer, supplemented with 10 mM
827 ascorbate and 5 mM Trolox, for 5 min at 4 °C with end-over-end agitation. Samples
828 were then sheared as described previously (Hendrickson et al. 2016) using a
829 Branson Digital Sonifier 250 (Emerson Industrial Automation) at 10% amplitude for
830 three 30-s intervals (0.7 s on + 1.3 s off), with 30-s resting steps between intervals.
831 Samples were held in ice-cold metal thermal blocks throughout sonication. Lysates
832 were then clarified by centrifugation at 15,000 $\times g$ for 5 min at 4 °C, moved to fresh
833 tubes and diluted with 1 mL Native lysis buffer (NLB: 25mM Tris pH 7.4, 150 mM
834 KCl, 0.5% NP-40, 5 mM EDTA), supplemented with ascorbate and trolox), each. For
835 each sample, 20% was removed as “input;” to the remainder was added 50 μ L of
836 streptavidin-coated magnetic bead slurry (Pierce) that had been equilibrated by two
837 washes in 1:1 RIPA:NLB. Samples were incubated for 2 h at 4 °C with end-over-end
838 agitation. Beads were subsequently washed with the following series of buffers (1
839 mL each, 5 min per wash, 4 °C, with gentle end-over-end agitation): (1) RIPA buffer,
840 supplemented with trolox and ascorbate, (2) RIPA buffer without radical quenchers,
841 (3) high salt buffer (1 M KCl, 50 mM Tris, pH 8.0, 5 mM EDTA), (4) urea buffer (2 M
842 Urea, 50 mM Tris, pH 8.0, 5 mM EDTA), (5) RIPA Buffer, (6) 1:1 RIPA: NLB, (7) NLB,
843 and (8) TE (10 mM Tris, pH 7.4, 1 mM EDTA).

844 Enriched RNAs were released from the beads by proteolysis in 100 μ L of
845 Elution Buffer (2% N-lauryl sarcoside, 10mM EDTA, 5mM DTT, in 1X PBS,
846 supplemented with 200 μ g proteinase K (Ambion) and 4 U RNaseOUT) at 42 °C for
847 1h, followed by 55 °C for 1h, as previously described (Hendrickson et al. 2016).
848 Eluted samples were cleaned up using Agencourt RNAClean XP magnetic beads
849 (Beckman Coulter), following the manufacturer’s 1.5 mL tube format protocol, and
850 eluted into 85 μ L H₂O. Thereafter, contaminating DNA was removed by digestion
851 with 5 U RQ1 RNase-free DNase I (Promega) in 100 μ L of the manufacturer’s
852 supplied buffer (1X final concentration) at 37 °C for 30 min. Purified RNAs were
853 again cleaned up using Agencourt RNAClean XP magnetic beads, as above, and
854 eluted into 30 μ L H₂O. The concentration and integrity of all samples was measured
855 using an Agilent 2100 Bioanalyzer, following the “RNA Nano” or “RNA Pico”
856 protocols, where appropriate. Samples were not heat-cooled prior to loading
857 Bioanalyzer chips.

858

859 **Quantitative RT-PCR**

860 For quantitative RT-PCR (qRT-PCR, [Figures S1A, S1C, 3E, and S3A](#)) RNA samples
861 were reverse transcribed using the SuperScript III Reverse Transcriptase kit
862 (ThermoFisher Scientific), priming with random hexamers (ThermoFisher
863 Scientific) according to the manufacturer's protocol. Samples were diluted with
864 water, mixed with gene specific primers ([Table S3](#)), and Rox-normalized FastStart
865 Universal SYBR Green Master Mix (Roche), and aliquotted into 384-well plates. qRT-
866 PCR was performed on an Applied Biosystems 7900HT Fast real time PCR
867 instrument, in quadruplicate. All threshold cycles (C_t , calculated per well) and
868 efficiencies (ϵ , calculated per primer pair), were calculated from "clipped" data,
869 using Real time qPCR Miner (Zhao and Fernald 2005). Raw C_t values were corrected
870 to account for the differences in sample volume, and percent yields were calculated
871 via the ΔC_t method:

$$872 \quad \text{yield} = 100 \times (1 + \epsilon)^{\Delta C_t}$$

$$873 \quad \dots \text{wherein, } \Delta C_t = {}^{corr}C_{t_{input}} - {}^{corr}C_{t_{RIP}}$$

874
875
876
877 Experimental uncertainties were calculated as described previously (Shechner et al.
878 2015). Given $D = A - B$, uncertainty was calculated using the formula:

$$879 \quad \sigma_D = \sqrt{(\sigma_A)^2 + (\sigma_B)^2}$$

880
881
882 ...wherein σ_A and σ_B are the measurement errors of A and B, respectively. For P, the
883 product or quotient of values A and B, uncertainty was calculated using the formula:

$$884 \quad \sigma_P = P \times \sqrt{\left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2}$$

885
886 The uncertainties of other functions, $f(x)$, were calculated using the first derivative
887 approximation:

$$888 \quad \sigma_{f(x)} = \sigma_x \times f'(x)$$

889
890 Sample sizes were determined in accordance with standard practices used in similar
891 experiments in the literature; no sample-size estimates were performed to ensure
892 adequate power to detect a prespecified effect size. Experiments were neither
893 randomized nor blinded to experimental conditions. Each samples contained four
894 replicates and no samples were excluded from analysis.

895 The experiments for [Figures S1A, S1C, 3E, and S3A](#) were performed once. For
896 statistical analysis on [Figure 3E](#), percent yield of 6 target genes were compared
897 against percent yield of 6 non-target genes using paired t-test for both HRP-KDEL
898 and ERM-APEX2. For comparison between ERM-APEX2 and APEX2-NES, 12 target
899 and non-target genes were compared against each other using paired t-test.

900

901 **Library preparation, sequencing, and quantification**

902 Purified RNA samples were depleted of ribosomal RNA using the Ribo-Zero Gold
903 rRNA removal kit (Illumina), using 1 μ L or 2 μ L of Ribo-Zero rRNA Removal Solution
904 for samples with ≤ 16 ng or ≥ 50 ng total input RNA mass, respectively, in 20 μ L final
905 reaction volume. Eluted RNA was cleaned up with Agencourt RNAClean XP beads
906 and eluted with 19.5 μ L of Elute, Prime, Fragment mix from the TruSeq RNA sample
907 preparation kit, v2 (Illumina). Thereafter, libraries were prepared using the TruSeq
908 RNA sample preparation kit, according to the manufacturer's instructions, starting
909 from "Incubate RFP" step. Each library was given a unique index during synthesis.
910 Library concentration and quality were confirmed on an Agilent 2100 Bioanalyzer,
911 using "DNA High Sensitivity" kits.

912 Indexed libraries were pooled in equimolar concentrations, with no more
913 than ten libraries per pool, and subjected to 50 cycles of paired end sequencing,
914 followed indexing, on two lanes of Illumina HiSeq 2500 flow cells, run in rapid mode
915 (Genomics Core, Broad Institute of Harvard and MIT).

916 In general, the experiments for each construct were performed in three
917 biological replicates. The mito-APEX experiment in [Figure S1B](#) and the mito-APEX2
918 negative control experiment (omit H₂O₂) in [Figure S1F](#) was performed in two
919 biological replicates.

920

921 **Quantification of RNA Abundances and Folds Enrichment; Assembly of True** 922 **positive and False positive lists**

923 Deep sequencing reads were mapped to human genome assembly hg19 and UCSC
924 known genes using TopHat2, set to default options(Kim et al. 2013). Transcript- and
925 gene-level abundances were quantified and Cuffdiff2, set to default options(Trapnell
926 et al. 2013), and 0.01 was added to all quantified values for ENCODE data.

927 Enrichment analysis (*e.g.* [Figures 1D, 3F, S1B, E, F, S2E, F, and S3B](#)), was restricted to
928 RNAs with FPKMs ≥ 1.0 following streptavidin pulldown. Fold enrichments were
929 calculated as follows:

930

$$931 \quad \text{Fold Enrichment} = \log_2 \left[\frac{FPKM_{\text{Post streptavidin enrichment}}}{FPKM_{\text{Pre streptavidin enrichment}}} \right]$$

932

933 To call significantly enriched genes from our data, ENCODE data(Dunham et
934 al. 2012), and ER-fractionation data(Reid and Nicchitta 2012), threshold cutoffs
935 were determined using Receiver Operator Characteristic (ROC) analysis(Fawcett
936 2006), employing sets of true-positive and false-positive genes identified as
937 described below. At each fold enrichment value of the data, true positive rate (TPR –
938 fraction of detected true positive genes above the fold enrichment value) and false
939 positive rate (FPR – fraction of detected false positive genes above the fold
940 enrichment value) are calculated. The fold enrichment value that maximizes TPR-
941 FPR is chosen as the fold enrichment cutoff. In mitochondrial and ER-associated
942 APEX-RIP experiments, ROC analysis was based on fold enrichment values; in the
943 nuclear-cytoplasmic partitioning experiment, it was based on calculated nuclear
944 preference scores.

945 Nuclear preference scores were calculated as follows. To each gene, i , we
946 assigned coordinates, $(x,y) = (\log_2 N_i, \log_2 C_i)$, where N_i and C_i denote to the fold-
947 enrichments from nuclear- (NLS) and cytosolic- (NES) APEX-RIP, respectively. In
948 this space, the line $y=x$ (or rather, $\log_2 N = \log_2 C$) corresponds to all genes that were
949 equally enriched in both nuclear and cytosolic APEX-RIP, and hence do not
950 preferentially reside in either compartment. The nuclear preference score for gene i
951 (NPS $_i$), is therefore defined as the minimum distance between its coordinates and
952 the line $\log_2 N = \log_2 C$. This is equivalent to calculating the distance between points
953 $(x_1,y_1) = (\log_2 N_i, \log_2 C_i)$ and $(x_2,y_2) = (0.5(\log_2 N_i + \log_2 C_i), 0.5(\log_2 N_i + \log_2 C_i))$. Hence:
954

$$955 \quad NPS_i = \sqrt{(\log_2 N_i - 0.5(\log_2 N_i + \log_2 C_i))^2 + (\log_2 C_i - 0.5(\log_2 N_i + \log_2 C_i))^2}$$

956

957 ...which reduces to:

$$958 \quad NPS_i = \left(\frac{1}{\sqrt{2}}\right)(\log_2 N_i - \log_2 C_i)$$

959

960 The true and false positive gene sets needed for ROC analysis were defined as
961 follows:

962 (1) For mitochondrial APEX-RIP, true positives corresponded to the thirteen
963 mitochondrial-encoded mRNAs; false positive RNAs corresponded to nuclear-
964 encoded long non-coding RNAs.

965 (2) For the nuclear and cytosolic partitioning experiment, true and false
966 positive gene lists were compiled using available ENCODE human cell line (NHEK-
967 Normal Human Epidermal Keratinocytes) nuclear-cytoplasmic fractionation
968 data (Dunham et al. 2012). We calculated fold-enrichments for RNAs in each
969 compartment (scaled relative to the whole cell RNA, [Figure S2A](#)), and used these
970 values to derive Nuclear Preference Scores, as described above. True positive and
971 true negative nuclear RNAs were then defined as the 1000 transcripts with the
972 highest and lowest NPSs, respectively ([Figure 2B](#); [Table S2 tab 4](#)). Using these gene
973 lists to perform ROC analysis on the original ENCODE data produced a significance
974 threshold cutoff at an NPS of 1.107 ([Figure S2B-D](#)), and lists of the 5467 and 10130
975 RNAs called as being enriched in the nucleus and cytoplasm, respectively ([Table S2](#),
976 [tabs 4 and 5](#)).

977 (3) For ER-APEX-RIP, most true positive genes were defined using data from
978 ER-localized proximity-dependent ribosome profiling (Jan, Williams, and Weissman
979 2014), corresponding to all RNAs with input RPKM ≥ 5.0 , input count ≥ 12 , and
980 $\log_2(\text{fold enrichment}) \geq 0.904$ ([Table S3 tab 4](#)). Additional true positive genes were
981 predicted by Phobius as having secretory signals, and were absent in
982 MitoCarta (Pagliarini et al. 2008). (Rhee et al. 2013) False positive RNAs included all
983 genes lacking secretory signals, as predicted by Phobius, SignalP, and TMHMM.

984

985 **Coverage and Specificity analysis of nuclear, cytosolic, and ER-proximal RNAs**

986 To estimate the coverage (recall) and specificity of APEX-RIP at each subcellular
987 compartment, we assembled lists of established target and off-target genes tailored
988 for that compartment.

989 For analysis of the nuclear–cytosolic datasets (Figure 2F), our reference
990 nuclear gene list comprised 827 lncRNAs with average RNA pre-enrichment
991 abundances of 1.0 or greater. Our reference off-target list comprised the set of 1260
992 ER-proximal RNAs defined using proximity-restricted ribosome profiling (Jan,
993 Williams, and Weissman 2014).

994 For analysis of the ER-proximal dataset (Figure 3I–J), our reference gene list
995 comprised 71 mRNAs encoding ER-resident proteins (Table S3 tab 5). Our reference
996 off-target list comprised (7589) RNAs lacking secretory annotation, as assessed
997 using Phobius (Käll, Krogh, and Sonnhammer 2004), TMHMM (Krogh et al. 2001),
998 SignalP (Petersen et al. 2011), and which lacked the GOCC terms “Endoplasmic
999 reticulum,” “Golgi,” “membrane,” and “extracellular” (Ashburner et al. 2000).

1000 For analysis of contaminants in ER datasets (Figure S3E), the RNAs that
1001 encode proteins with no predicted secretory annotation by Phobius, TMHMM, and
1002 SignalP and lacked GOCC terms “Endoplasmic reticulum,” “Golgi,” “membrane,” and
1003 “extracellular” were submitted to DAVID Bioinformatics analysis (Huang, Sherman,
1004 and Lempicki 2009) to find Gene ontology term enrichment against human
1005 background. Only terms with p-values less than 0.05 were shown.
1006

1007 **Supplementary Table S1: Genetic constructs used in this study**

Name	Features	Promoter/Vector	Details
Mito-V5-APEX	NotI-mito-BamHI-V5-APEX-XhoI	CMV/pCDNA3	Mito is a 24-amino acid mitochondrial targeting sequence (MTS) derived from COX4. V5: GKPIP NPLLGLDST
mito-V5-APEX2	mito-BamHI-V5-APEX2-NheI	CMV/pLX304	
Mito-GFP	NotI-mito-BamHI-BFP-XhoI	CMV/pCDNA3	
V5-APEX2-NLS	NotI-V5-APEX2-EcoRI-3xNLS-NheI	CMV	NLS: DPKKKRKV
FLAG-APEX2-NES	BstBI-FLAG-APEX2-NES-XhoI	CMV/pLX304	NES: LQLPPLERLTLD
HRP-V5-KDEL	NotI-IgK-HRP-V5-KDEL-IRES-puromycin-XbaI	CMV	IgK is N-terminal signaling sequence that brings protein to ER (METDTLLLWVLLLWVPGSTG D). KDEL is ER-retaining sequence
ERM-APEX2-V5	BstBI-ERM-APEX2-V5-NheI	CMV/pLX304	ERM is ER membrane targeting sequence derived from N-terminal 27 amino acids of rabbit P450 C1 (MDPVVVLGLCLSLLLSLWK QSYGGG)

1008

1009 **Supplementary Table 2. Antibodies used for immunofluorescence**

Antibody	Source	Company	Catalog number	Dilution	Associated Figure
Anti V5	Mouse	Life Technologies	R960-25	1:1000	Figure 1B, 2A, and 3A
Anti FLAG	Mouse	Agilent	200472	1:500	Figure 2A
Anti Mouse-AlexaFluor488	Goat	Life Technologies	A-11029	1:1000	Figure 1B, 2A, and 3A

1010

1011 **Supplementary Table 3. qRT-PCR primers used in this study**

Primer/probe name	Sequence (5'-3')
MT-ND1 forward	CACCTCTAGCCTAGCCGTTT
MT-ND1 reverse	CCGATCAGGGCGTAGTTTGA
MT-ND2 forward	CTTAAACTCCAGCACCACGAC
MT-ND2 reverse	AGCTTGTTTTAGGTGCGAGA
MT-ND3 forward	CCGCGTCCCTTTCTCCATAA
MT-ND3 reverse	AGGGCTCATGGTAGGGGTAA
MT-ND4 forward	ACAACACAATGGGGCTCACT
MT-ND4 reverse	CCGGTAATGATGTCGGGGTT
MT-ND4L forward	TCGCTCACACCTCATATCCTC
MT-ND4L reverse	AGGGCGCAAAGACTAGTATGG
MT-ND5 forward	TCCATTGTCGCATCCACCTT
MT-ND5 reverse	GGTTGTTTGGGTTGTGGCTC
MT-ND6 forward	GGGTTGAGGTCTTGGTGAGT
MT-ND6 reverse	ACCAATCCTACCTCCATCGC
MT-CYTB forward	TCTTGCACGAAACGGGATCA
MT-CYTB reverse	CGAGGGCGTCTTTGATTGTG
MT-COX1 forward	TCCTTATTGAGCCGAGCTG
MT-COX1 reverse	ACAAATGCATGGGCTGTGAC
MT-COX2 forward	AACCAAACCACTTTCACCGC
MT-COX2 reverse	CGATGGGCATGAAACTGTGG
MT-COX3 forward	CTAATGACCTCCGGCCTAGC
MT-COX3 reverse	AGGCCTAGTATGAGGAGCGT
MT-ATP6 forward	TTGCTTCATTTCATTGCCCC
MT-ATP6 reverse	GGGTGGTGATTAGTCGGTTGT
MT-ATP8 forward	ACTACCACCTACCTCCCTCAC
MT-ATP8 reverse	GGCAATGAATGAAGCGAACAGA
MT-RNR1 forward	CATCCCCGTTCCAGTGAGTT
MT-RNR1 reverse	TGGCTAGGCTAAGCGTTTTGA
MT-RNR2 forward	CAGCCGCTATTAAGGTTTCGT
MT-RNR2 reverse	AAGGCGCTTTGTGAAGTAGG
GAPDH forward	TTTCGACAGTCAGCCGCATCTTCTT
GAPDH reverse	GCCCAATACGACCAAATCCGTTGA
XIST forward	CCCTACTAGCTCCTCGGACA
XIST reverse	ACACATGCAGCGTGGTATCT
EMC10 forward	TTCATTGAGCGCCTGGAGAT
EMC10 reverse	TTCATTGAGCGCCTGGAGAT
PCSK1N forward	GAGACACCCGACGTGGAC
PCSK1N reverse	AATCCGTCAGCAAGTACC
SSR2 forward	GTTTGGGATGCCAACGATGAG
SSR2 reverse	CTCCACGGCGTATCTGTTCA
TMX1 forward	ACGGACGAGAAGTGGAGAGA
TMX1 reverse	ATTTTGACAAGCAGGGCACC
SFT2D2 forward	CCATCTCCTCATGGGACCAG
SFT2D2 reverse	GCAGAACACAGGGTAAGTGC
EPT1 forward	TGGCTTTCTGCTGGTTCGTAT
EPT1 reverse	AATCCAAACCCAGTCAGGCA
DRAP1 forward	ACATCCCACCTGAAGCAGTG
DRAP1 reverse	GATGCCACCAGGTCTTCAA
FAU forward	TCCTAAGGTGGCCAAACAGG
FAU reverse	GTGGGCACAACGTTGACAAA

SUB1 forward	CGTCACTTCCGGTTCTCTGT
SUB1 reverse	TGATTTAGGCATCGCTTCGC
LSM6 forward	CGGACGACCAGTTGTGGTAA
LSM6 reverse	CCAGGACCCCTCGATAATCC
COPS2 forward	AGGAGGACTACGACCTGGAAT
COPS2 reverse	GCCGCTTTTGGGTCATCTTC
CGGBP1 forward	GCCTCGTCCACTTTCCTAA
CGGBP1 reverse	TCATGCCCTTACGTAGGATCGAG
BCA53 forward	TCTTGCTGCTCCACAGTTT
BCA53 reverse	CAAACACCAAGGAGGGGTCT
CEP128 forward	TACAGTAATGGACAGGCGGG
CEP128 reverse	TCCGGAGTTGGTCGATTGAT
MAD1L1 forward	CGAGTCTGCCATCGTCCAA
MAD1L1 reverse	GCACTCTCCACCTGCTTCTT
RAD51B forward	TTTGGACGAAGCCCTGCAT
RAD51B reverse	CACAACCTGGTGGACCTGTA
RBPMS forward	ACAGTCGCTCAGAAGCAGAG
RBPMS reverse	CGAAGCGGATGCCATTCAAA
TCF7 forward	TCAACAGCCCACATCCCAC
TCF7 reverse	AGAGGCCTGTGAACTTGCTT
HOOK2 forward	TTTGCTGAAAAGGAAGCTGGA
HOOK2 reverse	GCAACTCCAGATCTGCCTCA
MAN2C1 forward	ATGAGGCCCAACAAGTTCCTG
MAN2C1 reverse	TCTCATAGGTGGCCTGGGAA

1012

1013 **References**

- 1014 Alán, Lukáš, Jaroslav Zelenka, Jan Ježek, Andrea Dlasková, and Petr Ježek. 2010.
1015 “Fluorescent in Situ Hybridization of Mitochondrial DNA and RNA.” *Acta*
1016 *Biochimica Polonica* 57 (4): 403–8.
1017 <http://www.ncbi.nlm.nih.gov/pubmed/21125028>.
- 1018 Ashburner, Michael, Catherine A. Ball, Judith A. Blake, David Botstein, Heather
1019 Butler, J. Michael Cherry, Allan P. Davis, et al. 2000. “Gene Ontology: Tool for the
1020 Unification of Biology.” *Nature Genetics* 25 (1): 25–29. doi:10.1038/75556.
- 1021 Bahar Halpern, Keren, Inbal Caspi, Doron Lemze, Maayan Levy, Shanie Landen, Eran
1022 Elinav, Igor Ulitsky, and Shalev Itzkovitz. 2015. “Nuclear Retention of mRNA in
1023 Mammalian Tissues.” *Cell Reports* 13 (12): 2653–62.
1024 doi:10.1016/j.celrep.2015.11.036.
- 1025 Batish, Mona, Arjun Raj, and Sanjay Tyagi. 2011. “Single Molecule Imaging of RNA In
1026 Situ.” In *RNA Detection and Visualization*, 3–13. doi:10.1007/978-1-61779-005-
1027 8_1.
- 1028 Cabili, Moran N, Margaret C Dunagin, Patrick D McClanahan, Andrew Biesch, Olivia
1029 Padovan-Merhar, Aviv Regev, John L Rinn, and Arjun Raj. 2015. “Localization
1030 and Abundance Analysis of Human lncRNAs at Single-Cell and Single-Molecule
1031 Resolution.” *Genome Biology* 16 (1): 20. doi:10.1186/s13059-015-0586-4.
- 1032 Chapman-Smith, A, and J E Cronan. 1999. “Molecular Biology of Biotin Attachment to
1033 Proteins.” *The Journal of Nutrition* 129 (2S Suppl): 477S–484S.
1034 <http://www.ncbi.nlm.nih.gov/pubmed/10064313>.
- 1035 Chen, Chun-Kan, Mario Blanco, Constanza Jackson, Erik Aznauryan, Noah Ollikainen,
1036 Christine Surka, Amy Chow, Andrea Cerase, Patrick McDonel, and Mitchell
1037 Guttman. 2016. “Xist Recruits the X Chromosome to the Nuclear Lamina to
1038 Enable Chromosome-Wide Silencing.” *Science (New York, N.Y.)*, August.
1039 doi:10.1126/science.aae0047.
- 1040 Chen, K. H., A. N. Boettiger, J. R. Moffitt, S. Wang, and X. Zhuang. 2015. “Spatially
1041 Resolved, Highly Multiplexed RNA Profiling in Single Cells.” *Science* 348 (6233):
1042 aaa6090-aaa6090. doi:10.1126/science.aaa6090.
- 1043 Chi, Sung Wook, Julie B. Zang, Aldo Mele, and Robert B. Darnell. 2009. “Argonaute
1044 HITS-CLIP Decodes microRNA–mRNA Interaction Maps.” *Nature*, June.
1045 doi:10.1038/nature08170.
- 1046 Csordás, György, Christian Renken, Péter Várnai, Ludivine Walter, David Weaver,
1047 Karolyn F. Buttle, Tamás Balla, Carmen A. Mannella, and György Hajnóczky.
1048 2006. “Structural and Functional Features and Significance of the Physical
1049 Linkage between ER and Mitochondria.” *The Journal of Cell Biology* 174 (7):
1050 915–21. doi:10.1083/jcb.200604016.
- 1051 Damas, Nkerorema Djodji, Michela Marcatti, Christophe Côme, Lise Lotte
1052 Christensen, Morten Muhlig Nielsen, Roland Baumgartner, Helene Maria
1053 Gylling, et al. 2016. “SNHG5 Promotes Colorectal Cancer Cell Survival by
1054 Counteracting STAU1-Mediated mRNA Destabilization.” *Nature*
1055 *Communications* 7 (December): 13875. doi:10.1038/ncomms13875.
- 1056 Derrien, T., R. Johnson, G. Bussotti, A. Tanzer, S. Djebali, H. Tilgner, G. Guernec, et al.
1057 2012. “The GENCODE v7 Catalog of Human Long Noncoding RNAs: Analysis of
1058 Their Gene Structure, Evolution, and Expression.” *Genome Research* 22 (9):

- 1059 1775–89. doi:10.1101/gr.132159.111.
- 1060 Dunham, Ian, Anshul Kundaje, Shelley F. Aldred, Patrick J. Collins, Carrie A. Davis,
1061 Francis Doyle, Charles B. Epstein, et al. 2012. “An Integrated Encyclopedia of
1062 DNA Elements in the Human Genome.” *Nature* 489 (7414): 57–74.
1063 doi:10.1038/nature11247.
- 1064 Engreitz, Jesse M., Noah Ollikainen, and Mitchell Guttman. 2016. “Long Non-Coding
1065 RNAs: Spatial Amplifiers That Control Nuclear Structure and Gene Expression.”
1066 *Nature Reviews Molecular Cell Biology* 17 (12): 756–70.
1067 doi:10.1038/nrm.2016.126.
- 1068 Fawcett, Tom. 2006. “An Introduction to ROC Analysis.” *Pattern Recognition Letters*
1069 27 (8): 861–74. doi:10.1016/j.patrec.2005.10.010.
- 1070 Fox, C H, F B Johnson, J Whiting, and P P Roller. 1985. “Formaldehyde Fixation.”
1071 *Journal of Histochemistry & Cytochemistry* 33 (8): 845–53.
1072 doi:10.1177/33.8.3894502.
- 1073 Gilbert, Chris, and Jesper Q Svejstrup. 2006. “RNA Immunoprecipitation for
1074 Determining RNA-Protein Associations in Vivo.” *Current Protocols in Molecular*
1075 *Biology / Edited by Frederick M. Ausubel ... [et Al.]* Chapter 27 (August): Unit
1076 27.4. doi:10.1002/0471142727.mb2704s75.
- 1077 Gilbert, Christopher, Arnold Kristjuhan, G Sebastiaan Winkler, and Jesper Q
1078 Svejstrup. 2004. “Elongator Interactions with Nascent mRNA Revealed by RNA
1079 Immunoprecipitation.” *Molecular Cell* 14 (4): 457–64.
1080 <http://www.ncbi.nlm.nih.gov/pubmed/15149595>.
- 1081 Han, Shuo, Namrata D Udeshi, Thomas J Deerinck, Tanya Svinkina, Mark H Ellisman,
1082 Steven A Carr, and Alice Y Ting. 2017. “Proximity Biotinylation as a Method for
1083 Mapping Proteins Associated with mtDNA in Living Cells.” *Cell Chemical Biology*
1084 24 (3): 404–14. doi:10.1016/j.chembiol.2017.02.002.
- 1085 Hendrickson, David G, David R Kelley, Danielle Tenen, Bradley Bernstein, and John L
1086 Rinn. 2016. “Widespread RNA Binding by Chromatin-Associated Proteins.”
1087 *Genome Biology* 17 (February): 28. doi:10.1186/s13059-016-0878-3.
- 1088 Hocine, Sami, Pascal Raymond, Daniel Zenklusen, Jeffrey A Chao, and Robert H
1089 Singer. 2013. “Single-Molecule Analysis of Gene Expression Using Two-Color
1090 RNA Labeling in Live Yeast.” *Nature Methods* 10 (2): 119–21.
1091 doi:10.1038/nmeth.2305.
- 1092 Huang, Da Wei, Brad T Sherman, and Richard A Lempicki. 2009. “Systematic and
1093 Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics
1094 Resources.” *Nature Protocols* 4 (1): 44–57. doi:10.1038/nprot.2008.211.
- 1095 Hung, Victoria, Stephanie S Lam, Namrata D Udeshi, Tanya Svinkina, Gaelen Guzman,
1096 Vamsi K Mootha, Steven A Carr, and Alice Y Ting. 2017. “Proteomic Mapping of
1097 Cytosol-Facing Outer Mitochondrial and ER Membranes in Living Human Cells
1098 by Proximity Biotinylation.” *eLife* 6 (April). doi:10.7554/eLife.24463.
- 1099 Hung, Victoria, Peng Zou, Hyun-Woo Rhee, Namrata D. Udeshi, Valentin Cracan,
1100 Tanya Svinkina, Steven A. Carr, Vamsi K. Mootha, and Alice Y. Ting. 2014.
1101 “Proteomic Mapping of the Human Mitochondrial Intermembrane Space in Live
1102 Cells via Ratiometric APEX Tagging.” *Molecular Cell* 55 (2): 332–41.
1103 doi:10.1016/j.molcel.2014.06.003.
- 1104 Ingolia, Nicholas T, Sina Ghaemmaghami, John R S Newman, and Jonathan S

- 1105 Weissman. 2009. "Genome-Wide Analysis in Vivo of Translation with
1106 Nucleotide Resolution Using Ribosome Profiling." *Science (New York, N.Y.)* 324
1107 (5924): 218–23. doi:10.1126/science.1168978.
- 1108 Jan, Calvin H, Christopher C Williams, and Jonathan S Weissman. 2014. "Principles of
1109 ER Cotranslational Translocation Revealed by Proximity-Specific Ribosome
1110 Profiling." *Science (New York, N.Y.)* 346 (6210): 1257521.
1111 doi:10.1126/science.1257521.
- 1112 Jung, Hosung, Christos G. Gkogkas, Nahum Sonenberg, and Christine E. Holt. 2014.
1113 "Remote Control of Gene Function by Local Translation." *Cell* 157 (1): 26–40.
1114 doi:10.1016/j.cell.2014.03.005.
- 1115 Käll, Lukas, Anders Krogh, and Erik L L Sonnhammer. 2004. "A Combined
1116 Transmembrane Topology and Signal Peptide Prediction Method." *Journal of*
1117 *Molecular Biology* 338 (5): 1027–36. doi:10.1016/j.jmb.2004.03.016.
- 1118 Kim, Daehwan, Geo Pertea, Cole Trapnell, Harold Pimentel, Ryan Kelley, and Steven
1119 L Salzberg. 2013. "TopHat2: Accurate Alignment of Transcriptomes in the
1120 Presence of Insertions, Deletions and Gene Fusions." *Genome Biology* 14 (4):
1121 R36. doi:10.1186/gb-2013-14-4-r36.
- 1122 Kind, Jop, and Bas van Steensel. 2010. "Genome–nuclear Lamina Interactions and
1123 Gene Regulation." *Current Opinion in Cell Biology* 22 (3): 320–25.
1124 doi:10.1016/j.ceb.2010.04.002.
- 1125 Krogh, A, B Larsson, G von Heijne, and E L Sonnhammer. 2001. "Predicting
1126 Transmembrane Protein Topology with a Hidden Markov Model: Application to
1127 Complete Genomes." *Journal of Molecular Biology* 305 (3): 567–80.
1128 doi:10.1006/jmbi.2000.4315.
- 1129 Lam, Stephanie S, Jeffrey D Martell, Kimberli J Kamer, Thomas J Deerinck, Mark H
1130 Ellisman, Vamsi K Mootha, and Alice Y Ting. 2014. "Directed Evolution of APEX2
1131 for Electron Microscopy and Proximity Labeling." *Nature Methods* 12 (1): 51–
1132 54. doi:10.1038/nmeth.3179.
- 1133 Lee, Je Hyuk, Evan R Daugharthy, Jonathan Scheiman, Reza Kalhor, Ryoji Amamoto,
1134 Derek T Peters, Brian M Turczyk, et al. 2014. "Highly Multiplexed Subcellular
1135 RNA Sequencing in Situ." *Science (New York, N.Y.)* 343 (March): 1360–63.
1136 doi:10.1126/science.1250212.
- 1137 Lesnik, Chen, Adi Golani-Armon, and Yoav Arava. 2015. "Localized Translation near
1138 the Mitochondrial Outer Membrane: An Update." *RNA Biology* 12 (8): 801–9.
1139 doi:10.1080/15476286.2015.1058686.
- 1140 Lobingier, Braden T., Ruth Hüttenhain, Kelsie Eichel, Kenneth B. Miller, Alice Y. Ting,
1141 Mark von Zastrow, and Nevan J. Krogan. 2017. "An Approach to
1142 Spatiotemporally Resolve Protein Interaction Networks in Living Cells." *Cell*
1143 169 (2): 350–360.e12. doi:10.1016/j.cell.2017.03.022.
- 1144 Loh, Ken H., Philipp S. Stawski, Austin S. Draycott, Namrata D. Udeshi, Emily K.
1145 Lehrman, Daniel K. Wilton, Tanya Svinkina, et al. 2016. "Proteomic Analysis of
1146 Unbounded Cellular Compartments: Synaptic Clefts." *Cell* 166 (5): 1295–1307.
1147 doi:10.1016/j.cell.2016.07.041.
- 1148 Martell, Jeffrey D, Thomas J Deerinck, Yasemin Sancak, Thomas L Poulos, Vamsi K
1149 Mootha, Gina E Sosinsky, Mark H Ellisman, and Alice Y Ting. 2012. "Engineered
1150 Ascorbate Peroxidase as a Genetically Encoded Reporter for Electron

- 1151 Microscopy." *Nature Biotechnology* 30 (11). Nature Publishing Group: 1143–48.
1152 doi:10.1038/nbt.2375.
- 1153 Martell, Jeffrey D, Masahito Yamagata, Thomas J Deerinck, Sébastien Phan, Carolyn G
1154 Kwa, Mark H Ellisman, Joshua R Sanes, and Alice Y Ting. 2016. "A Split
1155 Horseradish Peroxidase for the Detection of Intercellular Protein–protein
1156 Interactions and Sensitive Visualization of Synapses." *Nature Biotechnology* 34
1157 (7): 774–80. doi:10.1038/nbt.3563.
- 1158 Mercer, Tim R, Shane Neph, Marcel E Dinger, Joanna Crawford, Martin a Smith,
1159 Anne-Marie J Shearwood, Eric Haugen, et al. 2011. "The Human Mitochondrial
1160 Transcriptome." *Cell* 146 (4). Elsevier Inc.: 645–58.
1161 doi:10.1016/j.cell.2011.06.051.
- 1162 Mick, David U, Rachel B Rodrigues, Ryan D Leib, Christopher M Adams, Allis S Chien,
1163 Steven P Gygi, and Maxence V Nachury. 2015. "Proteomics of Primary Cilia by
1164 Proximity Labeling." *Developmental Cell* 35 (4): 497–512.
1165 doi:10.1016/j.devcel.2015.10.015.
- 1166 Motamedi, Mohammad R, André Verdel, Serafin U Colmenares, Scott A Gerber,
1167 Steven P Gygi, and Danesh Moazed. 2004. "Two RNAi Complexes, RITS and
1168 RDRC, Physically Interact and Localize to Noncoding Centromeric RNAs." *Cell*
1169 119 (6): 789–802. doi:10.1016/j.cell.2004.11.034.
- 1170 Nagao, Asuteka, Narumi Hino-Shigi, and Tsutomu Suzuki. 2008. "Chapter 23
1171 Measuring mRNA Decay in Human Mitochondria." In , 489–99.
1172 doi:10.1016/S0076-6879(08)02223-4.
- 1173 Nelles, David A, Mark Y Fang, Mitchell R O'Connell, Jia L Xu, Sebastian J Markmiller,
1174 Jennifer A Doudna, and Gene W Yeo. 2016. "Programmable RNA Tracking in
1175 Live Cells with CRISPR/Cas9." *Cell* 165 (2): 488–96.
1176 doi:10.1016/j.cell.2016.02.054.
- 1177 Paek, Jaeho, Marian Kalocsay, Dean P. Staus, Laura Wingler, Roberta Pascolutti, Joao
1178 A. Paulo, Steven P. Gygi, and Andrew C. Kruse. 2017. "Multidimensional
1179 Tracking of GPCR Signaling via Peroxidase-Catalyzed Proximity Labeling." *Cell*
1180 169 (2): 338–349.e11. doi:10.1016/j.cell.2017.03.028.
- 1181 Pagliarini, David J., Sarah E. Calvo, Betty Chang, Sunil A. Sheth, Scott B. Vafai, Shao-En
1182 Ong, Geoffrey A. Walford, et al. 2008. "A Mitochondrial Protein Compendium
1183 Elucidates Complex I Disease Biology." *Cell* 134 (1): 112–23.
1184 doi:10.1016/j.cell.2008.06.016.
- 1185 Paige, J. S., K. Y. Wu, and S. R. Jaffrey. 2011. "RNA Mimics of Green Fluorescent
1186 Protein." *Science* 333 (6042): 642–46. doi:10.1126/science.1207339.
- 1187 Petersen, Thomas Nordahl, Søren Brunak, Gunnar von Heijne, and Henrik Nielsen.
1188 2011. "SignalP 4.0: Discriminating Signal Peptides from Transmembrane
1189 Regions." *Nature Methods* 8 (10): 785–86. doi:10.1038/nmeth.1701.
- 1190 Piechota, Janusz, Rafał Tomecki, Kamil Gewartowski, Roman Szczesny, Aleksandra
1191 Dmochowska, Marek Kudła, Lien Dybczyńska, Piotr P Stepien, and Ewa Bartnik.
1192 2006. "Differential Stability of Mitochondrial mRNA in HeLa Cells." *Acta*
1193 *Biochimica Polonica* 53 (1): 157–68.
- 1194 Reid, David W, and Christopher V Nicchitta. 2012. "Primary Role for Endoplasmic
1195 Reticulum-Bound Ribosomes in Cellular Translation Identified by Ribosome
1196 Profiling." *The Journal of Biological Chemistry* 287 (8): 5518–27.

- 1197 doi:10.1074/jbc.M111.312280.
- 1198 Rhee, Hyun-Woo, Peng Zou, Namrata D Udeshi, Jeffrey D Martell, Vamsi K Mootha,
1199 Steven a Carr, and Alice Y Ting. 2013. "Proteomic Mapping of Mitochondria in
1200 Living Cells via Spatially Restricted Enzymatic Tagging." *Science (New York,
1201 N.Y.)* 339 (6125): 1328–31. doi:10.1126/science.1230593.
- 1202 Rinn, J., and M. Guttman. 2014. "RNA and Dynamic Nuclear Organization." *Science*
1203 345 (6202): 1240–41. doi:10.1126/science.1252966.
- 1204 Ro, Seungil, Hsiu-Yen Ma, Chanjae Park, Nicole Ortogero, Rui Song, Grant W Hennig,
1205 Huili Zheng, et al. 2013. "The Mitochondrial Genome Encodes Abundant Small
1206 Noncoding RNAs." *Cell Research* 23 (6): 759–74. doi:10.1038/cr.2013.37.
- 1207 Roux, Kyle J, Dae In Kim, Manfred Raida, and Brian Burke. 2012. "A Promiscuous
1208 Biotin Ligase Fusion Protein Identifies Proximal and Interacting Proteins in
1209 Mammalian Cells." *The Journal of Cell Biology* 196 (6): 801–10.
1210 doi:10.1083/jcb.201112098.
- 1211 Shechner, David M, Ezgi Haciosuleyman, Scott T Younger, and John L Rinn. 2015.
1212 "Multiplexable, Locus-Specific Targeting of Long RNAs with CRISPR-Display."
1213 *Nature Methods* 12 (7): 664–70. doi:10.1038/nmeth.3433.
- 1214 Sterne-Weiler, T., R. T. Martinez-Nunez, J. M. Howard, I. Cvitovik, S. Katzman, M. A.
1215 Tariq, N. Pourmand, and J. R. Sanford. 2013. "Frac-Seq Reveals Isoform-Specific
1216 Recruitment to Polyribosomes." *Genome Research* 23 (10): 1615–23.
1217 doi:10.1101/gr.148585.112.
- 1218 Trapnell, Cole, David G Hendrickson, Martin Sauvageau, Loyal Goff, John L Rinn, and
1219 Lior Pachter. 2013. "Differential Analysis of Gene Regulation at Transcript
1220 Resolution with RNA-Seq." *Nature Biotechnology* 31 (1): 46–53.
1221 doi:10.1038/nbt.2450.
- 1222 Ule, Jernej, Kirk B Jensen, Matteo Ruggiu, Aldo Mele, Aljaz Ule, and Robert B Darnell.
1223 2003. "CLIP Identifies Nova-Regulated RNA Networks in the Brain." *Science*
1224 (*New York, N.Y.*) 302 (5648): 1212–15. doi:10.1126/science.1090095.
- 1225 Wilk, Ronit, Jack Hu, Dmitry Blotsky, and Henry M Krause. 2016. "Diverse and
1226 Pervasive Subcellular Distributions for Both Coding and Long Noncoding
1227 RNAs." *Genes & Development* 30 (5): 594–609. doi:10.1101/gad.276931.115.
- 1228 Williams, Christopher C, Calvin H Jan, and Jonathan S Weissman. 2014. "Targeting
1229 and Plasticity of Mitochondrial Proteins Revealed by Proximity-Specific
1230 Ribosome Profiling." *Science (New York, N.Y.)* 346 (6210): 748–51.
1231 doi:10.1126/science.1257522.
- 1232 Wishart, J F, and B S Madhava Rao. 2010. *Recent Trends in Radiation Chemistry*.
1233 Singapore: World Scientific.
- 1234 Zhao, Sheng, and Russell D Fernald. 2005. "Comprehensive Algorithm for
1235 Quantitative Real-Time Polymerase Chain Reaction." *Journal of Computational*
1236 *Biology: A Journal of Computational Molecular Cell Biology* 12 (8): 1047–64.
1237 doi:10.1089/cmb.2005.12.1047.
- 1238