

1 **An antisense RNA capable of modulating the
2 expression of the tumor suppressor microRNA-34a**

3
4 **Jason T. Serviss^{1*}, Felix Clemens Richter^{1,2}, Jimmy Van den Eynden³,**
5 **Nathanael Andrews¹, Miranda Houtman^{1,4}, Mattias Vesterlund¹, Laura**
6 **Schwarzmueller^{1,5}, Per Johnsson^{6,7}, Erik Larsson³, Dan Grandér¹†, Katja**
7 **Pokrovskaja Tamm¹**

8
9 ¹ Department of Oncology and Pathology, Karolinska Institutet, Stockholm, Sweden, SE-
10 17177

11 ² Kennedy Institute of Rheumatology, University of Oxford, Roosevelt Drive, Oxford OX3 7FY,
12 UK

13 ³ Department of Medical Biochemistry and Cell Biology, Institute of Biomedicine, The
14 Sahlgrenska Academy, University of Gothenburg, SE-405 30 Gothenburg, Sweden

15 ⁴ Rheumatology Unit, Department of Medicine, Karolinska University Hospital, Solna,
16 Karolinska Institutet, Stockholm, Sweden

17 ⁵ Laboratory for Experimental Oncology and Radiobiology (LEXOR), Center for Experimental
18 Molecular Medicine (CEMM), Academic Medical Center, Amsterdam, The Netherlands

19 ⁶ Ludwig Institute for Cancer Research, Stockholm, Sweden

20 ⁷ Department of Cell and Molecular Biology, Karolinska Institutet, Stockholm, Sweden

21
22 *** Correspondence:**

23 Jason T. Serviss jason.serviss@ki.se

24
25 **Abstract**

26
27 The microRNA-34a is a well-studied tumor suppressor microRNA (miRNA)
28 and a direct downstream target of TP53 with roles in several pathways
29 associated with oncogenesis, such as proliferation, cellular growth, and
30 differentiation. Due to its broad tumor suppressive activity, it is not surprising
31 that *miR34a* expression is altered in a wide variety of solid tumors and
32 hematological malignancies. However, the mechanisms by which *miR34a* is
33 regulated in these cancers is largely unknown. In this study, we find that a
34 long non-coding RNA transcribed antisense to the *miR34a* host gene, is
35 critical for *miR34a* expression and mediation of its cellular functions in multiple
36 types of human cancer. We name this long non-coding RNA *lncTAM34a*, and
37 characterize its ability to facilitate *miR34a* expression under different types of
38 cellular stress in both *TP53* deficient and wild type settings.

40 **Introduction**

41 In recent years advances in functional genomics have revolutionized our
42 understanding of the human genome. Evidence now points to the fact that
43 approximately 75% of the genome is transcribed but only ~1.2% of this is
44 responsible for encoding proteins (International Human Genome Sequencing
45 Consortium 2004, Djebali et al. 2012). Of these recently identified elements,
46 long non-coding (lnc) RNAs are defined as transcripts exceeding 200 base
47 pairs (bp) in length with a lack of a functional open reading frame. Some
48 lncRNAs are dually classified as antisense (as) RNAs that are expressed from
49 the same locus as a sense transcript in the opposite orientation. Current
50 estimates using high-throughput transcriptome sequencing, indicate that up to
51 20-40% of the approximately 20,000 protein-coding genes exhibit antisense
52 transcription (Chen et al. 2004, Katayama et al. 2005, Ozsolak et al. 2010).

53 Systematic large-scale studies have shown aberrant expression of asRNAs to
54 be associated with tumorigenesis (Balbin et al. 2015) and, although
55 characterization of several of these has identified asRNA-mediated regulation
56 of multiple well known tumorigenic factors (Yap et al. 2010, Johnsson et al.
57 2013), the vast majority of potential tumor-associated lncRNAs have not yet
58 been characterized. The known mechanisms by which asRNAs accomplish
59 their regulatory functions are diverse, and include recruitment of chromatin
60 modifying factors (Rinn et al. 2007, Johnsson et al. 2013), acting as
61 microRNA (miRNA) sponges (Memczak et al. 2013), and causing
62 transcriptional interference (Conley et al. 2012).

63

64 Responses to cellular stress, e.g. DNA damage, sustained oncogene

65 expression, and nutrient deprivation, are all tightly controlled cellular pathways
66 that are almost universally dysregulated in cancer. Cellular signaling, in
67 response to these types of stresses, often converges on the transcription
68 factor TP53 that regulates transcription of coding and non-coding downstream
69 targets. One important non-coding target of TP53 is the tumor suppressor
70 miRNA known as *miR34a* (Raver-Shapira et al. 2007).
71 Upon TP53 activation *miR34a* expression is increased allowing it to down-
72 regulate target genes involved in cellular pathways such as growth factor
73 signaling, apoptosis, differentiation, and cellular senescence (Lal et al. 2011,
74 Slabakova et al. 2017). Thus, *miR34a* is a crucial factor in mediating activated
75 TP53 response and, the fact that it is often deleted or down-regulated in
76 human cancers indicates, its tumor suppressive effect and makes it a valuable
77 prognostic marker (Cole et al. 2008, Gallardo et al. 2009, Zenz et al. 2009,
78 Cheng et al. 2010, Liu et al. 2011). Reduced *miR34a* transcription is mediated
79 via epigenetic regulation in many solid tumors, including colorectal-,
80 pancreatic-, and ovarian cancer (Vogt et al. 2011), as well as numerous types
81 of hematological malignancies (Chim et al. 2010). In addition, *miR34a* has
82 been shown to be transcriptionally regulated via TP53 homologs, TP63 and
83 TP73, other transcription factors, e.g. STAT3 and MYC, and, in addition, post-
84 transcriptionally through miRNA sponging by the NEAT1 lncRNA (Chang et al.
85 2008, Su et al. 2010, Agostini et al. 2011, Rokavec et al. 2015, Ding et al.
86 2017). Despite these findings, the mechanisms underlying *miR34a* regulation
87 in the context of oncogenesis have not yet been fully elucidated.

88

89 Studies across multiple cancer types have reported a decrease in oncogenic
90 phenotypes when *miR34a* expression is induced in a *TP53*-null background,
91 although endogenous mechanisms for achieving this have not yet been
92 discovered (Liu et al. 2011, Ahn et al. 2012, Yang et al. 2012, Stahlhut et al.
93 2015, Wang et al. 2015). In addition, previous reports from large-scale studies
94 interrogating global *TP53*-mediated regulation of lncRNAs have identified a
95 lncRNA (known as RP3-510D11.2 and LINC01759) originating in the
96 antisense orientation from the *miR34a* locus that is induced upon numerous
97 forms of cellular stress (Rashi-Elkeles et al. 2014, Hunten et al. 2015, Leveille
98 et al. 2015, Ashouri et al. 2016, Kim et al. 2017). Despite this, none of these
99 studies have functionally characterized this transcript, which we name *Long-*
100 *Non-Coding Transcriptional Activator of MiR34a* (lncTAM34a). In this study
101 we functionally characterize the *lncTAM34a* transcript, and find that it
102 positively regulates *miR34a* expression resulting in a decrease of several
103 tumorigenic phenotypes. Furthermore, we find that *lncTAM34a*-mediated up-
104 regulation of *miR34a* is sufficient to induce endogenous cellular mechanisms
105 counteracting several types of stress stimuli in a *TP53*-deficient background.
106 Finally, similar to the functional roles of antisense transcription at protein-
107 coding genes, we identify a rare example of an antisense RNA capable of
108 regulating a cancer-associated miRNA.

109

110 **Results**

111

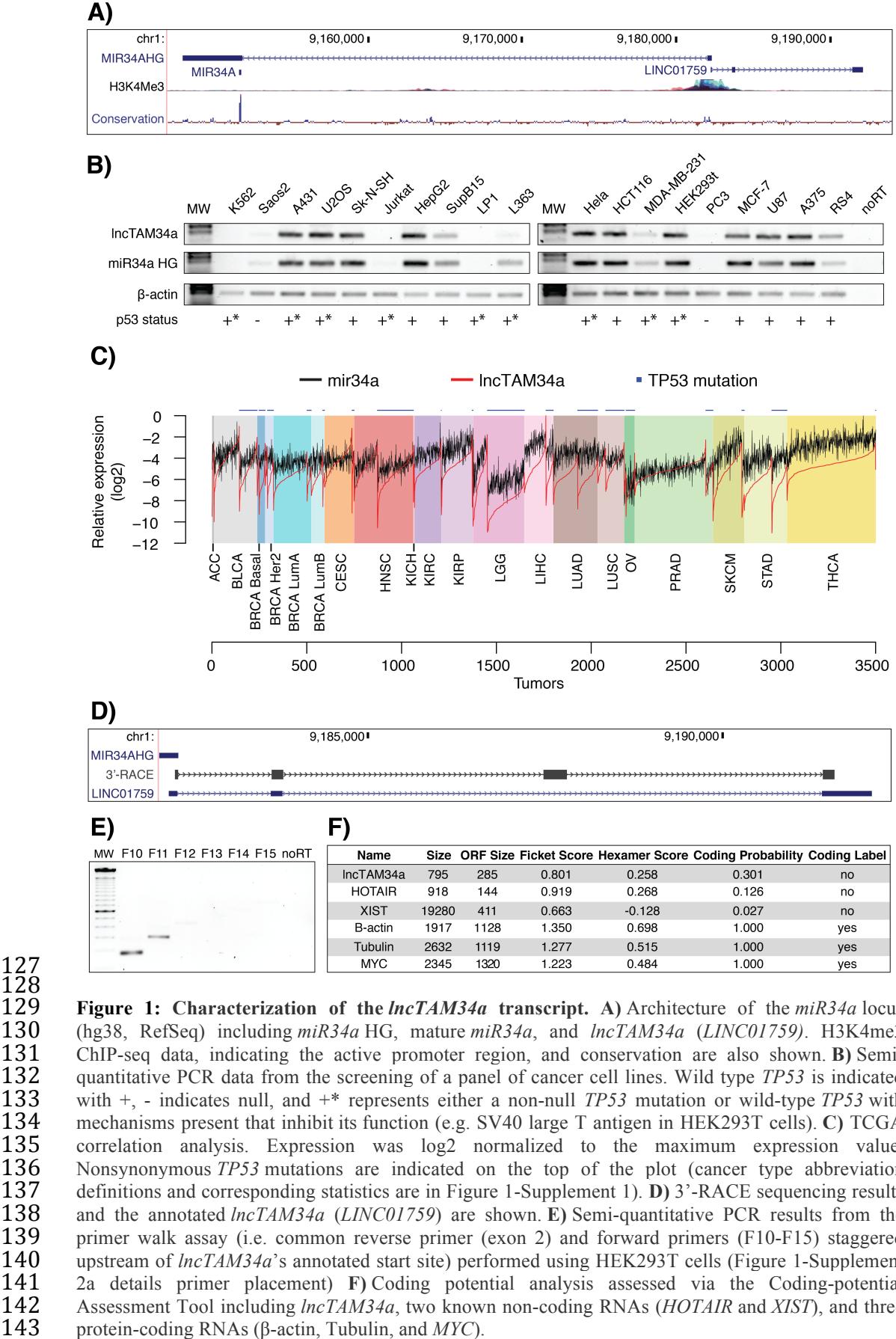
112 ***lncTAM34a* is a broadly expressed non-coding transcript whose levels**
113 **correlate with *miR34a* expression**

114

115 *lncTAM34a* is transcribed in a “head-to-head” orientation with approximately
116 100 base pair overlap with the *miR34a* host gene (HG) (**Fig. 1a**). Due to the

117 fact that sense/antisense pairs can be both concordantly and discordantly
118 expressed, we sought to evaluate this relationship in the case of *miR34a* HG
119 and its asRNA. Using a diverse panel of cancer cell lines, we detected co-
120 expression of both the *miR34a* HG and *IncTAM34a* (**Fig. 1b**). We used cell
121 lines with a known *TP53* status in the panel due to previous reports
122 that *miR34a* and *IncTAM34a* are known downstream targets of TP53. These
123 results indicate that *miR34a* HG and *IncTAM34a* are co-expressed and that
124 their expression levels correlate with *TP53* status, with *TP53*^{-/-} cells tending to
125 have decreased or undetectable expression of both transcripts.

126



144 We next sought to analyze primary cancer samples to examine whether a
145 correlation between *lncTAM34a* and *miR34a* expression levels could be
146 identified. We utilized RNA sequencing data from The Cancer Genome Atlas
147 (TCGA) after stratifying patients by cancer type, *TP53* status, and, in the case
148 of breast cancer, cancer subtypes. The results indicate that *lncTAM34a*
149 and *miR34a* expression are strongly correlated in the vast majority of cancer
150 types examined, both in the presence and absence of wild-type *TP53* (**Fig.**
151 **1c, Figure 1-Figure Supplement 1a**). The results also further confirm that
152 the expression levels of both *miR34a* and *lncTAM34a* are significantly
153 reduced in patients with nonsynonymous *TP53* mutations (**Figure 1-Figure**
154 **Supplement 1b**).

155

156 Next, we aimed to gain a thorough understanding of *lncTAM34a*'s molecular
157 characteristics and cellular localization. To experimentally determine the 3'
158 termination site for the *lncTAM34a* transcript we performed 3' rapid
159 amplification of cDNA ends (RACE) using the U2OS osteosarcoma cell line
160 that exhibited high endogenous levels of *lncTAM34a* in the cell panel
161 screening. Sequencing the cloned cDNA indicated that the transcripts 3'
162 transcription termination site is 525 bp upstream of the *lncTAM34a* transcript's
163 annotated termination site (**Fig. 1d**). Next, we characterized the *lncTAM34a* 5'
164 transcription start site by carrying out a primer walk assay, i.e. a common
165 reverse primer was placed in exon 2 and forward primers were gradually
166 staggered upstream of *lncTAM34a*'s annotated start site (**Figure 1-Figure**
167 **Supplement 2a**). Our results indicated that the 5' start site for *lncTAM34a* is
168 in fact approximately 90 bp (F11 primer) to 220 bp (F12 primer) upstream of

169 the annotated start site (**Fig. 1e**). Polyadenylation status was evaluated via
170 cDNA synthesis with either random nanomers or oligo(DT) primers followed
171 by semi-quantitative PCR which showed that *lncTAM34a* is polyadenylated
172 although the unspliced form seems to only be present in a polyadenylation
173 negative state (**Figure 1-Figure Supplement 2b**). Furthermore, we
174 investigated the propensity of *lncTAM34a* to be alternatively spliced in U2OS
175 cells, using PCR cloning followed by sequencing and found that the transcript
176 is post-transcriptionally spliced to form multiple isoforms (**Figure 1-Figure**
177 **Supplement 2c**). In order to evaluate the subcellular localization of
178 *lncTAM34a*, we made use of RNA sequencing data from five cancer cell lines
179 included in the ENCODE (Encode Project Consortium 2012) project that had
180 been fractionated into cytosolic and nuclear fractions. The analysis revealed
181 that the *lncTAM34a* transcript primarily localizes to the nucleus with only a
182 minor fraction in the cytosol (**Figure 1-Figure Supplement 2d**).

183

184 Lastly, we utilized several approaches to evaluate the coding potential of
185 the *lncTAM34a* transcript. The Coding-Potential Assessment Tool is a
186 bioinformatics-based tool that uses a logistic regression model to evaluate
187 coding-potential by examining open reading frame (ORF) length, ORF
188 coverage, Fickett score, and hexamer score (Wang et al. 2013). Results
189 indicated that *lncTAM34a* has a similar low coding capacity to known non-
190 coding transcripts such as *HOTAIR* and *XIST* (Fig. 1F). We further confirmed
191 these results using the Coding-Potential Calculator that uses a support vector
192 machine-based classifier and accesses an alternate set of discriminatory
193 features (**Figure 1-Figure Supplement 2e**) (Kong et al. 2007). Finally, we

194 downloaded mass spectrometry spectra for 11 cancer cell lines (Geiger et al.
195 2012), 7 of which were also present in the cell line panel above (**Fig. 1b**), and
196 searched it against a database of human protein sequences which also
197 contained the 6 frame translation of *lncTAM34a*. However, we did not manage
198 to detect any peptides matching the sequence in any of the 11 cell lines.
199 Taken together our results indicate that *lncTAM34a* is not a coding transcript
200 and that it is not translated to any significant degree.

201

202 **TP53-mediated regulation of *lncTAM34a* expression**

203 *miR34a* is a known downstream target of TP53 and has been previously
204 shown to exhibit increased expression within multiple contexts of cellular
205 stress. Several global analyses of TP53-regulated lncRNAs have also shown
206 *lncTAM34a* to be induced upon TP53 activation (Rashi-Elkeles et al. 2014,
207 Hunten et al. 2015, Leveille et al. 2015, Ashouri et al. 2016, Kim et al. 2017).
208 To confirm these results in our biological systems, we treated HEK293T,
209 embryonic kidney cells, and HCT116, colorectal cancer cells, with the DNA
210 damaging agent doxorubicin to activate TP53. QPCR-mediated
211 measurements of both *miR34a* HG and *lncTAM34a* indicated that their
212 expression levels were increased in response to doxorubicin treatment in both
213 cell lines (**Fig. 2a**). To assess whether TP53 was responsible for the increase
214 in *lncTAM34a* expression upon DNA damage, we treated *TP53*^{+/+} and *TP53*^{-/-}
215 HCT116 cells with increasing concentrations of doxorubicin and monitored the
216 expression of both *miR34a* HG and *lncTAM34a*. We observed a dose-
217 dependent increase in both *miR34a* HG and *lncTAM34a* expression levels
218 with increasing amounts of doxorubicin, revealing that these two transcripts

219 are co-regulated, although, this effect was largely abrogated in *TP53*^{-/-} cells
220 (**Fig. 2b**). These results indicate that TP53 activation increases *lncTAM34a*
221 expression upon DNA damage. Nevertheless, *TP53*^{-/-} cells also showed a
222 dose-dependent increase in both *miR34a* HG and *lncTAM34a*, suggesting
223 that additional factors, other than *TP53* are capable of initiating an increase in
224 expression of both of these transcripts upon DNA damage.

225

226
227
228
229
230
231
232
233
234
235
236
237
238
239

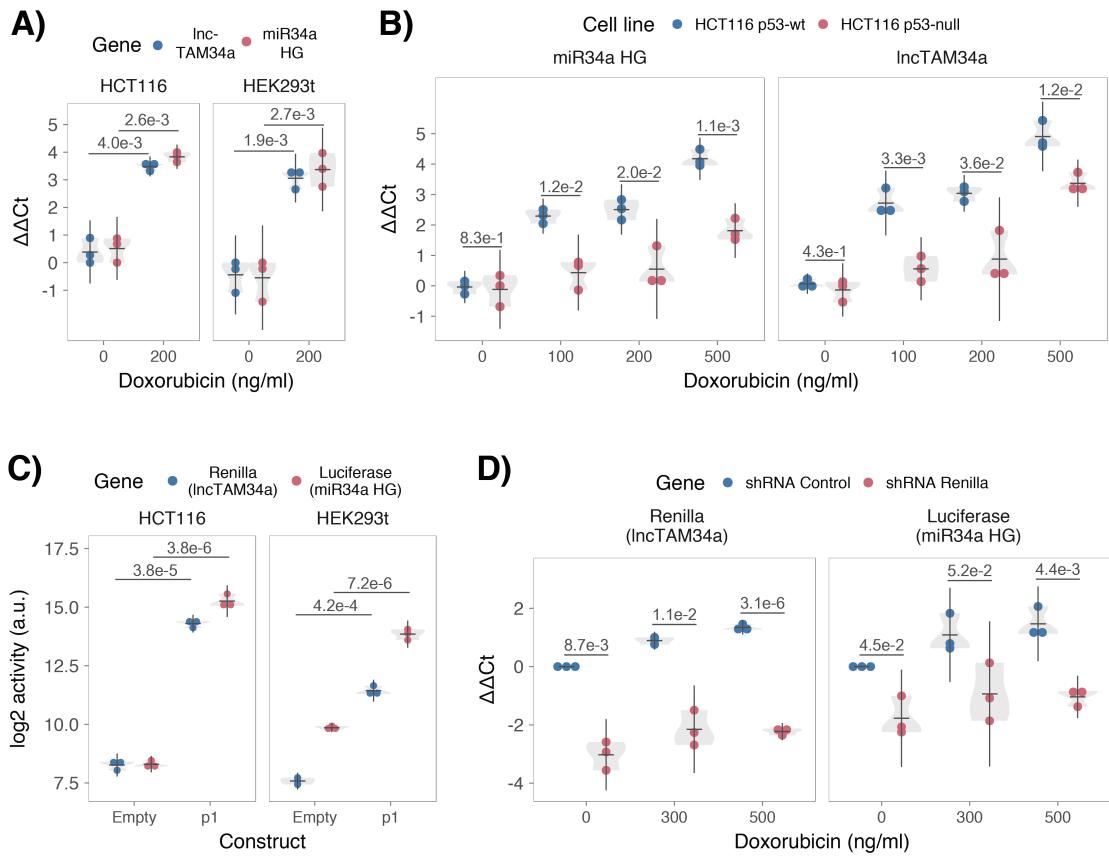


Figure 2: TP53-mediated regulation of the miR34a locus. **A)** Evaluating the effects of 24 hours of treatment with 200 ng/ml doxorubicin on *LncTAM34a* and *miR34a* HG in HCT116 and HEK293T cells.* **B)** Monitoring *miR34a* HG and *LncTAM34a* expression levels during 24 hours of doxorubicin treatment in *TP53*^{+/+} and *TP53*^{-/-} HCT116 cells.* **C)** Quantification of luciferase and renilla levels after transfection of HCT116 and HEK293T cells with the p1 construct (Figure 2-Supplement 2 contains a schematic representation of the p1 construct).* **D)** HCT116 cells were co-transfected with the p1 construct and shRNA renilla or shRNA control and subsequently treated with increasing doses of doxorubicin. 24 hours post-treatment, cells were harvested and renilla and luciferase levels were measured using QPCR.* Individual points represent results from independent experiments and the gray shadow indicates the density of those points. Error bars show the 95% CI, black horizontal lines represent the mean, and P values are shown over long horizontal lines indicating the comparison tested. All experiments in Figure 2 were performed in biological triplicate.

240 The head-to-head orientation of *miR34a* HG and *lncTAM34a*, suggests that
241 transcription is initiated from a single promoter in a bi-directional manner (**Fig**
242 **1a**). To investigate whether *miR34a* HG and *lncTAM34a* are transcribed from
243 the same promoter as divergent transcripts, we cloned the previously reported
244 *miR34a* HG promoter, a ~300 bp region including the TP53 binding site and
245 the majority of the first exon of both transcripts, into a luciferase/renilla dual
246 reporter vector (**Figure 2-Figure Supplement 1a-b**) (Raver-Shapira et al.
247 2007). We hereafter refer to this construct as p1. Upon transfection of p1 into
248 HCT116 and HEK293T cell lines we observed increases in both luciferase
249 and renilla indicating that *miR34a* HG and *lncTAM34a* expression can be
250 regulated by a single promoter contained within the p1 construct (**Fig. 2c**).
251

252 ***lncTAM34a* facilitates *miR34a* induction in response to DNA damage**
253 We hypothesized that *lncTAM34a* may regulate *miR34a* HG levels and, in
254 addition, that the overlapping regions of the sense and antisense transcripts
255 may mediate this regulation. Knockdown of endogenous *lncTAM34a* is
256 complicated by its various isoforms (**Figure 1-Figure Supplement 2c**). For
257 this reason, we utilized the p1 construct to evaluate the regulatory role of
258 *lncTAM34a* on *miR34a* HG. Accordingly, we first co-transfected the p1
259 construct, containing the overlapping region of the two transcripts, and two
260 different short hairpin (sh) RNAs targeting renilla into HEK293T cells and
261 subsequently measured luciferase and renilla expression. The results
262 indicated that shRNA-mediated knock-down of the p1-renilla transcript
263 (corresponding to *lncTAM34a*) caused p1-luciferase (corresponding
264 to *miR34a* HG) levels to concomitantly decrease (**Figure 2-Figure**

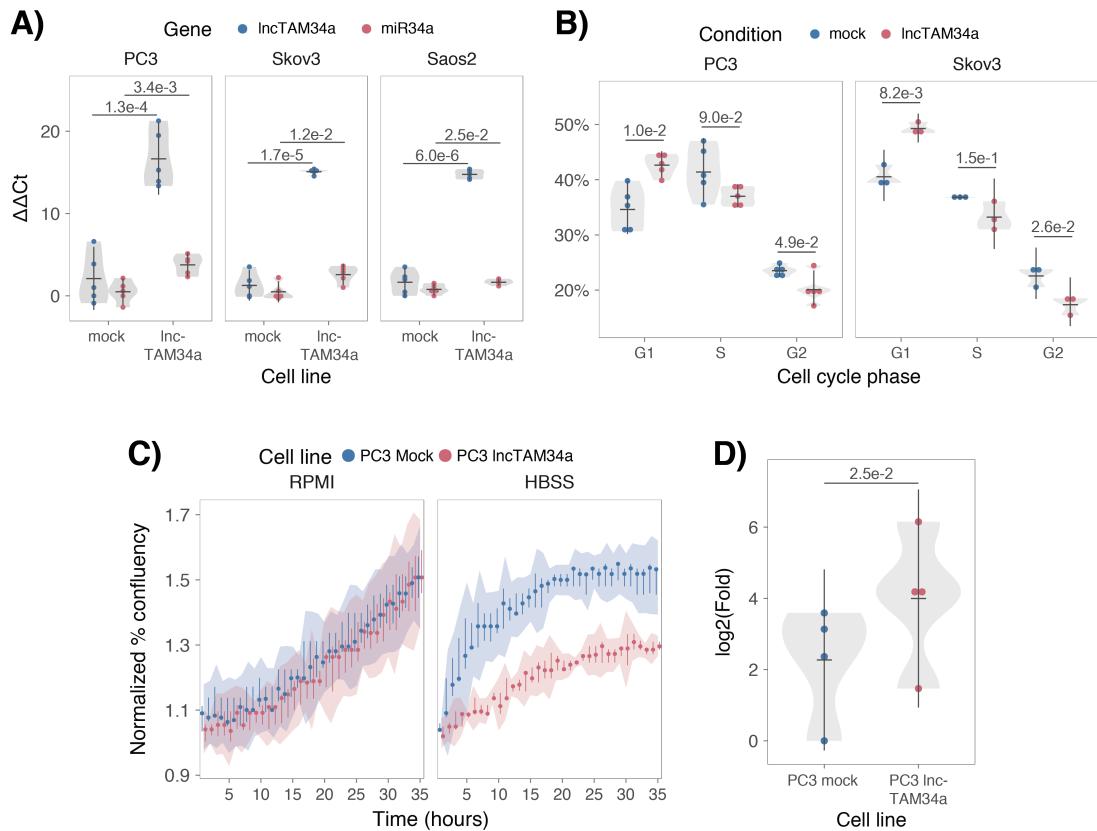
265 **Supplement 2).** The results suggest that *IncTAM34a* positively regulates
266 levels of *miR34a* HG and that the transcriptional product of *IncTAM34a* within
267 the p1 construct contributes to inducing a *miR34a* response. To further
268 support these conclusions and better understand the role of *IncTAM34a*
269 during TP53 activation, *TP53^{+/+}* HCT116 cells were co-transfected with p1
270 and shRNA renilla (2.1) and subsequently treated with increasing doses of
271 doxorubicin. Again, the results showed a concomitant reduction in luciferase
272 levels upon knock-down of p1-renilla i.e. the *IncTAM34a* corresponding
273 segment of the p1 transcript (**Fig. 2d**). Furthermore, the results showed that in
274 the absence of p1-renilla the expected induction of p1-luciferase in response
275 to TP53 activation by DNA damage is abrogated. Collectively these results
276 indicate that *IncTAM34a* positively regulates *miR34a* expression and
277 furthermore, suggests that it is crucial for an appropriate TP53-
278 mediated *miR34a* response to DNA damage.

279

280 ***IncTAM34a* can regulate *miR34a* host gene independently of TP53**

281 Despite the fact that TP53 regulates *miR34a* HG and *IncTAM34a* expression,
282 our results showed that other factors are also able to regulate this locus (**Fig.**
283 **2b**). Utilizing a lentiviral system, we stably over-expressed the *IncTAM34a*
284 transcript in three *TP53*-null cell lines, PC3 (prostate cancer), Saos2
285 (osteogenic sarcoma), and Skov3 (ovarian adenocarcinoma). We first
286 analyzed the levels of *IncTAM34a* in these stable cell lines, compared to
287 HEK293T cells, which have high endogenous levels of *IncTAM34a*. On
288 average, the over-expression was approximately 30-fold higher in the over-
289 expression cell lines than in HEK293T cells, roughly corresponding to

290 physiologically relevant levels in cells encountering a stress stimulus, such as
291 DNA damage (**Figure 3-Figure Supplement 1**). Analysis of *miR34a* levels in
292 the *lncTAM34a* over-expressing cell lines showed that this over-expression
293 resulted in a concomitant increase in the expression of *miR34a* in all three cell
294 lines (**Fig. 3a**). These results indicate that, in the absence of
295 *TP53*, *miR34a* expression may be rescued by activating *lncTAM34a*
296 expression.



297

298 **Figure 3: *IncTAM34a* positively regulates *miR34a* and its associated phenotypes.** A) QPCR-
299 mediated quantification of *miR34a* expression in cell lines stably over-expressing *IncTAM34a*.* B)
300 Cell cycle analysis comparing stably over-expressing *IncTAM34a* cell lines to the respective mock
301 control.* C) Analysis of cellular growth over time in *IncTAM34a* over-expressing PC3 cells. Points
302 represent the median from 3 independent experiments, the colored shadows indicate the 95%
303 confidence interval, and vertical lines show the minimum and maximum values obtained from the three
304 experiments. D) Differential phosphorylated polymerase II binding in *IncTAM34a* over-expressing PC3
305 cells.* Individual points represent results from independent experiments and the gray shadow
306 indicates the density of those points. Error bars show the 95% CI, black horizontal lines represent the
307 mean, and *P* values are shown over long horizontal lines indicating the comparison tested.
308

309 *miR34a* has been previously shown to regulate cell cycle progression, with
310 *miR34a* induction causing G1 arrest (Raver-Shapira et al. 2007, Tarasov et al.
311 2007). Cell cycle analysis via determination of DNA content showed a
312 significant increase in G1 phase cells and a concomitant decrease in G2
313 phase cells in the PC3 and Skov3 *IncTAM34a* over-expressing cell lines,
314 indicating G1 arrest (**Fig. 3b**). The effects of *miR34a* on the cell cycle are
315 mediated by its ability to target cell cycle regulators such as cyclin D1
316 (*CCND1*) (Sun et al. 2008). Quantification of both *CCND1* RNA expression
317 (**Figure 3-Figure Supplement 2a**) and protein levels (**Figure 3-Figure**
318 **Supplement 2b**) in the PC3 *IncTAM34a* over-expressing cell line showed a
319 significant decrease of *CCND1* levels compared to the mock control.
320 Collectively, these results indicate that *IncTAM34a*-mediated induction of
321 *miR34a* is sufficient to result in the corresponding *miR34a*-directed effects on
322 cell cycle.

323

324 *miR34a* is also a well-known inhibitor of cellular growth via its ability to
325 negatively regulate growth factor signaling. Furthermore, starvation has been
326 shown to induce *miR34a* expression causing inactivation of numerous pro-
327 survival growth factors (Lal et al. 2011). We further interrogated the effects
328 of *IncTAM34a* over-expression by monitoring the growth of the PC3 stable cell
329 lines in both normal and starvation conditions via confluence measurements
330 over a 35-hour period. Under normal growth conditions there is a small but
331 significant reduction ($P = 3.0\text{e-}8$; linear regression, **Fig. 3c**) in confluence in
332 the *IncTAM34a* over-expressing cell lines compared to mock control.
333 However, these effects on cell growth are drastically increased in starvation

334 conditions ($P = 9.5e-67$; linear regression; **Fig. 3c**). This is in agreement with
335 our previous results, and suggests that *lncTAM34a*-mediated increases
336 in *miR34a* expression are crucial under conditions of stress and necessary for
337 the initiation of an appropriate cellular response. In summary, we find that
338 over-expression of *lncTAM34a* is sufficient to increase *miR34a* expression
339 and gives rise to known phenotypes observed upon induction of *miR34a*.

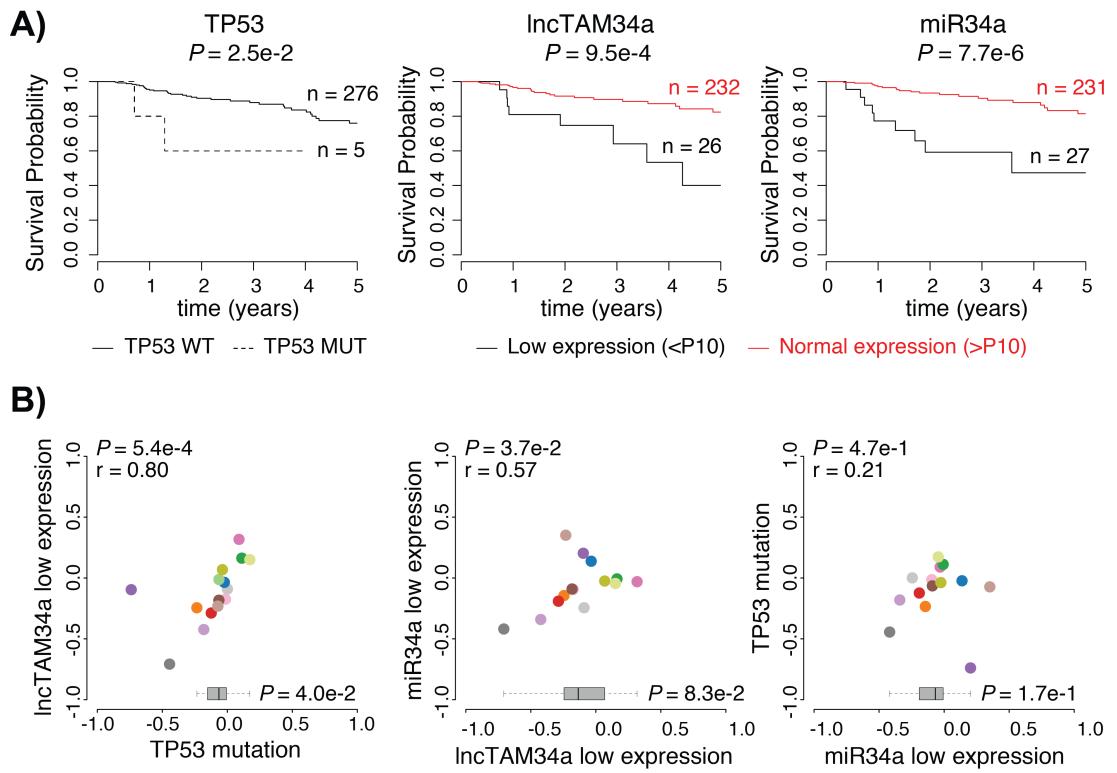
340

341 ***lncTAM34a* transcriptionally activates *miR34a* host gene**

342 Antisense RNAs have been reported to mediate their effects both via
343 transcriptional and post-transcriptional mechanisms. Due to the fact that
344 *miR34a* expression is undetected in wild type PC3 cells (**Fig. 1b**) but, upon
345 over-expression of *lncTAM34a*, increases to detectable levels, we
346 hypothesized that *lncTAM34a* is capable of regulating *miR34a* expression via
347 a transcriptional mechanism. To ascertain if this is actually the case, we
348 performed chromatin immunoprecipitation (ChIP) for phosphorylated
349 polymerase II (polII) at the *miR34a* HG promoter in both *lncTAM34a* over-
350 expressing and mock control cell lines. Our results indicated a clear increase
351 in phosphorylated polII binding at the *miR34a* promoter upon *lncTAM34a*
352 over-expression indicating the ability of *lncTAM34a* to transcriptionally
353 regulate *miR34a* levels (**Fig. 3d**).

354

355

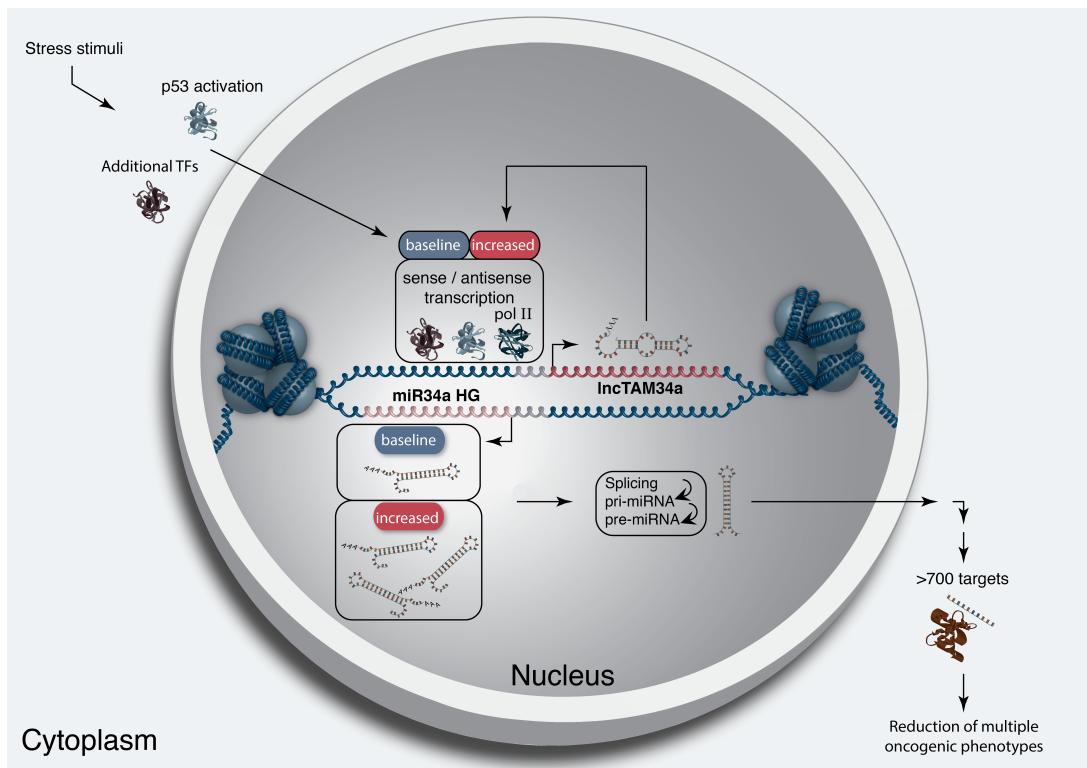


356 **Figure 4: Survival analysis in TCGA cancers.** A) Kaplan-Meier survival curves comparing the
 357 effects of *TP53*-mutated samples (left), low *lncTAM34a* expression (middle) and low *miR34a*
 358 expression (right) to control samples in papillary kidney cancer (results for other cancers in Figure 4-
 359 Supplement 1). Middle and Right panel include only *TP53* wild type patients where RNAseq data
 360 exists. B) Correlation analysis between the effects on the 5-year survival probability of *TP53*-mutated
 361 samples, low *lncTAM34a* expression, and low *miR34a* expression as indicated. For each variable the 5-
 362 year survival probability was compared to the control group (negative values indicate lower survival,
 363 positive values indicate higher survival). Spearman correlation coefficients are given on the top left of
 364 each plot. Each dot indicates one cancer type (see Fig. 1c for legend). Boxplots on the bottom
 365 summarize the effects for the parameter on the x-axis, with indication of P values, as calculated using
 366 paired Wilcoxon signed rank test. Low expression was defined as *TP53* non-mutated samples having
 367 expression values in the bottom 10th percentile.

368
369 **Low *lncTAM34a* expression levels are associated with decreased**
370 **survival**

371
372 As *TP53* mutations and low expression of *miR34a* have been associated with
373 worse prognosis in cancer, we compared survival rates of samples with low
374 expression of *lncTAM34a* (bottom 10th percentile) to control samples in 17
375 cancer types from TCGA (**Figure 4-Supplement 1**) (Gallardo et al. 2009,
376 Zenz et al. 2009, Liu et al. 2011). To correct for the effect of *TP53* mutations
377 we focused on non-*TP53* mutated samples, and noted a worse survival for the
378 low expression group in several cancers. This effect was most pronounced in
379 papillary kidney cancer (unadjusted $P=0.00095$; **Fig. 4a**). By systematically
380 comparing 5-year survival probabilities between the low expression group and
381 the control group for each cancer we found a median reduction of 5-year
382 survival probability of 9.6% ($P=0.083$; Wilcoxon signed rank test; **Fig. 4b**).
383 Furthermore, we found that *lncTAM34a* expression showed similar patterns in
384 terms of direction and strength of association with 5-year survival probability
385 as *miR34a* expression ($r=0.57$, $P=0.037$) and *TP53* mutations ($r=0.80$,
386 $P=0.00054$) across the different cancer types (**Fig. 4b**). Although these results
387 do not implicate any causal relationship, they do indicate a striking similarity
388 between the association of worse prognosis and *TP53* mutations, low
389 *miR34a*, and low *lncTAM34a* expression.

390



391

Figure 5: A graphical summary of the proposed *IncTAM34a* function. Stress stimuli, originating in the cytoplasm or nucleus, activate TP53 as well as additional factors. These factors then bind to the *miR34a* promoter and drive baseline transcription levels of the sense and antisense strands. *IncTAM34a* serves to further increase *miR34a* HG transcription levels resulting in enrichment of polymerase II at the *miR34a* promoter and a positive feed-forward loop. *miR34a* HG then, in turn, is spliced and processed in multiple steps before the mature *miR34a* binds to the RISC complex allowing it to repress its targets and exert its tumor suppressive effects.

392

393

394

395

396

397

398 **Discussion**

399
400 Multiple studies have previously shown asRNAs to be crucial for the
401 appropriate regulation of cancer-associated protein-coding genes and that
402 their dysregulation can lead to perturbation of tumor suppressive and
403 oncogenic pathways, as well as, cancer-related phenotypes (Yu et al. 2008,
404 Yap et al. 2010, Serviss et al. 2014, Balbin et al. 2015). Here we show that
405 asRNAs are also capable of regulating cancer-associated miRNAs resulting in
406 similar consequences as protein-coding gene dysregulation (**Fig. 4**).
407 Interestingly, we show that, both in the presence and absence of
408 *TP53*, *lncTAM34a* provides an additional regulatory level to control *miR34a*
409 expression in both homeostasis and upon encountering various forms of
410 cellular stress. Furthermore, we find that *lncTAM34a*-mediated increase in
411 *miR34a* expression is sufficient to drive the appropriate cellular responses to
412 these stress stimuli (**Fig. 2d and Fig. 3c**). Previous studies have exploited
413 various molecular biology methods to up-regulate *miR34a* expression in cells
414 lacking wild type *TP53* (Liu et al. 2011, Ahn et al. 2012, Yang et al. 2012,
415 Stahlhut et al. 2015, Wang et al. 2015). In this study, we demonstrate a novel,
416 endogenous mechanism of *miR34a* regulation that has similar phenotypic
417 outcomes as has been previously shown for *miR34a* induction in a *TP53*
418 deficient background.

419

420 In agreement with previous studies, we demonstrate that upon encountering
421 various types of cellular stress, TP53 in concert with additional factors initiates
422 transcription at the *miR34a* locus, thus increasing the levels of *lncTAM34a*
423 and *miR34a* (Rashi-Elkeles et al. 2014, Hunten et al. 2015, Leveille et al.

424 2015, Ashouri et al. 2016, Kim et al. 2017). We found that over-expression of
425 *IncTAM34a* leads to recruitment of polII to the *miR34a* promoter and
426 hypothesize that *IncTAM34a* may provide positive feedback for *miR34a*
427 expression whereby it serves as a scaffold for the recruitment of additional
428 factors that facilitate polII-mediated transcription. In this manner, *miR34a*
429 expression is induced, driving a shift towards a reduction in growth factor
430 signaling, senescence, and in some cases apoptosis. On the other hand, in
431 cells without functional TP53, other factors, which typically act independently
432 or in concert with TP53, may initiate transcription of the *miR34a* locus. Due to
433 the fact that *IncTAM34a* can alter *miR34a* expression in these cells, we
434 suggest that it is interacting with one of these additional factors, possibly
435 recruiting it to the *miR34a* locus in order to drive *miR34a* transcription, similar
436 to mechanisms described for other lncRNAs (Hung et al. 2011, Ng et al. 2012,
437 Ng et al. 2013). The head-to-head orientation of the *miR34a* HG and
438 *IncTAM34a* causes sequence complementarity between the RNA and the
439 promoter DNA, making targeting by direct binding an attractive mechanism.
440 Previous reports have also illustrated the ability of asRNAs to form hybrid
441 DNA:RNA R-loops and, thus, facilitate an open chromatin structure and the
442 transcription of the sense gene (Boque-Sastre et al. 2015). The fact that the
443 p1 construct only contains a small portion (~300 bp) of the *IncTAM34a*
444 transcript indicates that this portion is sufficient to give rise to at least a partial
445 *miR34a* inducing response and therefore, that *IncTAM34a* may be able to
446 facilitate *miR34a* expression independent of additional factors (**Fig 2d, Figure**
447 **2-Figure Supplement 2a**). Nevertheless, further work will need to be
448 performed to explore the mechanism whereby *IncTAM34a* regulates *miR34a*

449 gene expression.

450

451 An antisense transcript arising from the *miR34a* locus, *Lnc34a*, has been
452 previously reported to negatively regulate the expression of *miR34a* (Wang et
453 al. 2016). Although the *Lnc34a* and *IncTAM34a* transcripts share some
454 sequence similarity, we believe them to be separate RNAs that are,
455 potentially, different isoforms of the same gene. We utilized CAGE and
456 RNAseq data from the ENCODE project to evaluate the presence of
457 *IncTAM34a* and *Lnc34a* in 28 and 36 commonly used cancer cell lines,
458 respectively. Although the results show the presence of *IncTAM34a* in these
459 cell lines, we find no evidence for *Lnc34a* transcription (**Supplementary**
460 **Document 1**). These results are in line with the findings of Wang et al.
461 indicating that *Lnc34a* is highly expressed in colon cancer stem cell spheres
462 compared to all other cell types used in their study and may not be broadly
463 expressed in other tissues or tumor types. The fact that *IncTAM34a* and
464 *Lnc34a* would appear to have opposing roles in their regulation of *miR34a*,
465 further underlines the complexity of the regulation at this locus.

466

467 Clinical trials utilizing *miR34a* replacement therapy have previously been
468 conducted but, disappointingly, were terminated after adverse side effects of
469 an immunological nature were observed in several of the patients (Slabakova
470 et al. 2017). Although it is not presently clear if these side effects were caused
471 by *miR34a* or the liposomal carrier used to deliver the miRNA, the multitude of
472 evidence indicating *miR34a*'s crucial role in oncogenesis still makes its
473 therapeutic induction an interesting strategy and needs further investigation.

474 Our results indicate an association between survival probability and low
475 *lncTAM34a* expression making it an attractive candidate for controlled
476 preclinical studies. Due to the *lncTAM34a*-mediated positive feedback on
477 *miR34a* expression, initiation of this feedback mechanism may provide a
478 sustained *miR34a* induction in a relatively more robust manner than *miR34a*
479 replacement alone. In summary, our results have identified *lncTAM34a* as a
480 vital component in the regulation of *miR34a* and its particular importance in
481 typical examples of cellular stress encountered in cancer. On a broader level,
482 the conclusions drawn in this study provide an example of asRNA-mediated
483 regulation of a clinically relevant cancer-associated miRNA and contribute to
484 fundamental knowledge concerning *miR34a* regulation.

485

486 **Materials and Methods**

487 **Cell Culture**

488 All cell lines were cultured at 5% CO₂ and 37°C with HEK293T, Saos2, and
489 Skov3 cells cultured in DMEM high glucose (GE Healthcare Life Sciences,
490 Hyclone, Amersham, UK, Cat# SH30081), HCT116 and U2OS cells in
491 McCoy's 5a (ThermoFisher Scientific, Pittsburgh, MA, USA, Cat# SH30200),
492 and PC3 cells in RPMI (GE Healthcare Life Sciences, Hyclone, Cat#
493 SH3009602) and 2 mM L-glutamine (GE Healthcare Life Sciences, Hyclone,
494 Cat# SH3003402). All growth mediums were supplemented with 10% heat-
495 inactivated FBS (ThermoFisher Scientific, Gibco, Cat# 12657029) and 50
496 µg/ml of streptomycin (ThermoFisher Scientific, Gibco, Cat# 15140122) and
497 50 µg/ml of penicillin (ThermoFisher Scientific, Gibco, Cat# 15140122). All cell
498 lines were purchased from ATCC, tested negative for mycoplasma, and their

499 identity was verified via STR profiling.

500

501 **Bioinformatics, Data Availability, and Statistical Testing**

502 The USCS genome browser (Kent et al. 2002) was utilized for the
503 bioinformatic evaluation of antisense transcription utilizing the RefSeq
504 (O'Leary et al. 2016) gene annotation track.

505

506 All raw experimental data, code used for analysis, and supplementary
507 methods are available for review at (Serviss 2017) and are provided as an R
508 package. All analysis took place using the R statistical programming language
509 (Team 2017) using external packages that are documented in the package
510 associated with this article (Wilkins , Chang 2014, Wickham 2014, Therneau
511 2015, Wickham 2016, Allaire et al. 2017, Arnold 2017, Wickham 2017,
512 Wickham 2017, Wickham 2017, Xiao 2017, Xie 2017). The package facilitates
513 replication of the operating system and package versions used for the original
514 analysis, reproduction of each individual figure and figure supplement
515 included in the article, and easy review of the code used for all steps of the
516 analysis, from raw-data to figure.

517

518 The significance threshold (alpha) in this study was set to 0.05. Statistical
519 testing was performed using an unpaired two sample Student's t-test unless
520 otherwise specified.

521

522 **Coding Potential**

523 Protein-coding capacity was evaluated using the Coding-potential
524 Assessment Tool (Wang et al. 2013) and Coding-potential Calculator (Kong et

525 al. 2007) with default settings. Transcript sequences for use with Coding-
526 potential Assessment Tool were downloaded from the UCSC genome
527 browser using the Ensembl
528 accessions: *HOTAIR* (ENST00000455246), *XIST* (ENST00000429829), β-
529 actin (ENST00000331789), Tubulin (ENST00000427480),
530 and *MYC* (ENST00000377970). Transcript sequences for use with Coding-
531 potential Calculator were downloaded from the UCSC genome browser using
532 the following IDs: *HOTAIR* (uc031qho.1), β-actin (uc003sq.4).

533

534 **Peptide identification in MS/MS spectra**

535 Orbitrap raw MS/MS files for 11 human cell lines were downloaded from the
536 PRIDE repository (PXD002395; (Geiger et al. 2012)) converted to mzML
537 format using msConvert from the ProteoWizard tool suite (Holman et al.
538 2014). Spectra were then searched using MSGF+ (v10072) (Kim et al. 2014)
539 and Percolator (v2.08) (Granholm et al. 2014). All searches were done
540 against the human protein subset of Ensembl 75 in the Galaxy platform
541 (Boekel et al. 2015) supplemented with the 6 frame translation of both the
542 annotated (LOC102724571; hg38) and PCR cloned sequence of *IncTAM34a*
543 (supplementary data; (Serviss 2017)). MSGF+ settings included precursor
544 mass tolerance of 10 ppm, fully-tryptic peptides, maximum peptide length of
545 50 amino acids and a maximum charge of 6. Fixed modification was
546 carbamidomethylation on cysteine residues; a variable modification was used
547 for oxidation on methionine residues. Peptide Spectral Matches found at 1%
548 FDR (false discovery rate) were used to infer peptide identities. The output
549 from all searches are available in (Serviss 2017).

550

551 **shRNAs**

552 shRNA-expressing constructs were cloned into the U6M2 construct using the
553 BgIII and KpnI restriction sites as previously described (Amarzguioui et al.
554 2005). shRNA constructs were transfected using Lipofectamine 2000 or 3000
555 (ThermoFisher Scientific, Cat# 12566014 and L3000015). The sequences
556 targeting renilla is as follows: shRenilla 1.1 (AAT ACA CCG CGC TAC TGG
557 C), shRenilla 2.1 (TAA CGG GAT TTC ACG AGG C).

558

559 **Bi-directional Promoter Cloning**

560 The overlapping region (p1) corresponds with the sequence previously
561 published as the TP53 binding site in (Raver-Shapira et al. 2007) which we
562 synthesized, cloned into the pLucRluc construct (Polson et al. 2011), and
563 sequenced to verify its identity.

564

565 **Promoter Activity**

566 Cells were co-transfected with the p1 renilla/firefly bidirectional promoter
567 construct (Polson et al. 2011) and GFP by using Lipofectamine 2000 (Life
568 Technologies, Cat# 12566014). The expression of GFP and luminescence
569 was measured 24 h post transfection by using the Dual-Glo Luciferase Assay
570 System (Promega, Cat# E2920) and detected by the GloMax-Multi+ Detection
571 System (Promega, Cat# SA3030). The expression of luminescence was
572 normalized to GFP.

573

574 **Generation of U6-expressed *IncTAM34a* Lentiviral Constructs**

575 The U6 promoter was amplified from the U6M2 cloning plasmid (Amarzguioui
576 et al. 2005) and ligated into the Not1 restriction site of the pHIV7-IMPDH2
577 vector (Turner et al. 2012). *IncTAM34a* was PCR amplified and subsequently
578 cloned into the Nhe1 and Pac1 restriction sites in the pHIV7-IMPDH2-U6
579 plasmid.

580

581 **Lentiviral Particle production, infection, and selection**

582 Lentivirus production was performed as previously described in (Turner et al.
583 2012). Briefly, HEK293T cells were transfected with viral and expression
584 constructs using Lipofectamine 2000 (ThermoFisher Scientific, Cat#
585 12566014), after which viral supernatants were harvested 48 and 72 hours
586 post-transfection. Viral particles were concentrated using PEG-IT solution
587 (Systems Biosciences, Palo Alto, CA, USA. Cat# LV825A-1) according to the
588 manufacturer's recommendations. HEK293T cells were used for virus titration
589 and GFP expression was evaluated 72hrs post-infection via flow cytometry
590 (LSRII, BD Biosciences, San Jose, CA, USA) after which TU/ml was
591 calculated.

592

593 Stable lines were generated by infecting cells with a multiplicity of infection of
594 1 and subsequently initiating 1-2 µM mycophenolic acid-based (Merck,
595 Kenilworth, NJ, USA. Cat# M5255) selection 48-72 hours post-infection. Cells
596 were expanded as the selection process was monitored via flow cytometry
597 analysis (LSRII, BD Biosciences) of GFP and selection was terminated once
598 > 90% of the cells were GFP positive. Quantification of *IncTAM34a* over-
599 expression and *miR34a* was performed in biological quintuplet for all cell

600 lines.

601

602 **Western Blotting**

603 Samples were lysed in 50 mM Tris-HCl (Sigma Aldrich, St. Louis, MO, USA).

604 Cat# T2663), pH 7.4, 1% NP-40 (Sigma Aldrich, Cat# I8896), 150 mM NaCl

605 (Sigma Aldrich, Cat# S5886), 1 mM EDTA (Promega, Madison, WI, USA).

606 Cat# V4231), 1% glycerol (Sigma Aldrich, Cat# G5516), 100 µM vanadate

607 (Sigma Aldrich, Cat# S6508), protease inhibitor cocktail (Roche Diagnostics,

608 Basel, Switzerland, Cat# 004693159001) and PhosSTOP (Roche

609 Diagnostics, Cat# 04906837001). Lysates were subjected to SDS-PAGE and

610 transferred to PVDF membranes. The proteins were detected by western blot

611 analysis by using an enhanced chemiluminescence system (Western

612 Lightning-ECL, PerkinElmer, Waltham, MA, USA. Cat# NEL103001EA).

613 Antibodies used were specific for CCND1 1:1000 (Cell Signaling, Danvers,

614 MA, USA. Cat# 2926), and GAPDH 1:5000 (Abcam, Cambridge, UK, Cat#

615 ab9485). All western blot quantifications were performed using ImageJ

616 (Schneider et al. 2012).

617

618 **RNA Extraction and cDNA Synthesis**

619 For downstream SYBR green applications, RNA was extracted using the

620 RNeasy mini kit (Qiagen, Venlo, Netherlands, Cat# 74106) and subsequently

621 treated with DNase (Ambion Turbo DNA-free, ThermoFisher Scientific, Cat#

622 AM1907). 500ng RNA was used for cDNA synthesis using MuMLV

623 (ThermoFisher Scientific, Cat# 28025013) and a 1:1 mix of oligo(dT) and

624 random nanomers.

625
626 For analysis of miRNA expression with Taqman, samples were isolated with
627 TRIzol reagent (ThermoFisher Scientific, Cat# 15596018) and further
628 processed with the miRNeasy kit (Qiagen, Cat# 74106). cDNA synthesis was
629 performed using the TaqMan MicroRNA Reverse Transcription Kit
630 (ThermoFisher Scientific, Cat# 4366597) using the corresponding oligos
631 according to the manufacturer's recommendations.

632

633 **QPCR and PCR**

634 PCR was performed using the KAPA2G Fast HotStart ReadyMix PCR Kit
635 (Kapa Biosystems, Wilmington, MA, USA, Cat# KK5601) with corresponding
636 primers. QPCR was carried out using KAPA 2G SYBRGreen (Kapa
637 Biosystems, Cat# KK4602) using the Applied Biosystems 7900HT machine
638 with the cycling conditions: 95 °C for 3 min, 95 °C for 3 s, 60 °C for 30 s.

639

640 QPCR for miRNA expression analysis was performed according to the primer
641 probe set manufacturers recommendations (ThermoFisher Scientific) and
642 using the TaqMan Universal PCR Master Mix (ThermoFisher Scientific, Cat#
643 4304437) with the same cycling scheme as above. Primer and probe sets for
644 TaqMan were also purchased from ThermoFisher Scientific (Life
645 Technologies at time of purchase, TaqMan® MicroRNA Assay, hsa-miR-34a,
646 human, Cat# 4440887, Assay ID: 000426 and Control miRNA Assay, RNU48,
647 human, Cat# 4440887, Assay ID: 001006).

648

649 The ΔΔCt method was used to quantify gene expression. All QPCR-based

650 experiments were performed in at least technical duplicate. Primers for all
651 PCR-based experiments are listed in **Supplementary Document 2** and
652 arranged by figure.

653

654 **Cell Cycle Distribution**

655 Cells were washed in PBS and fixed in 4% paraformaldehyde at room
656 temperature overnight. Paraformaldehyde was removed, and cells were re-
657 suspended in 95% EtOH. The samples were then rehydrated in distilled
658 water, stained with DAPI and analyzed by flow cytometry on a LSRII (BD
659 Biosciences) machine. Resulting cell cycle phases were quantified using the
660 ModFit software (Verity Software House, Topsham, ME, USA). Experiments
661 were performed in biological quadruplet (PC3) or triplicate (Skov3). The log2
662 fraction of cell cycle phase was calculated for each replicate and a two
663 sample t-test was utilized for statistical testing.

664

665 **3' Rapid Amplification of cDNA Ends**

666 3'-RACE was performed as described as previously in (Johnsson et al. 2013).
667 Briefly, U2OS cell RNA was polyA-tailed using yeast polyA polymerase
668 (ThermoFisher Scientific, Cat# 74225Z25KU) after which cDNA was
669 synthesized using oligo(dT) primers. Nested-PCR was performed first using a
670 forward primer in *lncTAM34a* exon 1 and a tailed oligo(dT) primer followed by
671 a second PCR using an alternate *lncTAM34a* exon 1 primer and a reverse
672 primer binding to the tail of the previously used oligo(dT) primer. PCR
673 products were gel purified and cloned the Strata Clone Kit (Agilent
674 Technologies, Santa Clara, CA, USA. Cat# 240205), and sequenced.

675

676 **Chromatin Immunoprecipitation**

677 The ChIP was performed as previously described in (Johnsson et al. 2013)
678 with the following modifications. Cells were crosslinked in 1% formaldehyde
679 (Merck, Cat# 1040039025), quenched with 0.125M glycine (Sigma Aldrich,
680 Cat# G7126), and lysed in cell lysis buffer comprised of: 5mM PIPES (Sigma
681 Aldrich, Cat# 80635), 85mM KCL (Merck, Cat# 4936), 0.5% NP40 (Sigma
682 Aldrich, Cat# I8896), protease inhibitor (Roche Diagnostics, Cat#
683 004693159001). Samples were then sonicated in 50mM TRIS-HCL pH 8.0
684 (Sigma Aldrich, MO, USA, Cat# T2663) 10mM EDTA (Promega, WI, USA,
685 Cat# V4231), 1% SDS (ThermoFisher Scientific, Cat# AM9822), and protease
686 inhibitor (Roche Diagnostics, Cat# 004693159001) using a Bioruptor
687 Sonicator (Diagenode, Denville, NJ, USA). Samples were incubated over
688 night at 4°C with the polII antibody (Abcam, Cat# ab5095) and subsequently
689 pulled down with Salmon Sperm DNA/Protein A Agarose (Millipore, Cat# 16-
690 157) beads. DNA was eluted in an elution buffer of 1% SDS (ThermoFisher
691 Scientific, Cat# AM9822) 100mM NaHCO3 (Sigma Aldrich, Cat# 71631),
692 followed by reverse crosslinking, RNaseA (ThermoFisher Scientific, Cat#
693 1692412) and protease K (New England Biolabs, Ipswich, MA, USA, Cat#
694 P8107S) treatment. The DNA was eluted using Qiagen PCR purification kit
695 (Cat# 28106) and quantified via QPCR. QPCR was performed in technical
696 duplicate using the standard curve method and reported absolute values. The
697 fraction of input was subsequently calculated using the mean of the technical
698 replicates followed by calculating the fold over the control condition. Statistical
699 testing was performed using 4 biological replicates with the null hypothesis

700 that the true log₂ fold change values were equal to zero.

701

702 **Confluency Analysis**

703 Cells were incubated in the Spark Multimode Microplate (Tecan, Männedorf,
704 Switzerland) reader for 48 hours at 37°C with 5% CO₂ in a humidity chamber
705 in either normal medium or HBSS (ThermoFisher Scientific, Cat# 14025092) .

706 Confluency was measured every hour using bright-field microscopy and the
707 percentage of confluency was reported via the plate reader's inbuilt algorithm.

708 Percentage of confluency was normalized to the control sample in each
709 condition (shown in figure) and then ranked to move the data to a linear scale.

710 Using the mean of the technical duplicates in three biological replicates, the
711 rank was then used to construct a linear model, of the dependency of the rank
712 on the time and cell lines variables for each growth condition. Reported *P*
713 values are derived from the t-test, testing the null hypothesis that the
714 coefficient estimate of the cell line variable is equal to 0.

715

716 **Pharmacological Compounds**

717 Doxorubicin was purchased from Teva (Petah Tikva, Israel, cat. nr. 021361).

718

719 **Cellular Localization Analysis**

720 Quantified RNAseq data from 11 cell lines from the GRCh38 assembly was
721 downloaded from the ENCODE project database and quantifications for
722 *IncTAM34a* (ENSG00000234546), GAPDH (ENSG00000111640), and
723 MALAT1 (ENSG00000251562) were extracted. Cell lines for which data was
724 downloaded include: A549, GM12878, HeLa-S3, HepG2, HT1080, K562

725 MCF-7, NCI-H460, SK-MEL-5, SK-N-DZ, SK-N-SH. Initial exploratory analysis
726 revealed that several cell lines should be removed from the analysis due to a)
727 a larger proportion of GAPDH in the nucleus than cytoplasm or b) variation of
728 *lncTAM34a* expression is too large to draw conclusions, or c) they have no or
729 low (<6 TPM) *lncTAM34a* expression. Furthermore, only polyadenylated
730 libraries were used in the final analysis, due to the fact that the cellular
731 compartment enrichment was improved in these samples. All analyzed genes
732 are reported to be polyadenylated. In addition, only samples with 2 biological
733 replicates were retained. For each cell type, gene, and biological replicate the
734 fraction of transcripts per million (TPM) in each cellular compartment was
735 calculated as the fraction of TPM in the specific compartment by the total
736 TPM. The mean and standard deviation for the fraction was subsequently
737 calculated for each cell type and cellular compartment and this information
738 was represented in the final figure.

739

740 **CAGE Analysis**

741 All available CAGE data from the ENCODE project (Consortium 2012) for 36
742 cell lines was downloaded from the UCSC genome browser (Kent et al. 2002)
743 for genome version hg19. Of these, 28 cell lines had CAGE transcription start
744 sites (TSS) mapping to the plus strand of chromosome 1 and in regions
745 corresponding to 200 base pairs upstream of the *Lnc34a* start site (9241796 -
746 200) and 200 base pairs upstream of the GENCODE annotated *lncTAM34a*
747 start site (9242263 + 200). These cell lines included: HFDPC, H1-hESC,
748 HMEpC, HAoEC, HPIEpC, HSaVEC, GM12878, hMSC-BM, HUVEC,
749 AG04450, hMSC-UC, IMR90, NHDF, SK-N-SH_RA, BJ, HOB, HPC-PL,

750 HAoAF, NHEK, HVMF, HWP, MCF-7, HepG2, hMSC-AT, NHEM.f_M2,
751 SkMC, NHEM_M2, and HCH. In total 74 samples were included. 17 samples
752 were polyA-, 47 samples were polyA+, and 10 samples were total RNA. In
753 addition, 34 samples were whole cell, 15 enriched for the cytosolic fraction, 15
754 enriched for the nucleolus, and 15 enriched for the nucleus. All CAGE
755 transcription start sites were plotted and the RPKM of the individual reads was
756 used to color each read to indicate their relative abundance. In cases where
757 CAGE TSS spanned identical regions, the RPMKs of the regions were
758 summed and represented as one CAGE TSS in the figure. In addition, a
759 density plot shows the distribution of the CAGE reads in the specified
760 interval.

761

762 Splice Junction Analysis

763 All available whole cell (i.e. non-fractionated) spliced read data originating
764 from the Cold Spring Harbor Lab in the ENCODE project (Consortium 2012)
765 for 38 cell lines was downloaded from the UCSC genome browser (Kent et al.
766 2002). Of these cell lines, 36 had spliced reads mapping to the plus strand of
767 chromosome 1 and in the region between the *Lnc34a* start (9241796) and
768 transcription termination (9257102) site (note that *lncTAM34a* resides totally
769 within this region). Splice junctions from the following cell lines were included
770 in the final figure: A549, Ag04450, Bj, CD20, CD34 mobilized, Gm12878,
771 H1hesc, Haoaf, Haoec, Hch, Helas3, Hepg2, Hfdpc, Hmec, Hmepc, Hmscat,
772 Hmscbm, Hmscuc, Hob, Hpcpl, Hpiepc, Hsavec, Hsmmm, Huvec, Hvmf, Hwp,
773 Imr90, Mcf7, Monocd14, Nhdf, Nhek, Nhemfm2, Nhemm2, Nhlf, Skmc, and
774 Sknsh. All splice junctions were included in the figure and colored according

775 to the number of reads corresponding to each. In cases where identical reads
776 were detected multiple times, the read count was summed and represented
777 as one read in the figure.

778

779 **TCGA Data Analysis**

780 RNAseq data and copy number data were downloaded from TCGA and
781 processed as described previously (Ashouri et al. 2016). Briefly, RNAseq data
782 were aligned to the human hg19 assembly and quantified using GENCODE
783 (v19) annotated HTSeq-counts and FPKM normalizations. Expression data
784 from *miR34a* and *IncTAM34a* (identified as RP3-510D11.2) were used for
785 further analysis. Copy number amplitudes for GENCODE genes were
786 determined from segmented copy-number data. Samples that were diploid for
787 *IncTAM34a* were identified as those samples that had copy number
788 amplitudes between -0.1 and 0.1.

789

790 Somatic mutation data were downloaded from the Genomics Data Commons
791 data portal (GDC) as mutation annotation format (maf) files, called using
792 Mutect2 on 30/10/2017 (v7) (Grossman et al. 2016).

793

794 Survival analysis was performed on TCGA vital state and follow-up data,
795 downloaded from GDC on 27/10/2017 using the R survival package
796 (Therneau 2015).

797

798 **Acknowledgments**

799 The authors would like to kindly thank Martin Enge for his critical review of the
800 manuscript and fruitful discussions. Dan Grandér, who played a significant

801 role in the conceptualization and supervision of this project, sadly passed
802 away before the initial submission of the manuscript. May he rest in peace.

803

804 **Competing Interests**

805
806 The authors declare no competing interests.

807

808

809 **Funding**

810
811 This work has been supported by the Swedish Research Council [521-2012-
812 2037], Swedish Cancer Society [150768], Cancer Research Foundations of
813 Radiumhemmet [144063] and the Swedish Childhood Cancer Foundation
814 [PR2015-0009].

815

816

817 **Figure Supplements**

818
819 Figure 1-Supplement 1: TCAG expression levels and correlation analysis
820 statistics.
821
822 Figure 1-Supplement 2: Molecular characteristics of *lncTAM34a*.
823
824 Figure 2-Supplement 1: A schematic representation of the p1 construct.
825
826 Figure 2-Supplement 2: Evaluating the effects of *lncTAM34a* down-regulation.
827
828 Figure 3-Supplement 1: Physiological relevance of *lncTAM34a* over-
829 expression.
830
831 Figure 3-Supplement 2: Effects of *lncTAM34a* over-expression on cyclin D1.
832
833 Figure 4-Supplement 1: Survival analysis in 17 cancers from TCGA.
834
835 Supplementary Document 1: Evaluating the relationship between *lncTAM34a*
836 and *Lnc34a*.*
837
838 Supplementary Document 2: A table of primers used in this study.*
839
840 *Please note that in the initial submission these documents are included after
841 the References section of the pdf file.

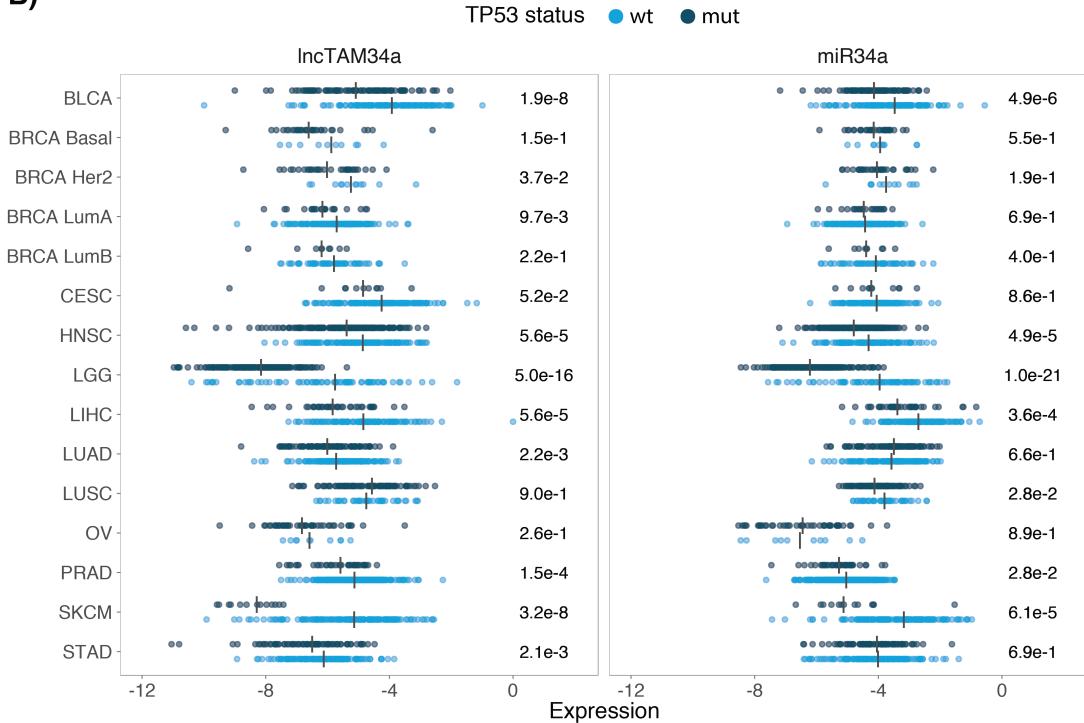
842

843 **Supplementary Figures**

A)

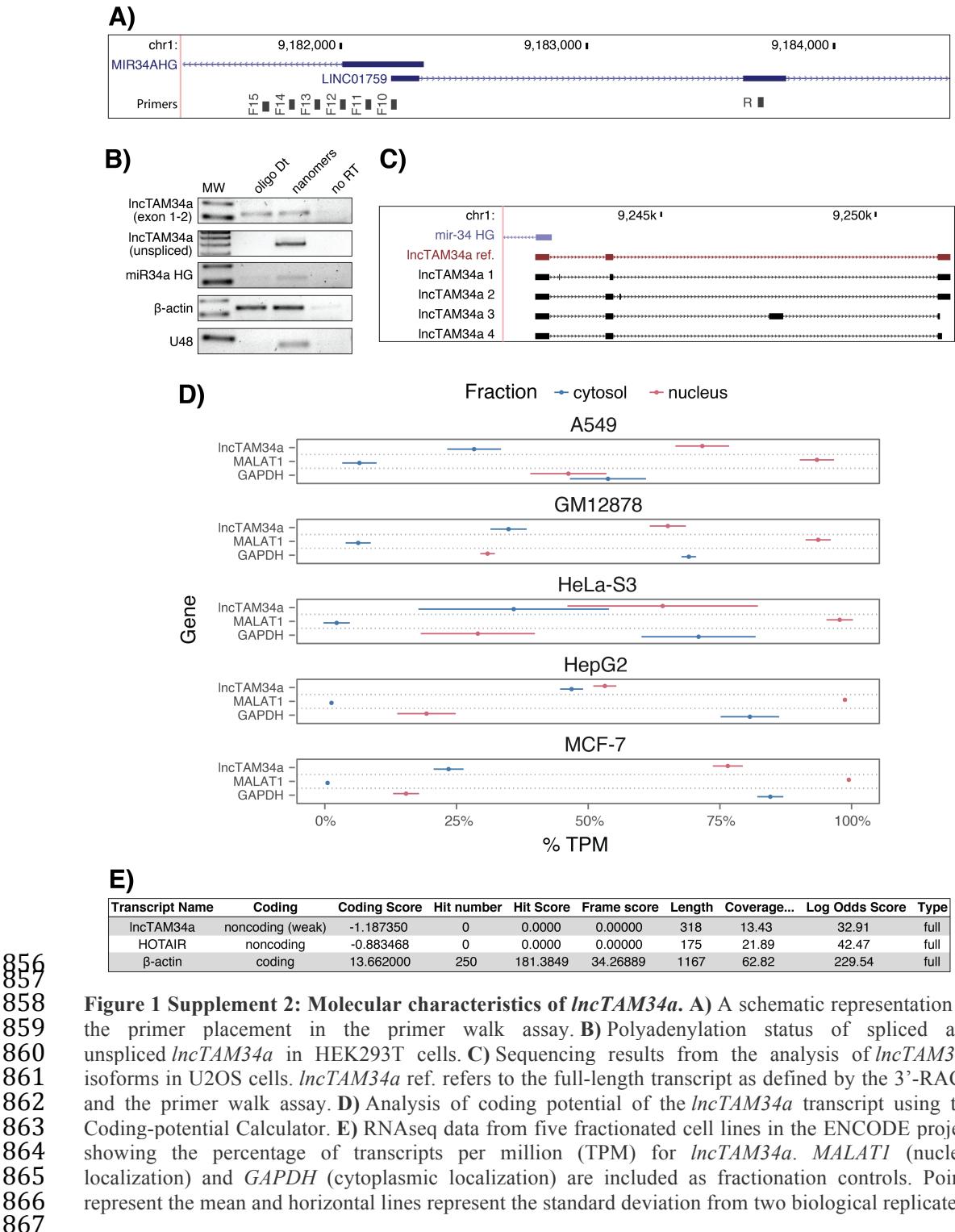
cancer	all n	all rho	all p	TP53wt n	TP53wt rho	TP53wt p	TP53mut n	TP53mut rho	TP53mut p
Adrenocortical carcinoma (ACC)	10	0.55	1.04e-01	10	0.55	1.04e-01	NA	NA	NA
Bladder Urothelial Carcinoma (BLCA)	228	0.51	7.89e-17	134	0.45	3.86e-08	94	0.43	1.73e-05
Breast invasive carcinoma (BRCA) Basal	42	0.57	9.54e-05	10	0.62	6.02e-02	32	0.57	7.41e-04
Breast invasive carcinoma (BRCA) Her2	44	0.15	3.39e-01	12	0.22	4.85e-01	32	0.07	7.10e-01
Breast invasive carcinoma (BRCA) LumA	199	0.34	8.22e-07	177	0.34	2.96e-06	22	0.49	2.31e-02
Breast invasive carcinoma (BRCA) LumB	70	0.17	1.57e-01	61	0.15	2.53e-01	9	0.17	6.78e-01
Cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC)	156	0.14	8.37e-02	145	0.16	5.45e-02	11	-0.05	9.03e-01
Head and Neck squamous cell carcinoma (HNSC)	313	0.54	8.38e-25	123	0.61	0.00e+00	190	0.45	9.68e-11
Kidney Chromophobe (KICH)	5	0.60	3.50e-01	5	0.60	3.50e-01	NA	NA	NA
Kidney renal clear cell carcinoma (KIRC)	142	0.35	2.06e-05	141	0.34	4.41e-05	NA	NA	NA
Kidney renal papillary cell carcinoma (KIRP)	167	0.45	9.16e-10	163	0.45	2.04e-09	4	0.80	3.33e-01
Brain Lower Grade Glioma (LGG)	271	0.63	9.92e-32	76	0.73	0.00e+00	195	0.39	2.26e-08
Liver hepatocellular carcinoma (LIHC)	153	0.56	3.64e-14	114	0.52	4.18e-09	39	0.45	3.95e-03
Lung adenocarcinoma (LUAD)	234	0.28	1.15e-05	128	0.36	2.87e-05	106	0.23	1.91e-02
Lung squamous cell carcinoma (LUSC)	139	0.23	6.74e-03	42	0.04	7.93e-01	97	0.33	9.91e-04
Ovarian serous cystadenocarcinoma (OV)	56	0.23	8.37e-02	10	0.84	4.46e-03	46	0.15	3.31e-01
Prostate adenocarcinoma (PRAD)	413	0.47	1.33e-23	375	0.46	6.13e-21	38	0.45	4.58e-03
Skin Cutaneous Melanoma (SKCM)	165	0.65	5.43e-21	152	0.61	7.85e-17	13	0.43	1.40e-01
Stomach adenocarcinoma (STAD)	225	0.37	8.23e-09	145	0.37	5.71e-06	80	0.42	1.03e-04
Thyroid carcinoma (THCA)	469	0.46	1.07e-25	467	0.46	4.06e-26	NA	NA	NA

B)

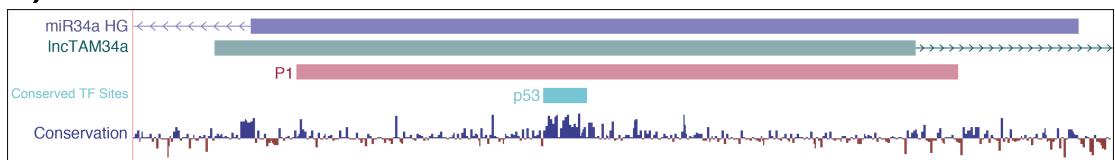


844
845

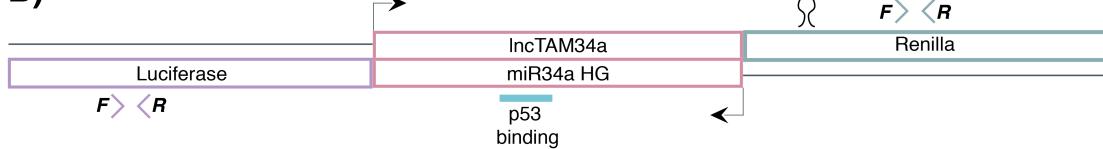
846 **Figure 1 Supplement 1: TCGA normalized expression levels and correlation analysis statistics.**
847 A) Spearman's rho and P values (p) from the correlation analysis in Figure 1a between *miR34a* and
848 *lncTAM34a* expression in *TP53* wild type (wt) and mutated (mut) samples within TCGA cancer types.
849 NA indicates not applicable, due to a lack of data for the specific group. B) Expression levels of
850 *miR34a* and *lncTAM34a* in *TP53* wt and nonsynonymous mutation samples. Expression was quantified
851 by the log2 ratio of expression of the gene to its maximal expression value. Vertical lines indicate the
852 median. P values are indicated on the right side of each panel and are derived from comparing the
853 *TP53* wild type samples to the samples with a nonsynonymous mutation using a two-sided Wilcoxon
854 signed rank test. Only cancers that had at least 5 samples per group were included. In addition, only
855 samples that were diploid at the *miR34a* locus were used for the analysis to avoid copy number bias.



A)

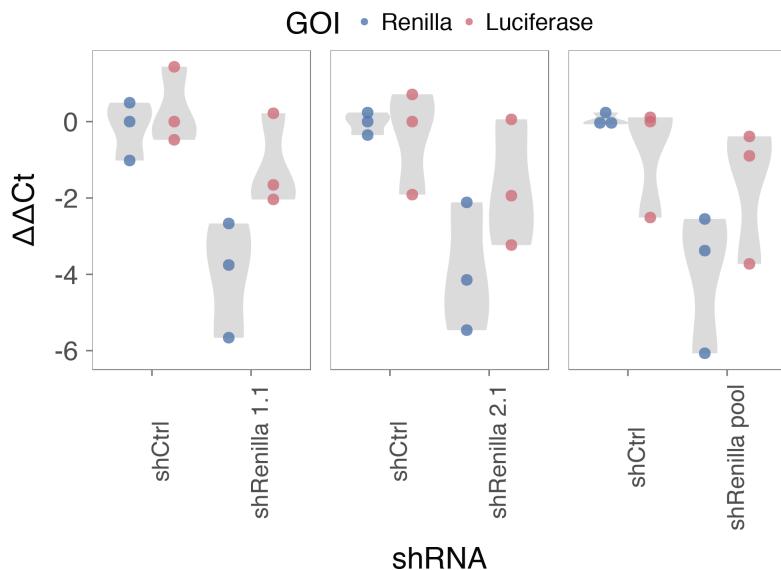


B)



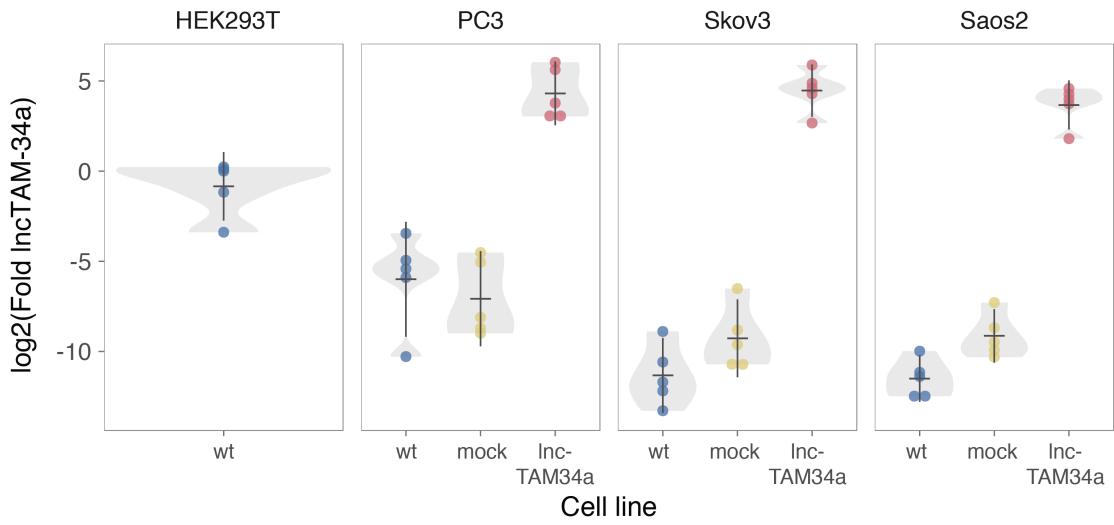
868
869

870 **Figure 2 Supplement 1: A schematic representation of the p1 construct. A)** A UCSC genome
871 browser illustration indicating the location of the promoter region cloned into the p1 construct
872 including the conserved TP53-binding site. **B)** A representative picture of the p1 construct including
873 forward (F) and reverse (R) primer locations and the renilla shRNA targeting site.



874
875
876
877
878
879
880

Figure 2 Supplement 2: Evaluating the effects of *lncTAM34a* down-regulation. HEK293T cells were co-transfected with the p1 construct and either shRenilla or shControl. Renilla and luciferase levels were measured with QPCR 48 hours after transfection. Individual points represent independent experiments with the gray shadow indicating the density of the points. The experiment was performed in biological triplicate.



881
882
883
884
885

Figure 3 Supplement 1: Physiological relevance of *lncTAM34a* over-expression. Comparison of *lncTAM34a* expression in HEK293T cells (high endogenous *lncTAM34a*), and the wild-type (wt), mock, and *lncTAM34a* over-expressing stable cell lines.

886
887
888
889
890

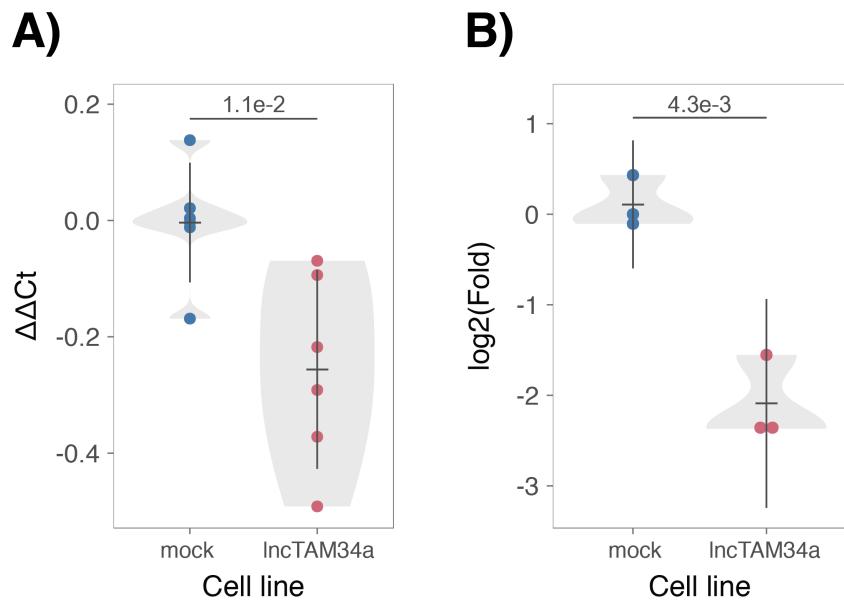
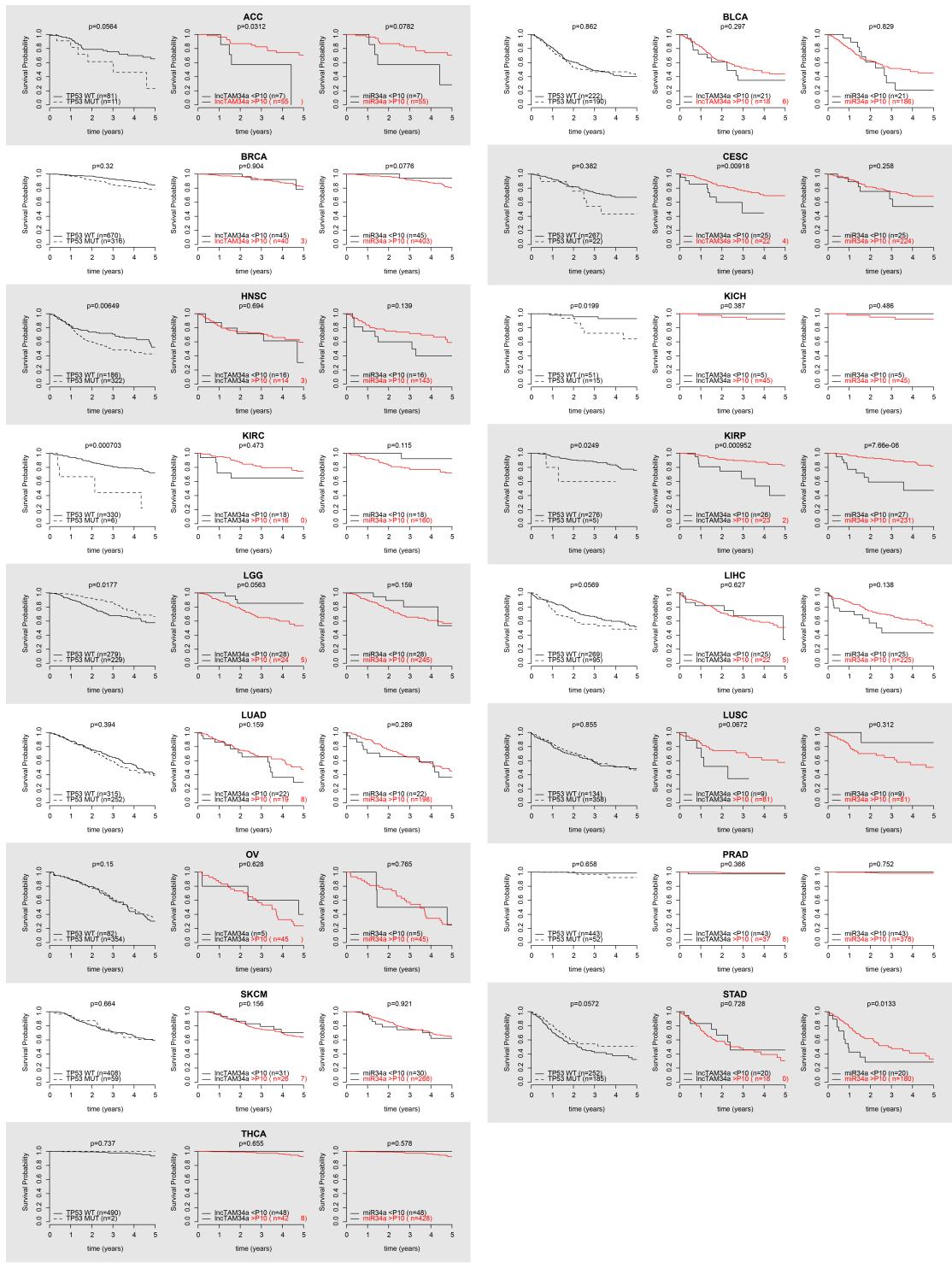


Figure 3 Supplement 2: Effects of *lncTAM34a* over-expression on cyclin D1. CCND1 expression (A) and western blot quantification of protein levels (B) in *lncTAM34a* over-expressing PC3 stable cell lines. Experiments were performed in biological sextuplets (A) or triplicates (B).



891

Figure 4-Supplement 1: Survival analysis in 17 cancers from TCGA. Kaplan-Meier survival curves comparing the survival of *TP53*-mutated samples (left), low *IncTAM34a* expression (middle) and low *miR34a* expression (right) to control samples in 17 cancer types from TCGA. Low expression was defined as *TP53* non-mutated samples having expression values in the bottom 10th percentile.

892
893
894
895

896

897

898 **References**

899

- 900 Agostini, M., P. Tucci, R. Killick, E. Candi, B. S. Sayan, P. Rivetti di Val Cervo, P.
 901 Nicotera, F. McKeon, R. A. Knight, T. W. Mak and G. Melino (2011). "Neuronal
 902 differentiation by TAp73 is mediated by microRNA-34a regulation of synaptic
 903 protein targets." *Proc Natl Acad Sci U S A* **108**(52): 21093-21098. DOI:
 904 10.1073/pnas.1112061109
- 905
- 906 Ahn, Y. H., D. L. Gibbons, D. Chakravarti, C. J. Creighton, Z. H. Rizvi, H. P. Adams, A.
 907 Pertsemlidis, P. A. Gregory, J. A. Wright, G. J. Goodall, E. R. Flores and J. M. Kurie
 908 (2012). "ZEB1 drives prometastatic actin cytoskeletal remodeling by
 909 downregulating miR-34a expression." *J Clin Invest* **122**(9): 3170-3183. DOI:
 910 10.1172/JCI63608
- 911
- 912 Allaire, J., Y. Xie, J. McPherson, J. Luraschi, K. Ushey, A. Atkins, H. Wickham, J.
 913 Cheng and W. Chang (2017). rmarkdown: Dynamic Documents for R. R package
 914 version 1.8. <https://CRAN.R-project.org/package=rmarkdown>
- 915
- 916 Amarzguioui, M., J. J. Rossi and D. Kim (2005). "Approaches for chemically
 917 synthesized siRNA and vector-mediated RNAi." *FEBS Lett* **579**(26): 5974-5981.
 918 DOI: 10.1016/j.febslet.2005.08.070
- 919
- 920 Arnold, J. B. (2017). ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'. R
 921 package version 3.4.0. <https://CRAN.R-project.org/package=ggthemes>
- 922
- 923 Ashouri, A., V. I. Sayin, J. Van den Eynden, S. X. Singh, T. Papagiannakopoulos and
 924 E. Larsson (2016). "Pan-cancer transcriptomic analysis associates long non-
 925 coding RNAs with key mutational driver events." *Nature Communications* **7**:
 926 13197. DOI: 10.1038/ncomms13197
- 927
- 928 Balbin, O. A., R. Malik, S. M. Dhanasekaran, J. R. Prensner, X. Cao, Y. M. Wu, D.
 929 Robinson, R. Wang, G. Chen, D. G. Beer, A. I. Nesvizhskii and A. M. Chinnaiyan
 930 (2015). "The landscape of antisense gene expression in human cancers." *Genome*
 931 *Res* **25**(7): 1068-1079. DOI: 10.1101/gr.180596.114
- 932
- 933 Boekel, J., J. M. Chilton, I. R. Cooke, P. L. Horvatovich, P. D. Jagtap, L. Kall, J. Lehtio,
 934 P. Lukasse, P. D. Moerland and T. J. Griffin (2015). "Multi-omic data analysis using
 935 Galaxy." *Nat Biotechnol* **33**(2): 137-139. DOI: 10.1038/nbt.3134
- 936
- 937 Boque-Sastre, R., M. Soler, C. Oliveira-Mateos, A. Portela, C. Moutinho, S. Sayols, A.
 938 Villanueva, M. Esteller and S. Guil (2015). "Head-to-head antisense transcription
 939 and R-loop formation promotes transcriptional activation." *Proc Natl Acad Sci U*
 940 *S A* **112**(18): 5785-5790. DOI: 10.1073/pnas.1421197112
- 941
- 942 Chang, T. C., D. Yu, Y. S. Lee, E. A. Wentzel, D. E. Arking, K. M. West, C. V. Dang, A.
 943 Thomas-Tikhonenko and J. T. Mendell (2008). "Widespread microRNA
 944 repression by Myc contributes to tumorigenesis." *Nat Genet* **40**(1): 43-50. DOI:
 945 10.1038/ng.2007.30

946
947 Chang, W. (2014). extrafont: Tools for using fonts. R package version 0.17.
948 <https://CRAN.R-project.org/package=extrafont>
949
950 Chen, J., M. Sun, W. J. Kent, X. Huang, H. Xie, W. Wang, G. Zhou, R. Z. Shi and J. D.
951 Rowley (2004). "Over 20% of human transcripts might form sense-antisense
952 pairs." *Nucleic Acids Res* **32**(16): 4812-4820. DOI: 10.1093/nar/gkh818
953
954 Cheng, J., L. Zhou, Q. F. Xie, H. Y. Xie, X. Y. Wei, F. Gao, C. Y. Xing, X. Xu, L. J. Li and S.
955 S. Zheng (2010). "The impact of miR-34a on protein output in hepatocellular
956 carcinoma HepG2 cells." *Proteomics* **10**(8): 1557-1572. DOI:
957 10.1002/pmic.200900646
958
959 Chim, C. S., K. Y. Wong, Y. Qi, F. Loong, W. L. Lam, L. G. Wong, D. Y. Jin, J. F. Costello
960 and R. Liang (2010). "Epigenetic inactivation of the miR-34a in hematological
961 malignancies." *Carcinogenesis* **31**(4): 745-750. DOI: 10.1093/carcin/bgq033
962
963 Cole, K. A., E. F. Attiyeh, Y. P. Mosse, M. J. Laquaglia, S. J. Diskin, G. M. Brodeur and
964 J. M. Maris (2008). "A functional screen identifies miR-34a as a candidate
965 neuroblastoma tumor suppressor gene." *Mol Cancer Res* **6**(5): 735-742. DOI:
966 10.1158/1541-7786.MCR-07-2102
967
968 Conley, A. B. and I. K. Jordan (2012). "Epigenetic regulation of human cis-natural
969 antisense transcripts." *Nucleic Acids Res* **40**(4): 1438-1445. DOI:
970 10.1093/nar/gkr1010
971
972 Encode Project Consortium (2012). "An integrated encyclopedia of DNA
973 elements in the human genome." *Nature* **489**(7414): 57-74. DOI:
974 10.1038/nature11247
975
976 Ding, N., H. Wu, T. Tao and E. Peng (2017). "NEAT1 regulates cell proliferation
977 and apoptosis of ovarian cancer by miR-34a-5p/BCL2." *Onco Targets Ther* **10**:
978 4905-4915. DOI: 10.2147/OTT.S142446
979
980 Djebali, S., C. A. Davis, A. Merkel, A. Dobin, T. Lassmann, A. Mortazavi, A. Tanzer, J.
981 Lagarde, W. Lin, F. Schlesinger, C. Xue, G. K. Marinov, J. Khatun, B. A. Williams, C.
982 Zaleski, J. Rozowsky, M. Roder, F. Kokocinski, R. F. Abdelhamid, T. Alioto, I.
983 Antoshechkin, M. T. Baer, N. S. Bar, P. Batut, K. Bell, I. Bell, S. Chakrabortty, X.
984 Chen, J. Chrast, J. Curado, T. Derrien, J. Drenkow, E. Dumais, J. Dumais, R.
985 Duttagupta, E. Falconnet, M. Fastuca, K. Fejes-Toth, P. Ferreira, S. Foissac, M. J.
986 Fullwood, H. Gao, D. Gonzalez, A. Gordon, H. Gunawardena, C. Howald, S. Jha, R.
987 Johnson, P. Kapranov, B. King, C. Kingswood, O. J. Luo, E. Park, K. Persaud, J. B.
988 Preall, P. Ribeca, B. Risk, D. Robyr, M. Sammeth, L. Schaffer, L. H. See, A. Shahab, J.
989 Skancke, A. M. Suzuki, H. Takahashi, H. Tilgner, D. Trout, N. Walters, H. Wang, J.
990 Wrobel, Y. Yu, X. Ruan, Y. Hayashizaki, J. Harrow, M. Gerstein, T. Hubbard, A.
991 Reymond, S. E. Antonarakis, G. Hannon, M. C. Giddings, Y. Ruan, B. Wold, P.
992 Carninci, R. Guigo and T. R. Gingeras (2012). "Landscape of transcription in
993 human cells." *Nature* **489**(7414): 101-108. DOI: 10.1038/nature11233
994

- 995 Gallardo, E., A. Navarro, N. Vinolas, R. M. Marrades, T. Diaz, B. Gel, A. Quera, E.
996 Bandres, J. Garcia-Foncillas, J. Ramirez and M. Monzo (2009). "miR-34a as a
997 prognostic marker of relapse in surgically resected non-small-cell lung cancer."
998 Carcinogenesis **30**(11): 1903-1909. DOI: 10.1093/carcin/bgp219
999
- 1000 Geiger, T., A. Wehner, C. Schaab, J. Cox and M. Mann (2012). "Comparative
1001 proteomic analysis of eleven common cell lines reveals ubiquitous but varying
1002 expression of most proteins." Mol Cell Proteomics **11**(3): M111 014050. DOI:
1003 10.1074/mcp.M111.014050
1004
- 1005 Granholm, V., S. Kim, J. C. Navarro, E. Sjolund, R. D. Smith and L. Kall (2014). "Fast
1006 and accurate database searches with MS-GF+Percolator." J Proteome Res **13**(2):
1007 890-897. DOI: 10.1021/pr400937n
1008
- 1009 Grossman, R. L., A. P. Heath, V. Ferretti, H. E. Varmus, D. R. Lowy, W. A. Kibbe and
1010 L. M. Staudt (2016). "Toward a Shared Vision for Cancer Genomic Data." N Engl J
1011 Med **375**(12): 1109-1112. DOI: 10.1056/NEJMp1607591
1012
- 1013 Holman, J. D., D. L. Tabb and P. Mallick (2014). "Employing ProteoWizard to
1014 Convert Raw Mass Spectrometry Data." Curr Protoc Bioinformatics **46**: 13 24 11-
1015 19. DOI: 10.1002/0471250953.bi1324s46
1016
- 1017 Hung, T., Y. Wang, M. F. Lin, A. K. Koegel, Y. Kotake, G. D. Grant, H. M. Horlings, N.
1018 Shah, C. Umbricht, P. Wang, Y. Wang, B. Kong, A. Langerod, A. L. Borresen-Dale, S.
1019 K. Kim, M. van de Vijver, S. Sukumar, M. L. Whitfield, M. Kellis, Y. Xiong, D. J. Wong
1020 and H. Y. Chang (2011). "Extensive and coordinated transcription of noncoding
1021 RNAs within cell-cycle promoters." Nat Genet **43**(7): 621-629. DOI:
1022 10.1038/ng.848
1023
- 1024 Hunten, S., M. Kaller, F. Drepper, S. Oeljeklaus, T. Bonfert, F. Erhard, A. Dueck, N.
1025 Eichner, C. C. Friedel, G. Meister, R. Zimmer, B. Warscheid and H. Hermeking
1026 (2015). "p53-Regulated Networks of Protein, mRNA, miRNA, and lncRNA
1027 Expression Revealed by Integrated Pulsed Stable Isotope Labeling With Amino
1028 Acids in Cell Culture (pSILAC) and Next Generation Sequencing (NGS) Analyses."
1029 Mol Cell Proteomics **14**(10): 2609-2629. DOI: 10.1074/mcp.M115.050237
1030
- 1031 International Human Genome Sequencing Consortium. (2004). "Finishing the
1032 euchromatic sequence of the human genome." Nature **431**(7011): 931-945. DOI:
1033 10.1038/nature03001
1034
- 1035 Johnsson, P., A. Ackley, L. Vidarsdottir, W. O. Lui, M. Corcoran, D. Grander and K.
1036 V. Morris (2013). "A pseudogene long-noncoding-RNA network regulates PTEN
1037 transcription and translation in human cells." Nat Struct Mol Biol **20**(4): 440-
1038 446. DOI: 10.1038/nsmb.2516
1039
- 1040 Katayama, S., Y. Tomaru, T. Kasukawa, K. Waki, M. Nakanishi, M. Nakamura, H.
1041 Nishida, C. C. Yap, M. Suzuki, J. Kawai, H. Suzuki, P. Carninci, Y. Hayashizaki, C.
1042 Wells, M. Frith, T. Ravasi, K. C. Pang, J. Hallinan, J. Mattick, D. A. Hume, L. Lipovich,
1043 S. Batalov, P. G. Engstrom, Y. Mizuno, M. A. Faghihi, A. Sandelin, A. M. Chalk, S.

- 1044 Mottagui-Tabar, Z. Liang, B. Lenhard, C. Wahlestedt, R. G. E. R. Group, G. Genome
1045 Science and F. Consortium (2005). "Antisense transcription in the mammalian
1046 transcriptome." *Science* **309**(5740): 1564-1566. DOI: 10.1126/science.1112009
1047
- 1048 Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler and D.
1049 Haussler (2002). "The human genome browser at UCSC." *Genome Res* **12**(6):
1050 996-1006. DOI: 10.1101/gr.229102. Article published online before print in May
1051 2002
1052
- 1053 Kim, K. H., H. J. Kim and T. R. Lee (2017). "Epidermal long non-coding RNAs are
1054 regulated by ultraviolet irradiation." *Gene* **637**: 196-202. DOI:
1055 10.1016/j.gene.2017.09.043
1056
- 1057 Kim, S. and P. A. Pevzner (2014). "MS-GF+ makes progress towards a universal
1058 database search tool for proteomics." *Nature Communications* **5**: 5277. DOI:
1059 10.1038/ncomms6277
1060
- 1061 Kong, L., Y. Zhang, Z. Q. Ye, X. Q. Liu, S. Q. Zhao, L. Wei and G. Gao (2007). "CPC:
1062 assess the protein-coding potential of transcripts using sequence features and
1063 support vector machine." *Nucleic Acids Res* **35**(Web Server issue): W345-349.
1064 DOI: 10.1093/nar/gkm391
1065
- 1066 Lal, A., M. P. Thomas, G. Altschuler, F. Navarro, E. O'Day, X. L. Li, C. Concepcion, Y.
1067 C. Han, J. Thiery, D. K. Rajani, A. Deutsch, O. Hofmann, A. Ventura, W. Hide and J.
1068 Lieberman (2011). "Capture of microRNA-bound mRNAs identifies the tumor
1069 suppressor miR-34a as a regulator of growth factor signaling." *PLoS Genet* **7**(11):
1070 e1002363. DOI: 10.1371/journal.pgen.1002363
1071
- 1072 Leveille, N., C. A. Melo, K. Rooijers, A. Diaz-Lagares, S. A. Melo, G. Korkmaz, R.
1073 Lopes, F. Akbari Moqadam, A. R. Maia, P. J. Wijchers, G. Geeven, M. L. den Boer, R.
1074 Kalluri, W. de Laat, M. Esteller and R. Agami (2015). "Genome-wide profiling of
1075 p53-regulated enhancer RNAs uncovers a subset of enhancers controlled by a
1076 lncRNA." *Nature Communications* **6**: 6520. DOI: 10.1038/ncomms7520
1077
- 1078 Liu, C., K. Kelnar, B. Liu, X. Chen, T. Calhoun-Davis, H. Li, L. Patrawala, H. Yan, C.
1079 Jeter, S. Honorio, J. F. Wiggins, A. G. Bader, R. Fagin, D. Brown and D. G. Tang
1080 (2011). "The microRNA miR-34a inhibits prostate cancer stem cells and
1081 metastasis by directly repressing CD44." *Nat Med* **17**(2): 211-215. DOI:
1082 10.1038/nm.2284
1083
- 1084 Memczak, S., M. Jens, A. Elefsinioti, F. Torti, J. Krueger, A. Rybak, L. Maier, S. D.
1085 Mackowiak, L. H. Gregersen, M. Munschauer, A. Loewer, U. Ziebold, M.
1086 Landthaler, C. Kocks, F. le Noble and N. Rajewsky (2013). "Circular RNAs are a
1087 large class of animal RNAs with regulatory potency." *Nature* **495**(7441): 333-
1088 338. DOI: 10.1038/nature11928
1089
- 1090 Ng, S. Y., G. K. Bogu, B. S. Soh and L. W. Stanton (2013). "The long noncoding RNA
1091 RMST interacts with SOX2 to regulate neurogenesis." *Mol Cell* **51**(3): 349-359.
1092 DOI: 10.1016/j.molcel.2013.07.017

- 1093
1094 Ng, S. Y., R. Johnson and L. W. Stanton (2012). "Human long non-coding RNAs
1095 promote pluripotency and neuronal differentiation by association with
1096 chromatin modifiers and transcription factors." *EMBO J* **31**(3): 522-533. DOI:
1097 10.1038/emboj.2011.459
1098
1099 O'Leary, N. A., M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad, R. McVeigh, B.
1100 Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y.
1101 Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C. M.
1102 Farrell, T. Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V. S. Joardar, V. K.
1103 Kodali, W. Li, D. Maglott, P. Masterson, K. M. McGarvey, M. R. Murphy, K. O'Neill,
1104 S. Pujar, S. H. Rangwala, D. Rausch, L. D. Riddick, C. Schoch, A. Shkeda, S. S. Storz,
1105 H. Sun, F. Thibaud-Nissen, I. Tolstoy, R. E. Tully, A. R. Vatsan, C. Wallin, D. Webb,
1106 W. Wu, M. J. Landrum, A. Kimchi, T. Tatusova, M. DiCuccio, P. Kitts, T. D. Murphy
1107 and K. D. Pruitt (2016). "Reference sequence (RefSeq) database at NCBI: current
1108 status, taxonomic expansion, and functional annotation." *Nucleic Acids Res*
1109 **44**(D1): D733-745. DOI: 10.1093/nar/gkv1189
1110
1111 Ozsolak, F., P. Kapranov, S. Foissac, S. W. Kim, E. Fishilevich, A. P. Monaghan, B.
1112 John and P. M. Milos (2010). "Comprehensive polyadenylation site maps in yeast
1113 and human reveal pervasive alternative polyadenylation." *Cell* **143**(6): 1018-
1114 1029. DOI: 10.1016/j.cell.2010.11.020
1115
1116 Polson, A., E. Durrett and D. Reisman (2011). "A bidirectional promoter reporter
1117 vector for the analysis of the p53/WDR79 dual regulatory element." *Plasmid*
1118 **66**(3): 169-179. DOI: 10.1016/j.plasmid.2011.08.004
1119
1120 Rashi-Elkeles, S., H. J. Warnatz, R. Elkon, A. Kupershtain, Y. Chobod, A. Paz, V.
1121 Amstislavskiy, M. Sultan, H. Safer, W. Nietfeld, H. Lehrach, R. Shamir, M. L. Yaspo
1122 and Y. Shiloh (2014). "Parallel profiling of the transcriptome, cistrome, and
1123 epigenome in the cellular response to ionizing radiation." *Sci Signal* **7**(325): rs3.
1124 DOI: 10.1126/scisignal.2005032
1125
1126 Raver-Shapira, N., E. Marciano, E. Meiri, Y. Spector, N. Rosenfeld, N. Moskovits, Z.
1127 Bentwich and M. Oren (2007). "Transcriptional activation of miR-34a
1128 contributes to p53-mediated apoptosis." *Mol Cell* **26**(5): 731-743. DOI:
1129 10.1016/j.molcel.2007.05.017
1130
1131 Rinn, J. L., M. Kertesz, J. K. Wang, S. L. Squazzo, X. Xu, S. A. Brugmann, L. H.
1132 Goodnough, J. A. Helms, P. J. Farnham, E. Segal and H. Y. Chang (2007).
1133 "Functional demarcation of active and silent chromatin domains in human HOX
1134 loci by noncoding RNAs." *Cell* **129**(7): 1311-1323. DOI:
1135 10.1016/j.cell.2007.05.022
1136
1137 Rokavec, M., M. G. Oner, H. Li, R. Jackstadt, L. Jiang, D. Lodygin, M. Kaller, D. Horst,
1138 P. K. Ziegler, S. Schwitalla, J. Slotta-Huspenina, F. G. Bader, F. R. Greten and H.
1139 Hermeking (2015). "Corrigendum. IL-6R/STAT3/miR-34a feedback loop
1140 promotes EMT-mediated colorectal cancer invasion and metastasis." *J Clin Invest*
1141 **125**(3): 1362. DOI: 10.1172/JCI81340

- 1142
1143 Schneider, C. A., W. S. Rasband and K. W. Eliceiri (2012). "NIH Image to ImageJ:
1144 25 years of image analysis." Nat Methods **9**(7): 671-675.
1145
1146 Serviss, J. T. (2017). miR34AasRNaproject.
1147 https://github.com/GranderLab/miR34a_asRNA_project
1148
1149 Serviss, J. T., P. Johnsson and D. Grander (2014). "An emerging role for long non-
1150 coding RNAs in cancer metastasis." Front Genet **5**: 234. DOI:
1151 10.3389/fgene.2014.00234
1152
1153 Slabakova, E., Z. Culig, J. Remsik and K. Soucek (2017). "Alternative mechanisms
1154 of miR-34a regulation in cancer." Cell Death Dis **8**(10): e3100. DOI:
1155 10.1038/cddis.2017.495
1156
1157 Stahlhut, C. and F. J. Slack (2015). "Combinatorial Action of MicroRNAs let-7 and
1158 miR-34 Effectively Synergizes with Erlotinib to Suppress Non-small Cell Lung
1159 Cancer Cell Proliferation." Cell Cycle **14**(13): 2171-2180. DOI:
1160 10.1080/15384101.2014.1003008
1161
1162 Su, X., D. Chakravarti, M. S. Cho, L. Liu, Y. J. Gi, Y. L. Lin, M. L. Leung, A. El-Naggar,
1163 C. J. Creighton, M. B. Suraokar, I. Wistuba and E. R. Flores (2010). "TAp63
1164 suppresses metastasis through coordinate regulation of Dicer and miRNAs."
1165 Nature **467**(7318): 986-990. DOI: 10.1038/nature09459
1166
1167 Sun, F., H. Fu, Q. Liu, Y. Tie, J. Zhu, R. Xing, Z. Sun and X. Zheng (2008).
1168 "Downregulation of CCND1 and CDK6 by miR-34a induces cell cycle arrest."
1169 FEBS Lett **582**(10): 1564-1568. DOI: 10.1016/j.febslet.2008.03.057
1170
1171 Tarasov, V., P. Jung, B. Verdoodt, D. Lodygin, A. Epanchintsev, A. Menssen, G.
1172 Meister and H. Hermeking (2007). "Differential regulation of microRNAs by p53
1173 revealed by massively parallel sequencing: miR-34a is a p53 target that induces
1174 apoptosis and G1-arrest." Cell Cycle **6**(13): 1586-1593. DOI:
1175 10.4161/cc.6.13.4436
1176
1177 Team, R. C. (2017). "R: A Language and Environment for Statistical Computing."
1178 from <https://www.R-project.org/>.
1179
1180 Therneau, T. M. (2015). A Package for Survival Analysis in S. version 2.38.
1181 <https://CRAN.R-project.org/package=survival>
1182
1183 Turner, A. M., A. M. Ackley, M. A. Matrone and K. V. Morris (2012).
1184 "Characterization of an HIV-targeted transcriptional gene-silencing RNA in
1185 primary cells." Hum Gene Ther **23**(5): 473-483. DOI: 10.1089/hum.2011.165
1186
1187 Vogt, M., J. Mundig, M. Gruner, S. T. Liffers, B. Verdoodt, J. Hauk, L.
1188 Steinstraesser, A. Tannapfel and H. Hermeking (2011). "Frequent concomitant
1189 inactivation of miR-34a and miR-34b/c by CpG methylation in colorectal,
1190 pancreatic, mammary, ovarian, urothelial, and renal cell carcinomas and soft

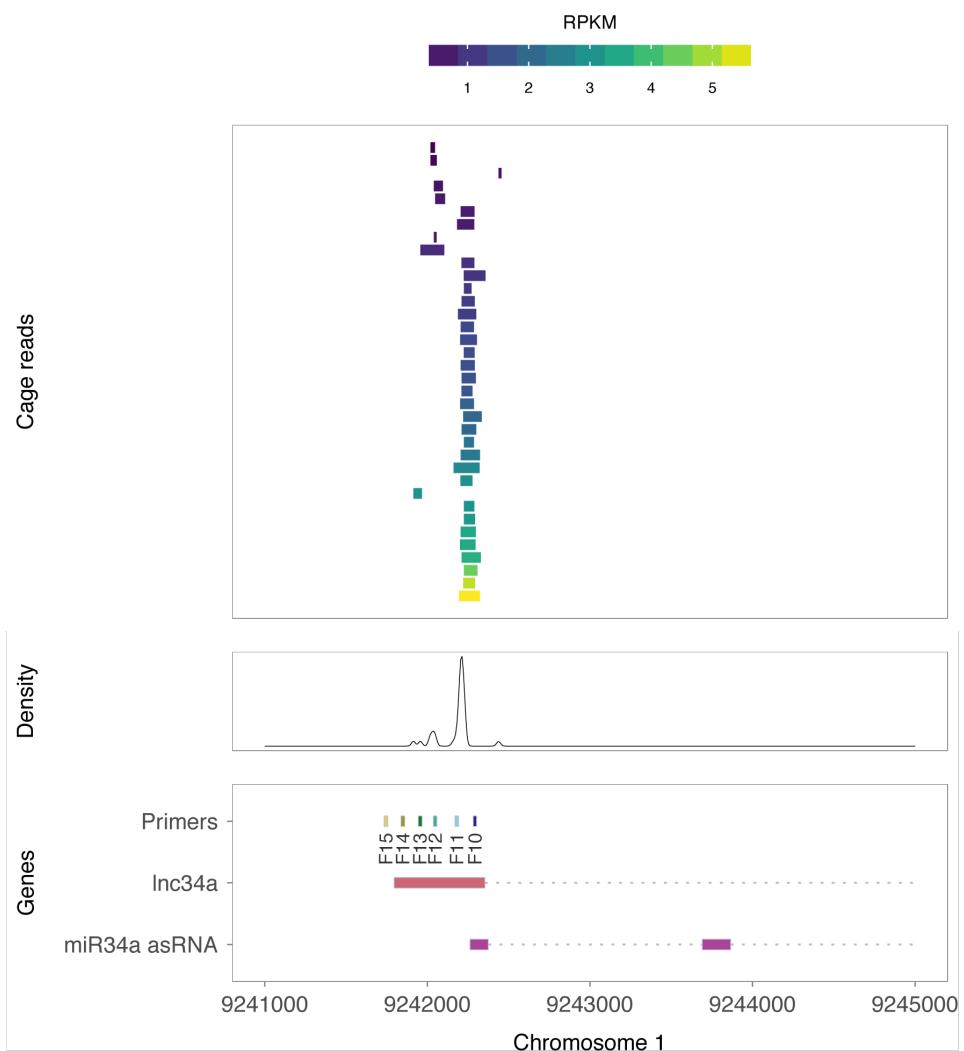
1191 tissue sarcomas." *Virchows Arch* **458**(3): 313-322. DOI: 10.1007/s00428-010-
1192 1030-5
1193
1194 Wang, L., P. Bu, Y. Ai, T. Srinivasan, H. J. Chen, K. Xiang, S. M. Lipkin and X. Shen
1195 (2016). "A long non-coding RNA targets microRNA miR-34a to regulate colon
1196 cancer stem cell asymmetric division." *eLife* **5**. DOI: 10.7554/eLife.14620
1197
1198 Wang, L., H. J. Park, S. Dasari, S. Wang, J. P. Kocher and W. Li (2013). "CPAT:
1199 Coding-Potential Assessment Tool using an alignment-free logistic regression
1200 model." *Nucleic Acids Res* **41**(6): e74. DOI: 10.1093/nar/gkt006
1201
1202 Wang, X., J. Li, K. Dong, F. Lin, M. Long, Y. Ouyang, J. Wei, X. Chen, Y. Weng, T. He
1203 and H. Zhang (2015). "Tumor suppressor miR-34a targets PD-L1 and functions
1204 as a potential immunotherapeutic target in acute myeloid leukemia." *Cell Signal*
1205 **27**(3): 443-452. DOI: 10.1016/j.cellsig.2014.12.003
1206
1207 Wickham, H. (2016). gtable: Arrange 'Grobs' in Tables. R package version 0.2.0.
1208 <https://CRAN.R-project.org/package=gtable>
1209
1210 Wickham, H. (2017). scales: Scale Functions for Visualization. R package version
1211 0.5.0. <https://CRAN.R-project.org/package=scales>
1212
1213 Wickham, H. (2017). tidyverse: Easily Install and Load the 'Tidyverse'. R package
1214 version 1.2.1. <https://CRAN.R-project.org/package=tidyverse>
1215
1216 Wickham, L. H. a. H. (2017). rlang: Functions for Base Types and Core R and
1217 'Tidyverse' Features. R package version 0.1.4. [https://CRAN.R-
1218 project.org/package=rlang](https://CRAN.R-project.org/package=rlang)
1219
1220 Wickham, S. M. B. a. H. (2014). magrittr: A Forward-Pipe Operator for R. R
1221 package version 1.5. <https://CRAN.R-project.org/package=magrittr>
1222
1223 Wilkins, D. ggenes: Draw Gene Arrow Maps in 'ggplot2'. R package version
1224 0.2.0.9003. <https://github.com/wilcox/ggenes>
1225
1226 Xiao, N. (2017). liftr: Containerize R Markdown Documents. R package version
1227 0.7. <https://CRAN.R-project.org/package=liftr>
1228
1229 Xie, Y. (2017). knitr: A General-Purpose Package for Dynamic Report Generation
1230 in R. R package version 1.17. <https://yihui.name/knitr/>
1231
1232 Yang, P., Q. J. Li, Y. Feng, Y. Zhang, G. J. Markowitz, S. Ning, Y. Deng, J. Zhao, S.
1233 Jiang, Y. Yuan, H. Y. Wang, S. Q. Cheng, D. Xie and X. F. Wang (2012). "TGF-beta-
1234 miR-34a-CCL22 signaling-induced Treg cell recruitment promotes venous
1235 metastases of HBV-positive hepatocellular carcinoma." *Cancer Cell* **22**(3): 291-
1236 303. DOI: 10.1016/j.ccr.2012.07.023
1237
1238 Yap, K. L., S. Li, A. M. Munoz-Cabello, S. Raguz, L. Zeng, S. Mujtaba, J. Gil, M. J.
1239 Walsh and M. M. Zhou (2010). "Molecular interplay of the noncoding RNA ANRIL

1240 and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional
1241 silencing of INK4a." Mol Cell **38**(5): 662-674. DOI: 10.1016/j.molcel.2010.03.021
1242
1243 Yu, W., D. Gius, P. Onyango, K. Muldoon-Jacobs, J. Karp, A. P. Feinberg and H. Cui
1244 (2008). "Epigenetic silencing of tumour suppressor gene p15 by its antisense
1245 RNA." Nature **451**(7175): 202-206. DOI: 10.1038/nature06468
1246
1247 Zenz, T., J. Mohr, E. Eldering, A. P. Kater, A. Buhler, D. Kienle, D. Winkler, J. Durig,
1248 M. H. van Oers, D. Mertens, H. Dohner and S. Stilgenbauer (2009). "miR-34a as
1249 part of the resistance network in chronic lymphocytic leukemia." Blood **113**(16):
1250 3801-3808. DOI: 10.1182/blood-2008-08-172254
1251
1252

Supplementary Document 1

An unannotated transcript, *Lnc34a*, arising from the antisense orientation of the *miR34a* locus and with a transcription start site >250 bp upstream of the annotated *IncTAM34as* start site, has been previously reported in a study examining colorectal cancer (Wang et al. 2016). Among the findings in Wang et al. the authors discover that *Lnc34a* negatively regulates miR34a expression via recruitment of DNMT3a, PHB2, and HDAC1 to the *miR34a* promoter. Although the *Lnc34a* and *IncTAM34a* transcripts share some sequence similarity, we believe them to be separate RNAs transcripts. Furthermore, *Lnc34a* may be highly context dependent and potentially only expressed at biologically significant levels in colon cancer stem cells, or other stem-like cells, in agreement with the conclusions drawn in the paper.

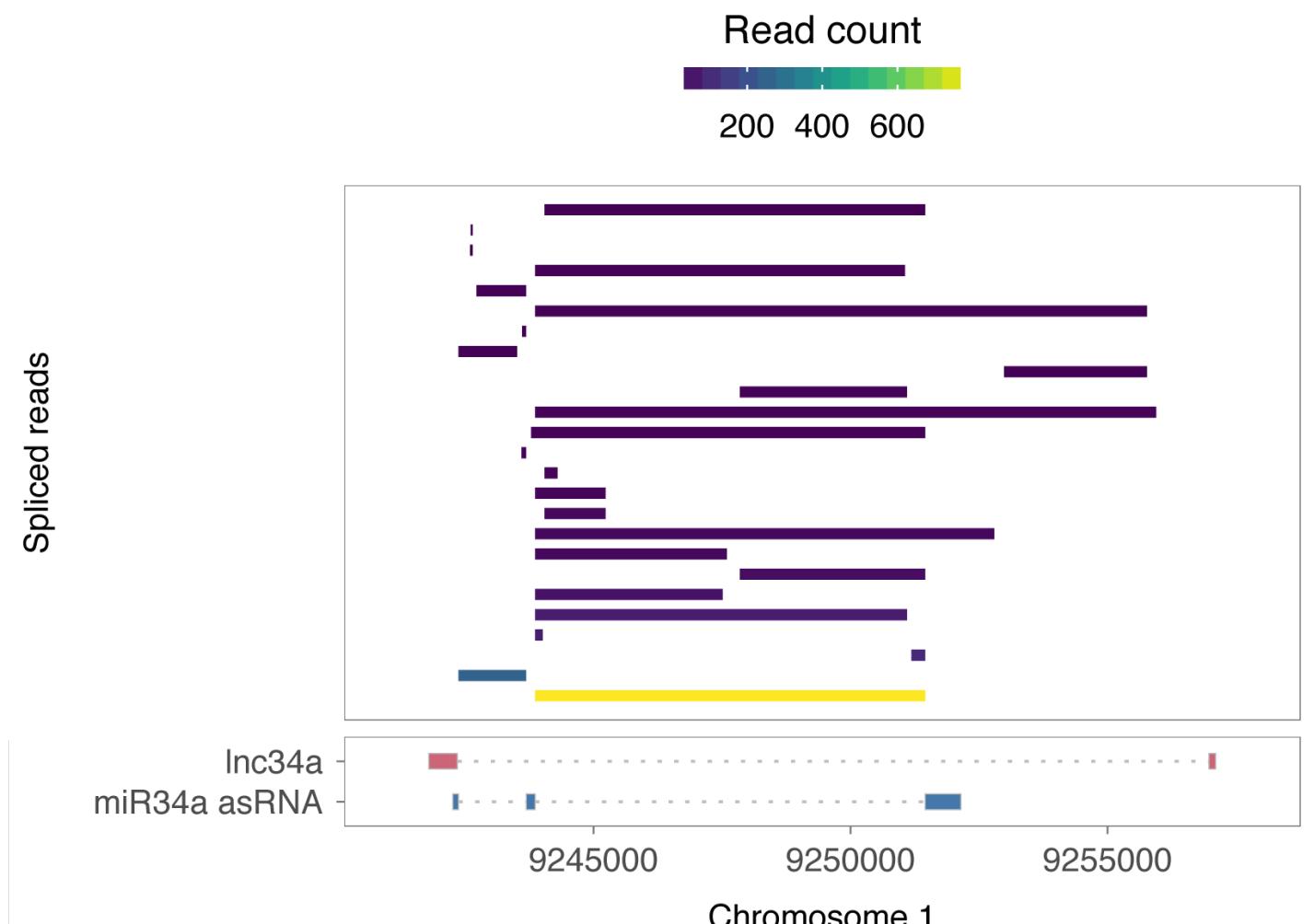
Several lines of evidence point to the fact that *IncTAM34a* and *Lnc34a* are not the same transcript and, in addition, that *Lnc34a* expression may be confined to a highly specific subset of colorectal cancer stem cells (CCSC). First, we were unable to detect transcription upstream of the 5' start site that was defined in the primer walk assay (**Fig. 1E** and **Supplementary Fig. 1B**) although the reported *Lnc34a* start site is >250 base pairs upstream of the F12 primer used in this assay. This could simply be due to the fact that *Lnc34a* is not expressed in HEK293t cells in which the assay was performed. To further investigate the existence of transcription start sites in the antisense orientation of the *miR34a* locus, we interrogated CAGE data from 28 cell lines.



Supplementary Document 1a: All available CAGE data from the ENCODE project for 36 cell lines was downloaded from the UCSC genome browser for genome version hg19. Of these, 28 cell lines had CAGE transcription start sites mapping to the plus strand of chromosome 1 and in regions corresponding to 200 base pairs upstream of the *Lnc34a* start site (9241796 - 200) and 200 base pairs upstream of the GENCODE annotated *IncTAM34a* start site (9242263 + 200). These cell lines included: HFDPC, H1-hESC, HMEpC, HAoEC, HPIEpC, HSaVEC, GM12878, hMSC-BM, HUVEC, AG04450, hMSC-UC, IMR90, NHDF, SK-N-SH_RA, BJ, HOB, HPC-PL, HAoAF, NHEK, HVMF, HWP, MCF-7, HepG2, hMSC-AT, NHEM.f_M2, SkMC, NHEM_M2, and HCH. In total 74 samples were included. 17 samples were polyA-, 47 samples were polyA+, and 10 samples were total RNA. In addition, 34 samples were whole cell, 15 enriched for the cytosolic fraction, 10 enriched for the nucleolus, and 15 enriched for the nucleus. All CAGE reads were plotted and the RPKM of the individual reads was used to colour each read to indicate their relative abundance (top panel). In addition, a density plot (middle panel) shows the distribution of the CAGE reads in the specified interval and the transcription start regions for *Lnc34a* and *IncTAM34a* as well as primer positions from the primer walk assay (bottom panel).

The results show a high density of CAGE tags aligning to the region corresponding to the annotated *IncTAM34a* start site. Several additional peaks, albeit with a much lower average expression, aligns slightly upstream of the annotated *IncTAM34a* start site, one of which, corresponds to the upstream start site detected in our primer walk analysis (**Figure 1e**). Despite this, we find no CAGE tags aligning at or upstream of the transcription start site of the *Lnc34a* transcript. This potentially indicates that *Lnc34a* is tightly regulated and specifically expressed in the CCSC context, as suggested by the authors. An alternative interpretation could be that *Lnc34a* expression is present in a subset of the examined cell lines although at levels too low to be detected. Finally, *Lnc34a* may not be 5'-capped precluding its detection by CAGE.

In order to detect *Lnc34a* expression in a manner that is not dependant on 5'-capping, we proceeded to examine spliced RNA sequencing reads from 36 cell lines, taking advantage of the fact that *Lnc34a* has an exon which is not present in any annotated or PCR cloned *IncTAM34a* isoforms.



Supplementary Document 1b: All available whole cell (i.e. non-fractionated) spliced read data originating from the Cold Spring Harbor Lab in the ENCODE project for 38 cell lines was downloaded from the UCSC genome browser. Of these cell lines, 36 had spliced reads mapping to the plus strand of chromosome 1 and in the region between the *Lnc34a* start (9241796) and transcription termination (9257102) site (note that *LncTAM34a* resides totally within this region). Splice junctions from the following cell lines were included in the final figure: A549, Ag04450, Bj, CD20, CD34 mobilized, Gm12878, H1hesc, Haoaf, Haoec, Hch, Helas3, Hepg2, Hfdpc, Hmec, Hmepc, Hmscat, Hmscbm, Hmscuc, Hob, Hpcpl, Hpiepc, Hsavec, Hsomm, Huvec, Hvmaf, Hwp, Imr90, Mcf7, Monocd14, Nhdf, Nhek, Nhemfm2, Nhemm2, Nhlf, Skmc, and Sknsh. All splice junctions were included in the figure and coloured according to the number of reads corresponding to each (top panel). In cases where the exact same read was detected multiple times the read count was summed and represented as one read in the figure. *LncTAM34a* and *Lnc34a* transcripts are represented for reference (bottom panel).

These results indicate that, although splice junctions corresponding to the annotated *IncTAM34a* transcript and multiple isoforms found via PCR cloning were detected, the data give no support for the presence of the splice junction between the first and second exon of *Lnc34a*. In summary, these results indicate that *Lnc34a* is unlikely to represent the same asRNA transcript as *IncTAM34a* and that its expression may be confined to CCSCs.

In addition, there are several other lines of evidence indicating that the asRNA described in our paper is distinct from *Lnc34a*. We noted several relevant comments in the public review that was published in conjunction with the work by Wang et al. The authors mention, and provide data, indicating that *Lnc34a* expression is not changed upon ectopic expression of TP53. In contrast, *IncTAM34a* is strongly regulated by TP53 as the evidence shows in our, as well as, others findings (Léveillé 2015, Rashi-Elkeles 2014, Hünten 2015, Ashouri 2016, Kim 2017). Furthermore, Wang et al. also mention in the public review that *Lnc34a* has a low expression level in HCT116 cells although we detect robust expression of *IncTAM34a* in this cell type (**Figure 1b**).

In summary, these results indicate that *Lnc34a* expression is not present in the cell types examined where there exists strong evidence for the presence *IncTAM34a* and, for these reasons, we believe *IncTAM34a* and *Lnc34a* to be individual antisense RNA transcripts.

Supplementary Document 2

Figure 1b

name	sequence
miR34a asRNA F1	AGC GGC ATC TCC TCC ACC TGA AA
miR34a asRNA R1	TTG CCT CGT GAG TCC AAG GAG AAT
miR34a HG F	TCT GCT CCA GTG GCT GAT GAG AAA
miR34a HG R	GTT CAC TGG CCT CAA AGT TGG CAT
β-actin Fwd	AGG TCA TCA CCA TTG GCA ATG AG
β-actin Rev	CTT TGC GGA TGT CCA CGT CA

Figure 1d

name	sequence
miR34a asRNA F10	ACG CGT CTC TCC AGC CCG GGA T
polyT T7 FAM	CAG TGA ATT GTA ATA CGA CTC ACT ATA GGG ACA TCC GTA GCT CGT CCA GGA CCC TTT TTT TTT TTT TTT VN
miR34a asRNA F1	AGC GGC ATC TCC TCC ACC TGA AA
FAM primer	CCG TAG CTC GTC CAG GAC CC

Figure 2a

name	sequence
β-actin Fwd	AGG TCA TCA CCA TTG GCA ATG AG
β-actin Rev	CTT TCG GGA TGT CCA CGT CA
miR34a HG F	TCT GCT CCA GTG GCT GAT GAG AAA
miR34a HG R	GTT CAC TGG CCT CAA AGT TGG CAT
miR34a AS F1	AGC GGC ATC TCC TCC ACC TGA AA
miR34a AS R1	TTG CCT CGT GAG TCC AAG GAG AAT

Figure 2b

name	sequence
β-actin Fwd	AGG TCA TCA CCA TTG GCA ATG AG
β-actin Rev	CTT TCG GGA TGT CCA CGT CA
miR34a HG F	TCT GCT CCA GTG GCT GAT GAG AAA
miR34a HG R	GTT CAC TGG CCT CAA AGT TGG CAT
miR34a AS F1	AGC GGC ATC TCC TCC ACC TGA AA
miR34a AS R1	TTG CCT CGT GAG TCC AAG GAG AAT

Figure 2c

name	sequence
Luc setII F	AAG ATT CAA AGT GCG CTG CTG
Luc setII R	TTG CCT GAT ACC TGG CAG ATG
Renilla pBiDir F1	TAA CGC GGC CTC TTC TTA TTT
Renilla pBiDir R1	GAT TTG CCT GAT TTG CCC ATA
β-actin Fwd	AGG TCA TCA CCA TTG GCA ATG AG
β-actin Rev	CTT TGC GGA TGT CCA CGT CA

Figure 2d

name	sequence
Luc setII F	AAG ATT CAA AGT GCG CTG CTG
Luc setII R	TTG CCT GAT ACC TGG CAG ATG
Renilla pBiDir F1	TAA CGC GGC CTC TTC TTA TTT
Renilla pBiDir R1	GAT TTG CCT GAT TTG CCC ATA
β-actin Fwd	AGG TCA TCA CCA TTG GCA ATG AG
β-actin Rev	CTT TGC GGA TGT CCA CGT CA

Figure 3a

Cloning primers

name	sequence
miR34aAS cloning F4	ACG CGT CTC TCC AGC CCG GGA T
miR34aAS cloning Ex3_1	AAT GAT GGC CGC AAC TAA TGA CGG

QPCR primers

name	sequence
β-actin Fwd	AGG TCA TCA CCA TTG GCA ATG AG
β-actin Rev	CTT TCG GGA TGT CCA CGT CA
miR34a AS F1	AGC GGC ATC TCC TCC ACC TGA AA
miR34a AS R1	TTG CCT CGT GAG TCC AAG GAG AAT

Figure 3d

name	sequence
miR34a ChIP F1	AAA GTT TGC AAA GAA GGA GGC GGG
miR34a ChIP R1	AGG GAA GAA AGA ACT AGC CGA GCA

Figure 1 Supplement 2a

name	sequence
miR34a AS F10	ACG CGT CTC TCC AGC CCG GGA T
miR34a AS F11	ATC TGC GTG GTC ACC GAG AAG CA
miR34a AS F12	CGC ACG GAC TGA GAA ACA CAA G
miR34a AS F13	ACG GAG GCT ACA CAA TTG AAC AGG
miR34a AS F14	AGG GAA GAA AGA ACT AGC CGA GCA
miR34a AS F15	CAT TTG CTG CAA TAT CAC CGT GGC
miR34a AS R1	TTG CCT CGT GAG TCC AAG GAG AAT

Figure 1 Supplement 2b

name	sequence
miR34a AS F1	AGC GGC ATC TCC TCC ACC TGA AA
miR34a AS R1	TTG CCT CGT GAG TCC AAG GAG AAT
miR34a AS int1 R1	TGC GCA AAC TAC GCG CTC T
miR34a HG F	TCT GCT CCA GTG GCT GAT GAG AAA
miR34a HG R	GTT CAC TGG CCT CAA AGT TGG CAT
β-actin Fwd	AGG TCA TCA CCA TTG GCA ATG AG
β-actin Rev	CTT TGC GGA TGT CCA CGT CA
U48 F	AGT GAT GAT GAC CCC AGG TA
U48 R	GGT CAG AGC GCT GCG GTG AT

Figure 1 Supplement 2c

name	sequence
miR34a AS F12	AAA CAC AAG CGT TTA CCT GGG TGC
miR34a AS R1	TTG CCT CGT GAG TCC AAG GAG AAT
miR34a AS R2	ATA GGT TCA TTT GCC CGA TGT GCC
miR34a AS R3	CCA CAG CTG TTG CTT CTG AAT GCT
miR34a AS Ex3 R1	TGA TGG CCG CAA CTA ATG ACG GAT

Figure 2 Supplement 2a

name	sequence
Luc setII F	AAG ATT CAA AGT GCG CTG CTG
Luc setII R	TTG CCT GAT ACC TGG CAG ATG
Renilla pBiDir F1	TAA CGC GGC CTC TTC TTA TTT
Renilla pBiDir R1	GAT TTG CCT GAT TTG CCC ATA
β-actin Fwd	AGG TCA TCA CCA TTG GCA ATG AG
β-actin Rev	CTT TGC GGA TGT CCA CGT CA

Figure 3 Supplement 1a

Cloning primers

name	sequence
miR34aAS cloning F4	ACG CGT CTC TCC AGC CCG GGA T
miR34aAS cloning Ex3_1	AAT GAT GGC CGC AAC TAA TGA CGG

QPCR primers

name	sequence
β-actin Fwd	AGG TCA TCA CCA TTG GCA ATG AG
β-actin Rev	CTT TCG GGA TGT CCA CGT CA
miR34a AS F1	AGC GGC ATC TCC TCC ACC TGA AA
miR34a AS R1	TTG CCT CGT GAG TCC AAG GAG AAT

Figure 3 Supplement 2a

name	sequence
CCND1 Fwd	CGT GGC CTC TAA GAT GAA GG
CCND1 Rev	CTG GCA TTT TGG AGA GGA AG
β-actin Fwd	AGG TCA TCA CCA TTG GCA ATG AG
β-actin Rev	CTT TGC GGA TGT CCA CGT CA