

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 Johansson Andrews¹, Miranda Houtman^{1,4}, Mattias**
6 **Vesterlund¹, Laura Schwarzmüller^{1,5}, Per Johnsson^{6,7}, Erik Larsson³,**
7 **Dan Grandér¹, Katja Pokrovskaia 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 and has 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 wildtype settings.

40

41 **Introduction**

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

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

64

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

89

90 Studies across multiple cancer types have reported a decrease in oncogenic
91 phenotypes when *miR34a* expression is induced in a *TP53*-null background,
92 although endogenous mechanisms for achieving this have not yet been
93 discovered (Liu et al. 2011, Ahn et al. 2012, Yang et al. 2012, Stahlhut et al.
94 2015, Wang et al. 2015). In addition, previous reports from large-scale studies
95 interrogating global *TP53*-mediated regulation of lncRNAs have identified a
96 lncRNA (known as RP3-510D11.2 and LINC01759) originating in the
97 antisense orientation from the *miR34a* locus which is induced upon numerous
98 forms of cellular stress (Rashi-Elkeles et al. 2014, Hunten et al. 2015, Leveille
99 et al. 2015, Ashouri et al. 2016, Kim et al. 2017). Despite this, none of these
100 studies have functionally characterized this transcript, which we name long-
101 non-coding Transcriptional Activator of MiR34a-34a (lncTAM34a). In this
102 study we functionally characterize the *lncTAM34a* transcript, and find that it
103 positively regulates *miR34a* expression resulting in a decrease of several
104 tumorigenic phenotypes. Furthermore, we find that *lncTAM34a*-mediated up-
105 regulation of *miR34a* is sufficient to induce endogenous cellular mechanisms
106 counteracting several types of stress stimuli in a *TP53*-deficient background.
107 Finally, similar to the functional roles of antisense transcription at protein-
108 coding genes, we identify a rare example of an antisense RNA capable of
109 regulating a cancer-associated miRNA.

110

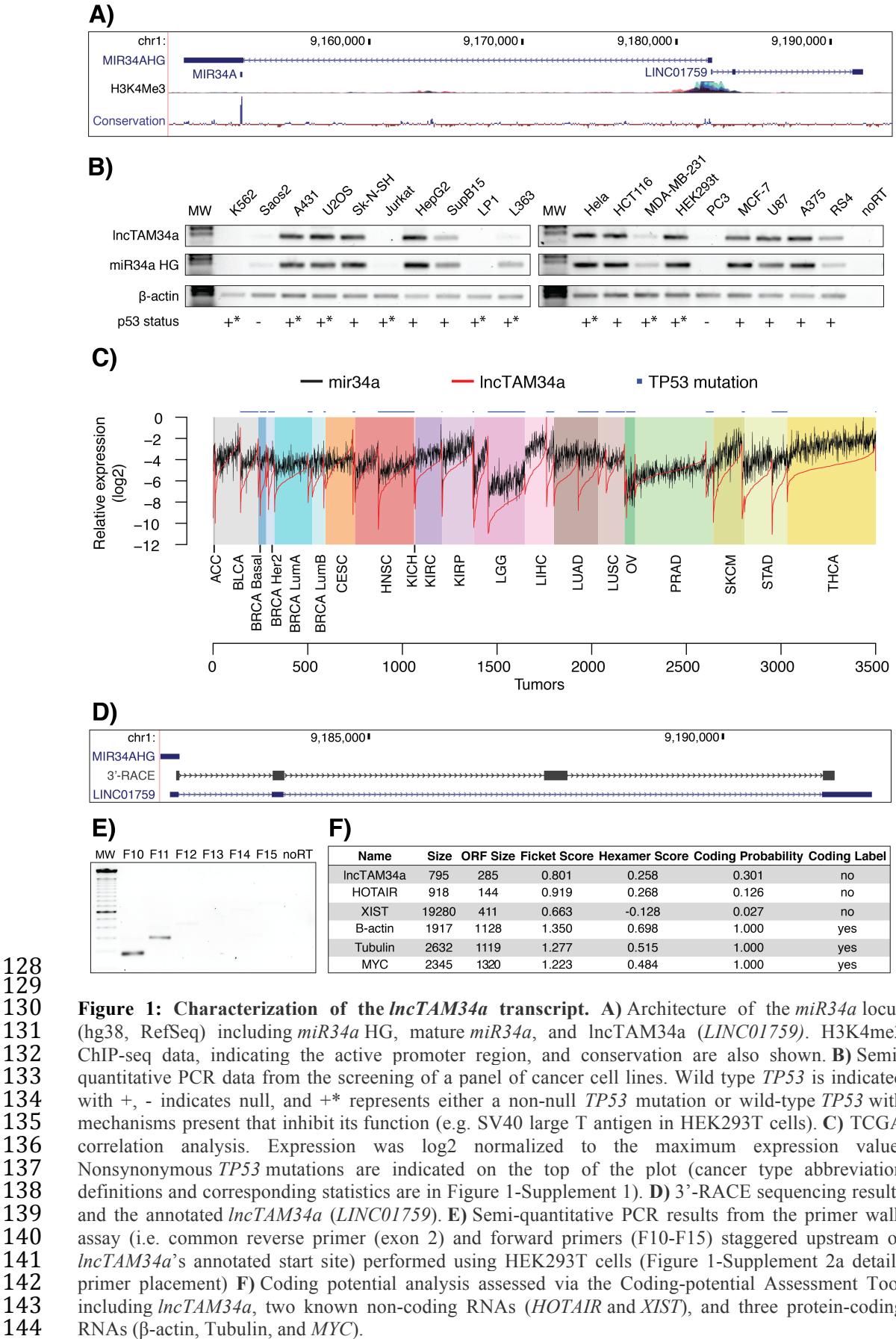
111 **Results**

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

115
116 *lncTAM34a* is transcribed in a “head-to-head” orientation with approximately
117 100 base pair overlap with the *miR34a* host gene (HG) (**Fig. 1a**). Due to the

fact that sense/antisense pairs can be both concordantly and discordantly expressed, we sought to evaluate this relationship in the case of *miR34a* HG and its asRNA. Using a diverse panel of cancer cell lines, we detected co-expression of both the *miR34a* HG and *lncTAM34a* (**Fig. 1b**). We used cell lines with a known *TP53* status in the panel due to previous reports that *miR34a* is a known downstream target of *TP53*. These results indicate that *miR34a* HG and *lncTAM34a* are co-expressed and that their expression levels correlate with *TP53* status, with *TP53*^{-/-} cells tending to have decreased or undetectable expression of both transcripts.

127



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

156

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

170 to 220bp (F12 primer) upstream of the annotated start site (**Fig. 1e**).
171 Polyadenylation status was evaluated via cDNA synthesis with either random
172 nanomers or oligo(DT) primers followed by semi-quantitative PCR which
173 showed that the *lncTAM34a* is polyadenylated although the unspliced form
174 seems to only be present in a polyadenylation negative state (**Figure 1-**
175 **Figure Supplement 2b**). Furthermore, we investigated the propensity
176 of *lncTAM34a* to be alternatively spliced in U2OS cells, using PCR cloning
177 followed by sequencing and found that the transcript is post-transcriptionally
178 spliced to form multiple isoforms (**Figure 1-Figure Supplement 2c**). In order
179 to evaluate the subcellular localization of *lncTAM34a*, we made use of RNA
180 sequencing data from five cancer cell lines included in the ENCODE
181 (Consortium 2012) project that had been fractionated into cytosolic and
182 nuclear fractions. The analysis revealed that the *lncTAM34a* transcript
183 primarily localizes to the nucleus with only a minor fraction in the cytosol
184 (**Figure 1-Figure Supplement 2d**).

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

195 features (**Figure 1-Figure Supplement 2e**) (Kong et al. 2007). Finally, we
196 downloaded mass spectrometry spectra for 11 cancer cell lines (Geiger et al.
197 2012), 7 of which were also present in the cell line panel above (**Fig. 1b**), and
198 searched it against a database of human protein sequences which also
199 contained the 6 frame translation of *IncTAM34a*. However, we did not manage
200 to detect any peptides matching the sequence in any of the 11 cell lines.
201 Taken together our results indicate that *IncTAM34a* is not a coding transcript
202 and that it is not translated to any significant degree.

203

204 **TP53-mediated regulation of *IncTAM34a* expression**

205 *miR34a* is a known downstream target of TP53 and has been previously
206 shown to exhibit increased expression within multiple contexts of cellular
207 stress. *IncTAM34a* has also been shown to be induced upon TP53 activation
208 in several global analyses of TP53-regulated lncRNAs (Rashi-Elkeles et al.
209 2014, Hunten et al. 2015, Leveille et al. 2015, Ashouri et al. 2016, Kim et al.
210 2017). To confirm these results in our biological systems, we treated
211 HEK293T, embryonic kidney cells, and HCT116, colorectal cancer cells, with
212 the DNA damaging agent doxorubicin to activate TP53. QPCR-mediated
213 measurements of both *miR34a* HG and asRNA indicated that their expression
214 levels were increased in response to doxorubicin treatment in both cell lines
215 (**Fig. 2a**). To assess whether TP53 was responsible for the increase
216 in *IncTAM34a* expression upon DNA damage, we treated *TP53^{+/+}* and *TP53^{-/-}*
217 HCT116 cells with increasing concentrations of doxorubicin and monitored the
218 expression of both *miR34a* HG and *IncTAM34a*. We observed a dose-
219 dependent increase in both *miR34a* HG and *IncTAM34a* expression levels

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

227

228
229
230
231
232
233
234
235
236
237
238
239
240
241

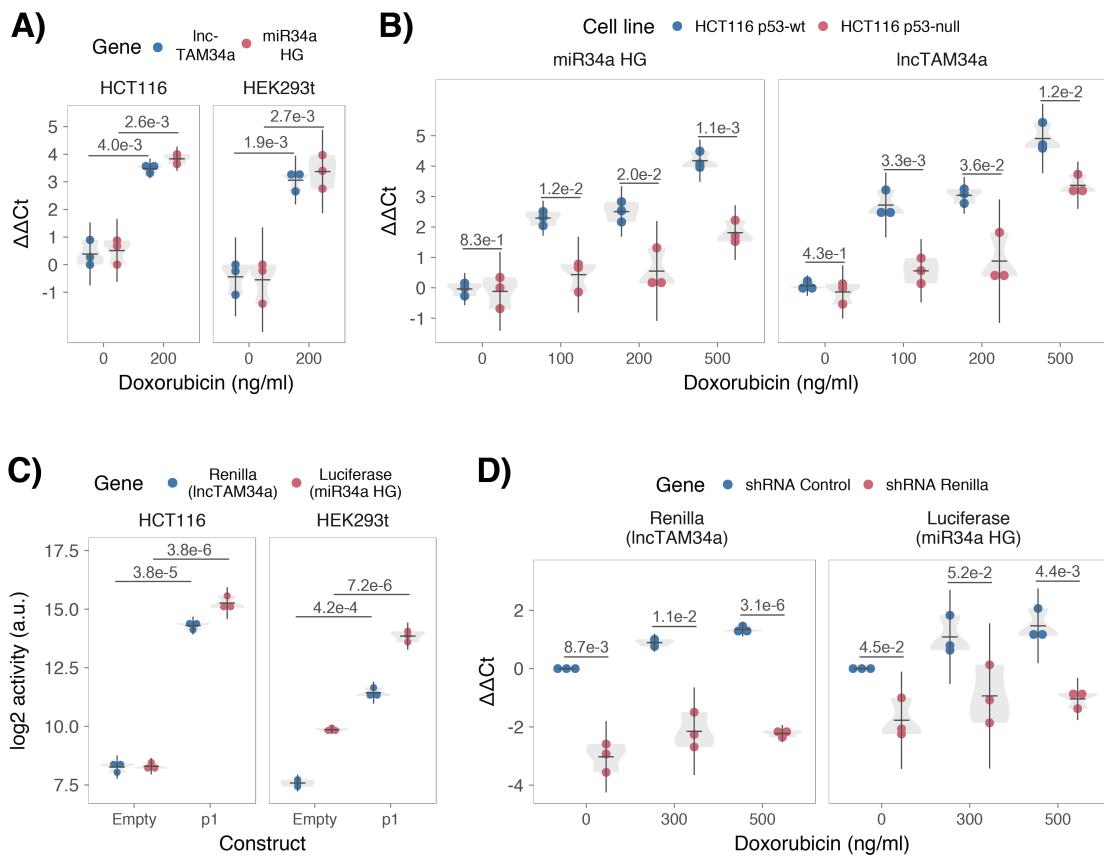


Figure 2: TP53-mediated regulation of the miR34a locus. **A)** Evaluating the effects of 24 hours of treatment with 200 ng/ml doxorubicin on *LncTAM34a* and *miR34a* HG in HCT116 and HEK293T cells.* **B)** Monitoring *miR34a* HG and *LncTAM34a* expression levels during 24 hours of doxorubicin treatment 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.

242 The head-to-head orientation of *miR34a* HG and *lncTAM34a*, suggests that
243 transcription is initiated from a single promoter in a bi-directional manner (**Fig**
244 **1a**). To investigate whether *miR34a* HG and *lncTAM34a* are transcribed from
245 the same promoter as divergent transcripts, we cloned the previously reported
246 *miR34a* HG promoter, including the TP53 binding site, into a luciferase/renilla
247 dual reporter vector which we hereafter refer to as p1 (**Figure 2-Figure**
248 **Supplement 1a-b**) (Raver-Shapira et al. 2007). Upon transfection of p1 into
249 HCT116 and HEK293T cell lines we observed increases in both luciferase
250 and renilla indicating that *miR34a* HG and *lncTAM34a* expression can be
251 regulated by a single promoter contained within the p1 construct (**Fig. 2c**).
252

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

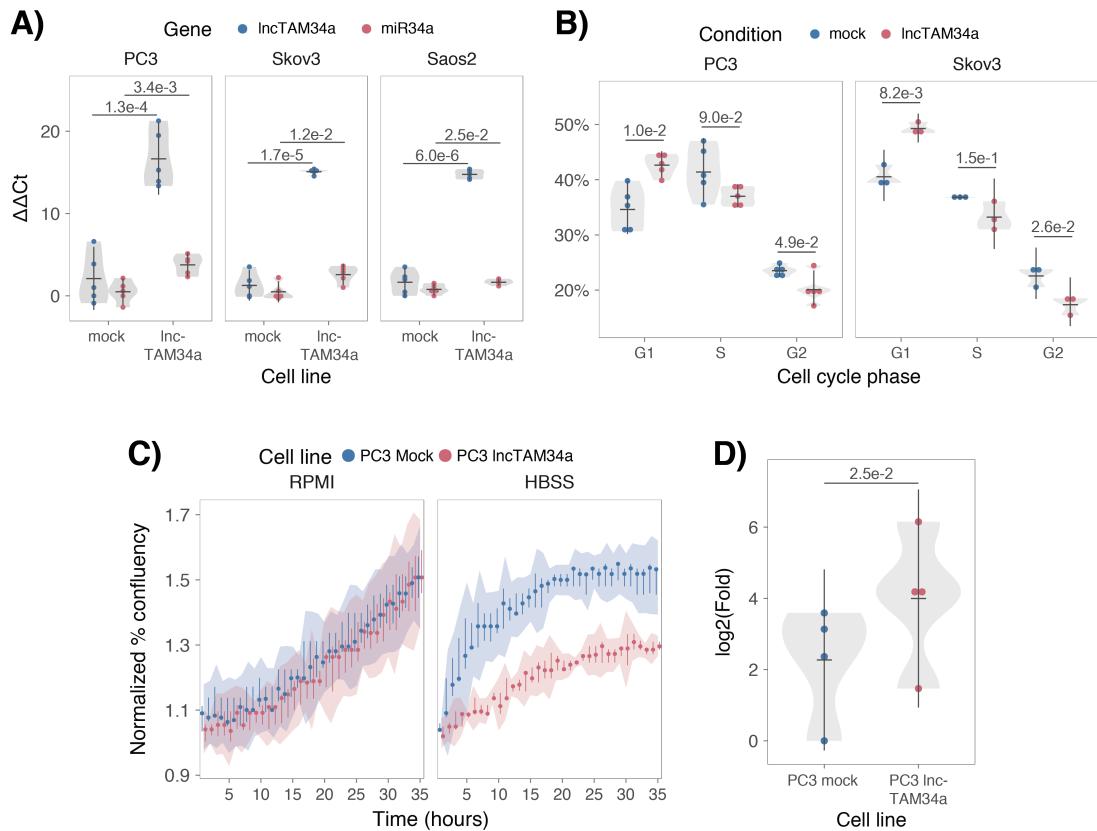
267 levels of *miR34a* HG and that the transcriptional product of the *lncTAM34a*
268 within the p1 construct contributes to inducing a *miR34a* response. To further
269 support these conclusions and better understand the role of *lncTAM34a*
270 during TP53 activation, *TP53^{+/+}* HCT116 cells were co-transfected with p1
271 and shRNA renilla (2.1) and subsequently treated with increasing doses of
272 doxorubicin. Again, the results showed a concomitant reduction in luciferase
273 levels upon knock-down of p1-renilla i.e. the *lncTAM34a* corresponding
274 segment of the p1 transcript (**Fig. 2d**). Furthermore, the results showed that in
275 the absence of p1-renilla the expected induction of p1-luciferase in response
276 to TP53 activation by DNA damage is abrogated. Collectively these results
277 indicate that *lncTAM34a* positively regulates *miR34a* expression and
278 furthermore, suggests that it is crucial for an appropriate TP53-
279 mediated *miR34a* response to DNA damage.

280

281 ***lncTAM34a* can regulate *miR34a* host gene independently of TP53**

282 Despite the fact that TP53 regulates *miR34a* HG and asRNA expression, our
283 results showed that other factors are also able to regulate this locus (**Fig. 2b**).
284 Utilizing a lentiviral system, we stably over-expressed the *lncTAM34a*
285 transcript in three *TP53*-null cell lines, PC3 (prostate cancer), Saos2
286 (osteogenic sarcoma), and Skov3 (ovarian adenocarcinoma). We first
287 analyzed the levels of *lncTAM34a* in these stable cell lines, compared to
288 HEK293T cells, which have high endogenous levels of *lncTAM34a*. On
289 average, the over-expression was approximately 30-fold higher in the over-
290 expression cell lines than in HEK293T cells, roughly corresponding to
291 physiologically relevant levels in cells encountering a stress stimulus, such as

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



298

299 **Figure 3: *IncTAM34a* positively regulates *miR34a* and its associated phenotypes.** A) QPCR-
300 mediated quantification of *miR34a* expression in cell lines stably over-expressing *IncTAM34a*.* B)
301 Cell cycle analysis comparing stably over-expressing *IncTAM34a* cell lines to the respective mock
302 control.* C) Analysis of cellular growth over time in *IncTAM34a* over-expressing PC3 cells. Points
303 represent the median from 3 independent experiments, the colored shadows indicate the 95%
304 confidence interval, and vertical lines show the minimum and maximum values obtained from the three
305 experiments. D) Differential phosphorylated polymerase II binding in *IncTAM34a* over-expressing PC3
306 cells.* *Individual points represent results from independent experiments and the gray shadow
307 indicates the density of those points. Error bars show the 95% CI, black horizontal lines represent the
308 mean, and p-values are shown over long horizontal lines indicating the comparison tested.
309

310 *miR34a* has been previously shown to regulate cell cycle progression, with
311 *miR34a* induction causing G1 arrest (Raver-Shapira et al. 2007, Tarasov et al.
312 2007). Cell cycle analysis via determination of DNA content showed a
313 significant increase in G1 phase cells and a concomitant decrease in G2
314 phase cells in the PC3 and Skov3 *lncTAM34a* over-expressing cell lines,
315 indicating G1 arrest (**Fig. 3b**). The effects of *miR34a* on the cell cycle are
316 mediated by its ability to target cell cycle regulators such as cyclin D1
317 (*CCND1*) (Sun et al. 2008). Quantification of both *CCND1* RNA expression
318 (**Figure 3-Figure Supplement 2a**) and protein levels (**Figure 3-Figure**
319 **Supplement 2b**) in the PC3 *lncTAM34a* over-expressing cell line showed a
320 significant decrease of *CCND1* levels compared to the mock control.
321 Collectively, these results indicate that *lncTAM34a*-mediated induction of
322 *miR34a* is sufficient to result in the corresponding *miR34a*-directed effects on
323 cell cycle.

324

325 *miR34a* is also a well-known inhibitor of cellular growth via its ability to
326 negatively regulate growth factor signaling. Furthermore, starvation has been
327 shown to induce *miR34a* expression causing down-regulation of numerous
328 pro-survival growth factors (Lal et al. 2011). We further interrogated the
329 effects of *lncTAM34a* over-expression by monitoring the growth of the PC3
330 stable cell lines in both normal and starvation conditions via confluency
331 measurements over a 35-hour period. Under normal growth conditions there
332 is a small but significant reduction ($P = 3.0\text{e-}8$; linear regression, **Fig. 3c**) in
333 confluency in the *lncTAM34a* over-expressing cell lines compared to mock
334 control. However, these effects on cell growth are drastically increased in

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

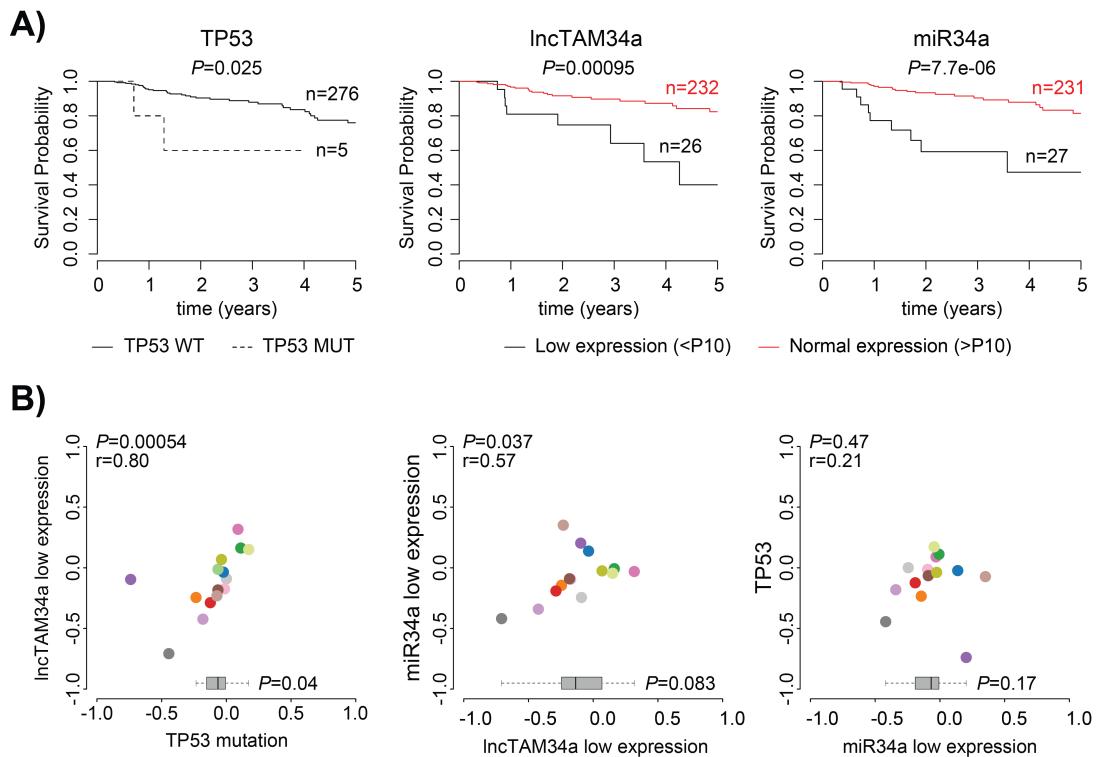
342

343 ***lncTAM34a* transcriptionally activates miR34a host gene**

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

356

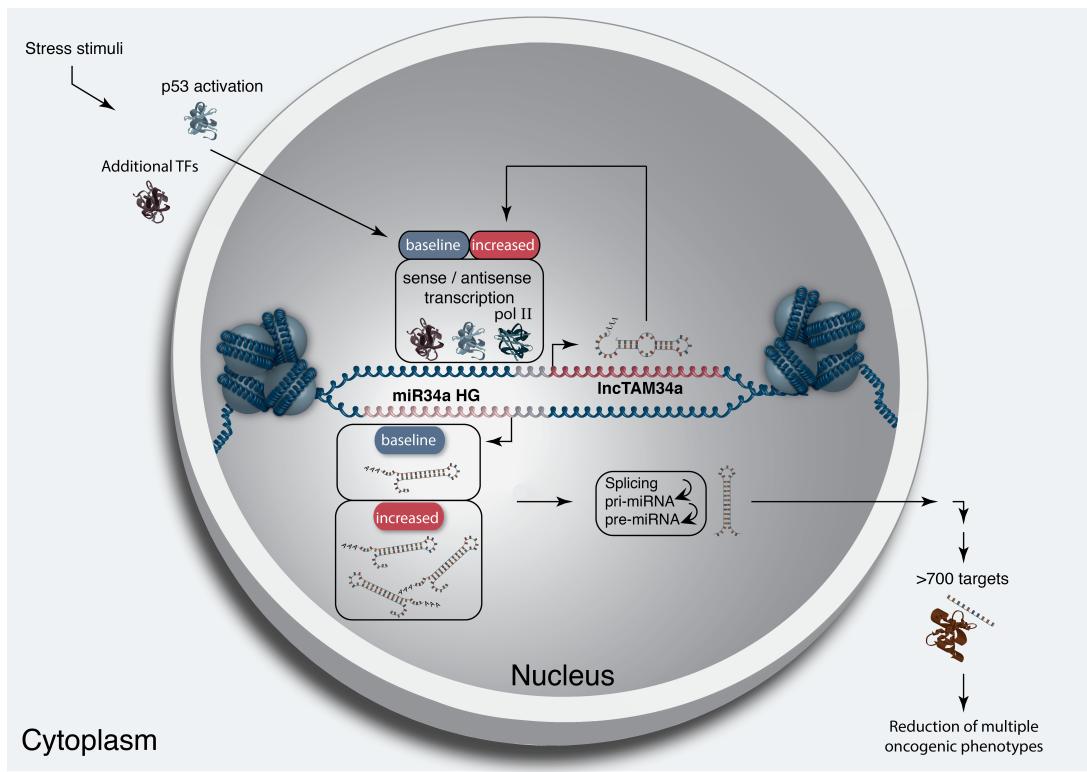
357



358 **Figure 4: Survival analysis.** A) Kaplan-Meier survival curves comparing the effects of *TP53*-mutated
359 samples (left), low *lncTAM34a* expression (middle) and low *miR34a* expression (right) to control
360 samples in papillary kidney cancer (results for other cancers in Figure 4-Supplement 1). B) Correlation
361 analysis between the effects on the 5-year survival probability of *TP53*-mutated samples, low
362 *lncTAM34a* expression, and low *miR34a* expression as indicated. For each variable the 5-year survival
363 probability was compared to the control group (negative value indicates lower survival, positive value
364 indicates higher survival). Spearman correlation coefficients are given on top left of each plot. Each dot
365 indicates one cancer type (see Fig.1c for legend). Boxplots on the bottom summarize the effects for the
366 parameter on the x-axis, with indication of p-values, as calculated using paired Wilcoxon signed rank
367 test. Low expression was defined as *TP53* non-mutated samples having expression values in the bottom
368 10th percentile.

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

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



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 increases *miR34a* HG transcription levels resulting in enrichment of polymerase II at the *miR34a* promoter and a positive feed-forward loop. *IncTAM34a*-mediated increases in *miR34a* HG potentially occur via direct DNA binding alone, by direct DNA binding and recruitment of additional factors, or through a yet unknown mechanism. *miR34a* HG then, in turn, is spliced and processed before being exported to the cytoplasm. The *miR34a* pre-miRNA then undergoes further processing before the mature *miR34a* binds to the RISC complex allowing it to bind and repress its targets and exert its tumor suppressive effects.

402 **Discussion**

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

422

423 In agreement with previous studies, we demonstrate that upon encountering
424 various types of cellular stress, *TP53* in concert with additional factors bind
425 and initiate transcription at the *miR34a* locus, thus increasing the levels of
426 *lncTAM34a* and, in addition, *miR34a*. We found that overexpression of
427 *lncTAM34a* leads to recruitment of polymerase II to the *miR34a* promoter and

428 hypothesize that *lncTAM34a* may provide positive feedback for *miR34a*
429 expression whereby it serves as a scaffold for the recruitment of additional
430 factors that facilitate polymerase II-mediated transcription. In this manner,
431 *miR34a* expression is induced and thus, drives a shift towards a reduction in
432 growth factor signaling, senescence, and in some cases apoptosis. On the
433 other hand, in cells without functional TP53, other factors, which typically act
434 independently or in concert with TP53, may initiate transcription of the *miR34a*
435 locus. In this scenario *lncTAM34a* could potentially be interacting directly with
436 one of these additional factors and recruiting it to the *miR34a* locus in order to
437 drive *miR34a* transcription. The head-to-head orientation of the *miR34a* HG
438 and *lncTAM34a* causes sequence complementarity between the RNA and the
439 promoter DNA, making this an attractive mechanism. Previous reports have
440 also illustrated the ability of asRNAs to form hybrid DNA:RNA R-loops and,
441 thus, facilitate an open chromatin structure and the transcription of the sense
442 gene (Boque-Sastre et al. 2015). The fact that the p1 construct only contains
443 a small portion (307 bp) of the *lncTAM34a* transcript indicates that this portion
444 is sufficient to give rise to at least a partial *miR34a* inducing response and
445 therefore, that *lncTAM34a* may be able to facilitate *miR34a* expression
446 independent of additional factors (**Fig 2d, Figure 2-Figure Supplement 2a**).
447 Nevertheless, further work will need to be performed to explore the
448 mechanism whereby *lncTAM34a* regulates *miR34a* gene expression.

449
450 An antisense transcript arising from the *miR34a* locus, *Lnc34a*, has been
451 previously reported to negatively regulate the expression of *miR34a* (Wang et
452 al. 2016). Although the *Lnc34a* and *lncTAM34a* transcripts share some

453 sequence similarity, we believe them to be separate RNAs that are,
454 potentially, different isoforms of the same gene. We utilized CAGE and
455 RNAseq data from the ENCODE project to evaluate the presence of
456 *IncTAM34a* and *Lnc34a* in 28 and 36 commonly used cancer cell lines,
457 respectively. Although the results show the presence of *IncTAM34a* in these
458 cell lines, we find no evidence for *Lnc34a* transcription (**Supplementary**
459 **Document 1**). These results are in line with the findings of Wang et al.
460 indicating that *Lnc34a* is highly expressed in colon cancer stem cell spheres
461 compared to other cell types used in their study and, furthermore suggests,
462 that these two transcripts are not commonly co-expressed. The fact
463 that *IncTAM34a* and *Lnc34a* would appear to have opposing roles in their
464 regulation of *miR34a*, further underlines the complexity of the regulation at
465 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 *IncTAM34a* expression making it an attractive candidate for controlled
476 preclinical studies. Due to *IncTAM34a*-mediated positive feedback on *miR34a*
477 expression, initiation of this feedback mechanism may be able to 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 player in the regulation of *miR34a* and its particular importance in typical
481 examples of cellular stress encountered in cancer. The conclusions drawn in
482 this study provide essential insight regarding asRNA-mediated regulation of
483 cancer-associated miRNAs and, contribute to fundamental knowledge
484 concerning *miR34a* regulation necessary for its efficient induction in clinical
485 settings.

486

487 **Materials and Methods**

488 **Cell Culture**

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

501 502 **Bioinformatics, Data Availability, and Statistical Testing**

503 The USCS genome browser (Kent et al. 2002) was utilized for the

504 bioinformatic evaluation of antisense transcription utilizing the RefSeq
505 (O'Leary et al. 2016) gene annotation track.

506

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

518

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

522

523 **Coding Potential**

524 Protein-coding capacity was evaluated using the Coding-potential
525 Assessment Tool (Wang et al. 2013) and Coding-potential Calculator (Kong et
526 al. 2007) with default settings. Transcript sequences for use with Coding-
527 potential Assessment Tool were downloaded from the UCSC genome
528 browser using the Ensembl

529 accessions: *HOTAIR* (ENST00000455246), *XIST* (ENST00000429829), β-
530 actin (ENST00000331789), Tubulin (ENST00000427480),
531 and *MYC* (ENST00000377970). Transcript sequences for use with Coding-
532 potential Calculator were downloaded from the UCSC genome browser using
533 the following IDs: *HOTAIR* (uc031qho.1), β-actin (uc003sq.4).

534

535 **Peptide identification in MS/MS spectra**

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

551

552 **shRNAs**

553 shRNA-expressing constructs were cloned into the U6M2 construct using the

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

559

560 **Bi-directional Promoter Cloning**

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

565

566 **Promoter Activity**

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

574

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

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

579 cloned into the Nhe1 and Pac1 restriction sites in the pHIV7-IMPDH2-U6
580 plasmid.

581

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

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

593

594 Stable lines were generated by infecting cells with a multiplicity of infection of
595 1 after which 1-2 µM mycophenolic acid (Merck, Kenilworth, NJ, USA. Cat#
596 M5255) selection was initiated 48-72 hours post-infection. Cells were
597 expanded as the selection process was monitored via flow cytometry analysis
598 (LSRII, BD Biosciences) of GFP and selection was terminated once > 90% of
599 the cells were GFP positive. Quantification of *IncTAM34a* over-expression and
600 *miR34a* was performed in biological quintuplet for all cell 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 $\Delta\Delta Ct$ 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 a two sample t-
663 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 log 2 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 Confluency was measured every hour using bright-field microscopy and the
706 percentage of confluence was reported via the plate reader's inbuilt algorithm.
707 Percentage of confluence was normalized to the control sample in each
708 condition (shown in figure) and then ranked to move the data to a linear scale.
709 Using the mean of the technical duplicates in three biological replicates, the
710 rank was then used to construct a linear model, of the dependency of the rank
711 on the time and cell lines variables for each growth condition. Reported p-
712 values are derived from the t-test, testing the null hypothesis that the
713 coefficient estimate of the cell line variable is equal to 0.

714

715 **Pharmacological Compounds**

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

717

718 **Cellular Localization Analysis**

719 Quantified RNAseq data from 11 cell lines from the GRCh38 assembly was
720 downloaded from the ENCODE project database and quantifications for
721 *IncTAM34a* (ENSG00000234546), GAPDH (ENSG00000111640), and
722 MALAT1 (ENSG00000251562) were extracted. Cell lines for which data was
723 downloaded include: A549, GM12878, HeLa-S3, HepG2, HT1080, K562
724 MCF-7, NCI-H460, SK-MEL-5, SK-N-DZ, SK-N-SH. Initial exploratory analysis
725 revealed that several cell lines should be removed from the analysis due to a)
726 a larger proportion of GAPDH in the nucleus than cytoplasm or b) variation of
727 *IncTAM34a* expression is too large to draw conclusions, or c) they have no or

728 low (<6 TPM) *lncTAM34a* expression. Furthermore, only polyadenylated
729 libraries were used in the final analysis, due to the fact that the cellular
730 compartment enrichment was improved in these samples. All analyzed genes
731 are reported to be polyadenylated. In addition, only samples with 2 biological
732 replicates were retained. For each cell type, gene, and biological replicate the
733 fraction of transcripts per million (TPM) in each cellular compartment was
734 calculated as the fraction of TPM in the specific compartment by the total
735 TPM. The mean and standard deviation for the fraction was subsequently
736 calculated for each cell type and cellular compartment and this information
737 was represented in the final figure.

738

739 **CAGE Analysis**

740 All available CAGE data from the ENCODE project (Consortium 2012) for 36
741 cell lines was downloaded from the UCSC genome browser (Kent et al. 2002)
742 for genome version hg19. Of these, 28 cell lines had CAGE transcription start
743 sites (TSS) mapping to the plus strand of chromosome 1 and in regions
744 corresponding to 200 base pairs upstream of the *lnc34a* start site (9241796 -
745 200) and 200 base pairs upstream of the GENCODE annotated *lncTAM34a*
746 start site (9242263 + 200). These cell lines included: HFDPC, H1-hESC,
747 HMEpC, HAoEC, HPIEpC, HSaVEC, GM12878, hMSC-BM, HUVEC,
748 AG04450, hMSC-UC, IMR90, NHDF, SK-N-SH_RA, BJ, HOB, HPC-PL,
749 HAoAF, NHEK, HVMF, HWP, MCF-7, HepG2, hMSC-AT, NHEM.f_M2,
750 SkMC, NHEM_M2, and HCH. In total 74 samples were included. 17 samples
751 were polyA-, 47 samples were polyA+, and 10 samples were total RNA. In
752 addition, 34 samples were whole cell, 15 enriched for the cytosolic fraction, 15

753 enriched for the nucleolus, and 15 enriched for the nucleus. All CAGE
754 transcription start sites were plotted and the RPKM of the individual reads was
755 used to color each read to indicate their relative abundance. In cases where
756 CAGE TSS spanned identical regions, the RPMKs of the regions were
757 summed and represented as one CAGE TSS in the figure. In addition, a
758 density plot shows the distribution of the CAGE reads in the specified
759 interval.

760

761 **Splice Junction Analysis**

762 All available whole cell (i.e. non-fractionated) spliced read data originating
763 from the Cold Spring Harbor Lab in the ENCODE project (Consortium 2012)
764 for 38 cell lines was downloaded from the UCSC genome browser (Kent et al.
765 2002). Of these cell lines, 36 had spliced reads mapping to the plus strand of
766 chromosome 1 and in the region between the *lnc34a* start (9241796) and
767 transcription termination (9257102) site (note that *lncTAM34a* resides totally
768 within this region). Splice junctions from the following cell lines were included
769 in the final figure: A549, Ag04450, Bj, CD20, CD34 mobilized, Gm12878,
770 H1hesc, Haoaf, Haoec, Hch, Helas3, Hepg2, Hfdpc, Hmec, Hmepc, Hmscat,
771 Hmscbm, Hmscuc, Hob, Hpcpl, Hpiepc, Hsavec, Hsomm, Huvec, Hvmaf, Hwp,
772 Imr90, Mcf7, Monocd14, Nhdf, Nhek, Nhemfm2, Nhemm2, Nhlf, Skmc, and
773 Sknsh. All splice junctions were included in the figure and colored according
774 to the number of reads corresponding to each. In cases where identical reads
775 were detected multiple times, the read count was summed and represented
776 as one read in the figure.

777

778 **TCGA Data Analysis**

779 RNA-Seq data and copy number data were downloaded from TCGA and
780 processed as described previously (Ashouri et al. 2016). Briefly, RNA-Seq
781 data were aligned to the human hg19 assembly and quantified using
782 GENCODE (v19) annotated HTSeq-counts and FPKM normalizations.
783 Expression data from *miR34a* and *lncTAM34a* (identified as RP3-510D11.2)
784 were used for further analysis. Copy number amplitudes for GENCODE genes
785 were determined from segmented copy-number data. Samples that were
786 diploid for *lncTAM34a* were identified as those samples that had copy number
787 amplitudes between -0.1 and 0.1.

788

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

792

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

796

797 **Acknowledgments**

798 The authors would like to kindly thank Martin Enge for his critical review of the
799 manuscript and fruitful discussions.

800

801 **Competing Interests**

802
803 The authors declare no competing interests.
804

805

806 **Funding**

807
808 This work has been supported by the Swedish Research Council [521-2012-
809 2037], Swedish Cancer Society [150768], Cancer Research Foundations of
810 Radiumhemmet [144063] and the Swedish Childhood Cancer Foundation
811 [PR2015-0009].

812
813

814 **Figure Supplements**

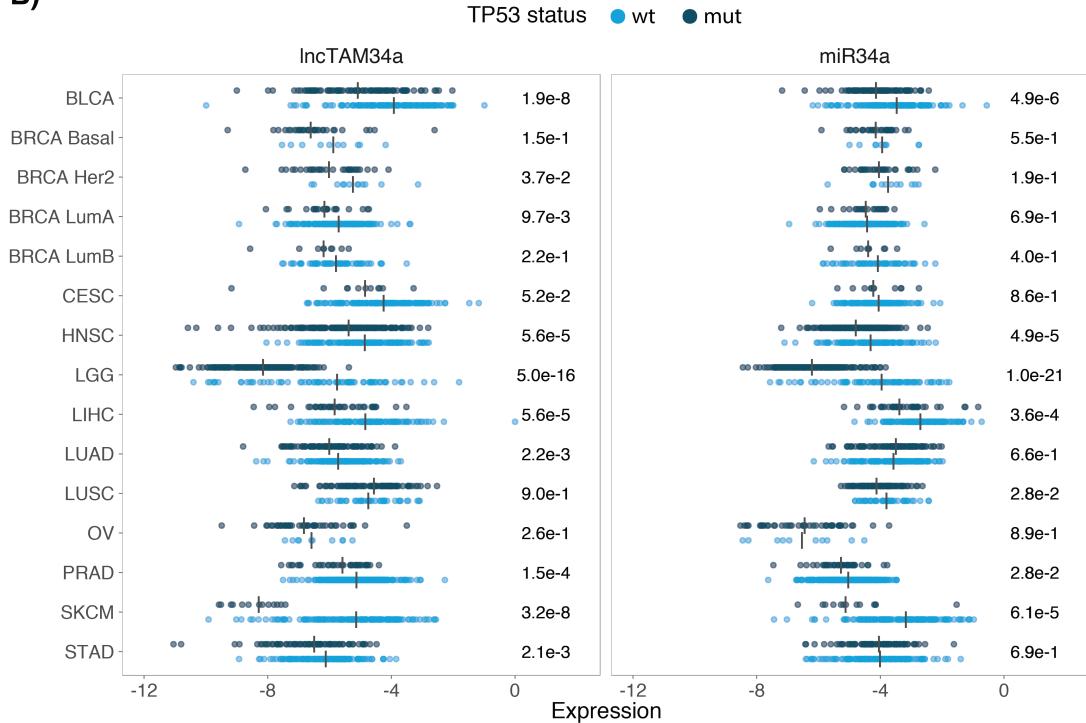
815
816 Figure 1-Supplement 1: TCAG expression levels and correlation analysis
817 statistics.
818
819 Figure 1-Supplement 2: Molecular characteristics of *IncTAM34a*.
820
821 Figure 2-Supplement 1: A schematic representation of the p1 construct.
822
823 Figure 2-Supplement 2: Evaluating the effects of *IncTAM34a* down-regulation.
824
825 Figure 3-Supplement 1: Physiological relevance of *IncTAM34a*
826 overexpression.
827
828 Figure 3-Supplement 2: Effects of *IncTAM34a* overexpression on cyclin D1.
829
830 Figure 4-Supplement 1: Survival analysis in 17 cancers from TCGA.
831
832 Supplementary Document 1: Evaluating the relationship between *IncTAM34a*
833 and *Inc34a*.
834
835 Supplementary Document 2: A table of primers used in this study.

836 **Supplementary Figures**

A)

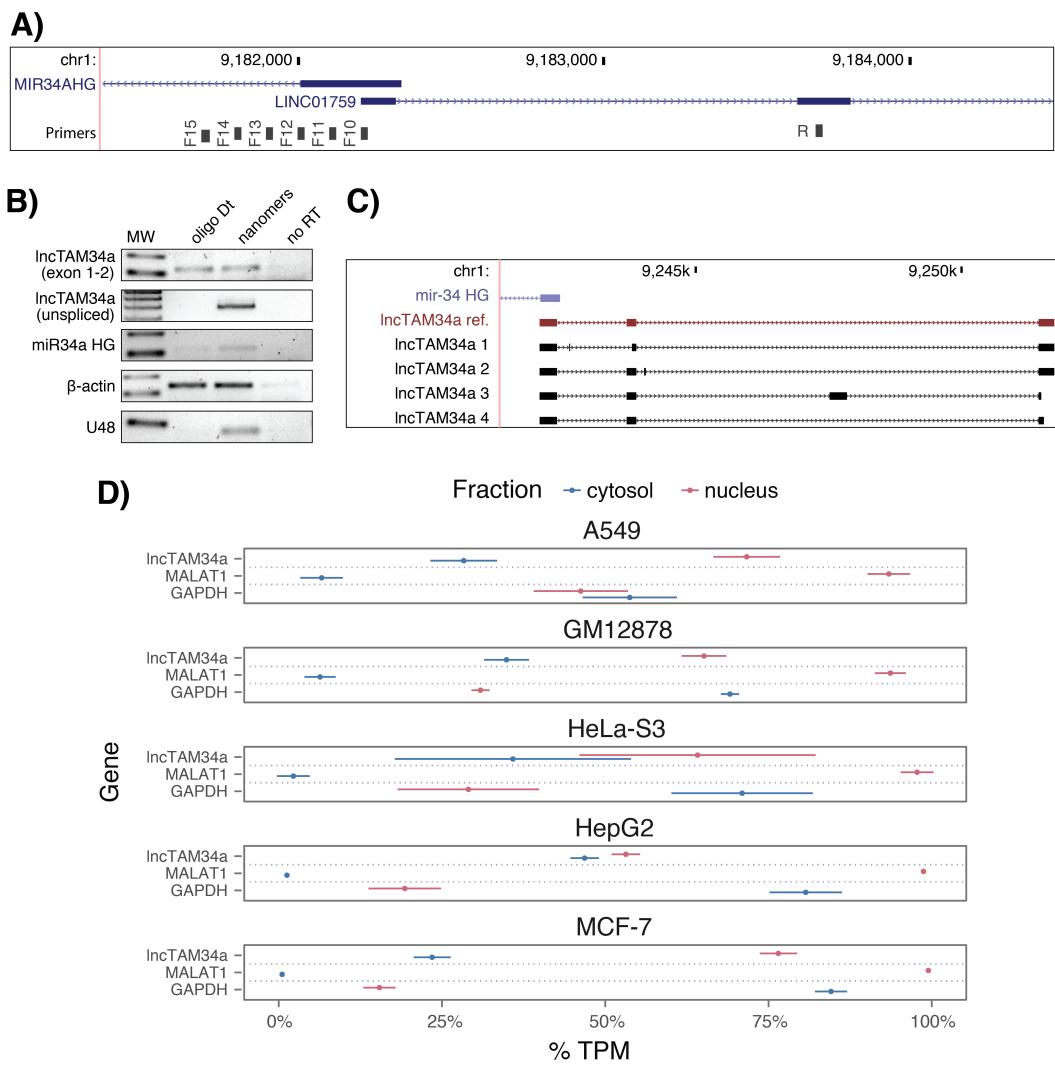
cancer	all n	all rho	all p	TP53wt n	TP53wt rho	TP53wt p	TP53mut n	TP53mut rho	TP53mut p
Adrenocortical carcinoma (ACC)	10	0.55	1.04e-01	10	0.55	1.04e-01	NA	NA	NA
Bladder Urothelial Carcinoma (BLCA)	228	0.51	7.89e-17	134	0.45	3.86e-08	94	0.43	1.73e-05
Breast invasive carcinoma (BRCA) Basal	42	0.57	9.54e-05	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)



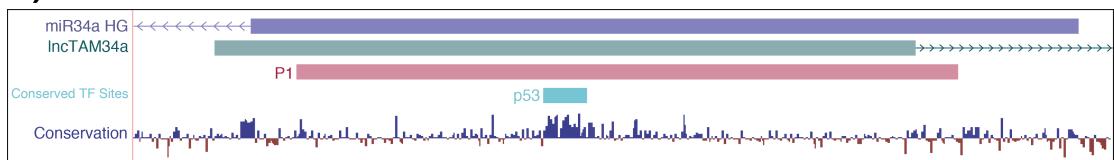
837
838

Figure 1 Supplement 1: TCGA normalized expression levels and correlation analysis statistics.
A) Spearman's rho and p-values (p) from the correlation analysis in Figure 1a between miR34a and IncTAM34a expression in TP53 wild type (wt) and mutated (mut) samples within TCGA cancer types. NA indicates not applicable, due to a lack of data for the specific group. **B)** Expression levels of *miR34a* and *IncTAM34a* in *TP53* wt and nonsynonymous mutation samples. Expression was quantified by the log2 ratio of expression of the gene to its maximal expression value. Vertical lines indicate the median. P-values are indicated on the right side of each panel and are derived from comparing the *TP53* wild type samples to the samples with a nonsynonymous mutation using a two-sided Wilcoxon signed rank test. Only samples that had at least 5 samples per comparison were included. In addition, only samples that were diploid at the *miR34a* locus and were used for the analysis to avoid copy number bias.

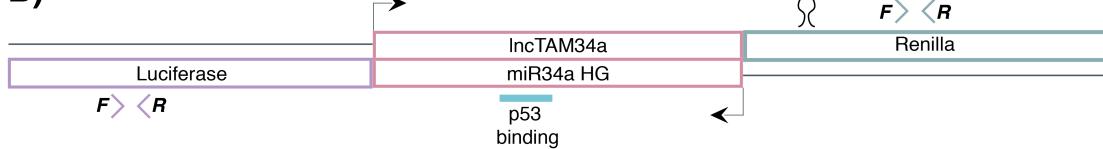


850
851
852 **Figure 1 Supplement 2: Molecular characteristics of lncTAM34a.** **A)** A schematic representation of
853 the primer placement in the primer walk assay. **B)** Polyadenylation status of spliced and
854 unspliced lncTAM34a in HEK293T cells. **C)** Sequencing results from the analysis of lncTAM34a
855 isoforms in U2OS cells. lncTAM34a ref. refers to the full-length transcript as defined by the 3'-RACE
856 and primer walk assay. **D)** Analysis of coding potential of the lncTAM34a transcript using the Coding-
857 potential Calculator. **E)** RNAseq data from five fractionated cell lines in the ENCODE project showing
858 the percentage of transcripts per million (TPM) for lncTAM34a. MALAT1 (nuclear localization) and
859 GAPDH (cytoplasmic localization) are included as fractionation controls. Points represent the mean
860 and horizontal lines represent the standard deviation from two biological replicates.
861

A)



B)



862
863
864
865
866
867

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

868
869
870
871
872
873
874

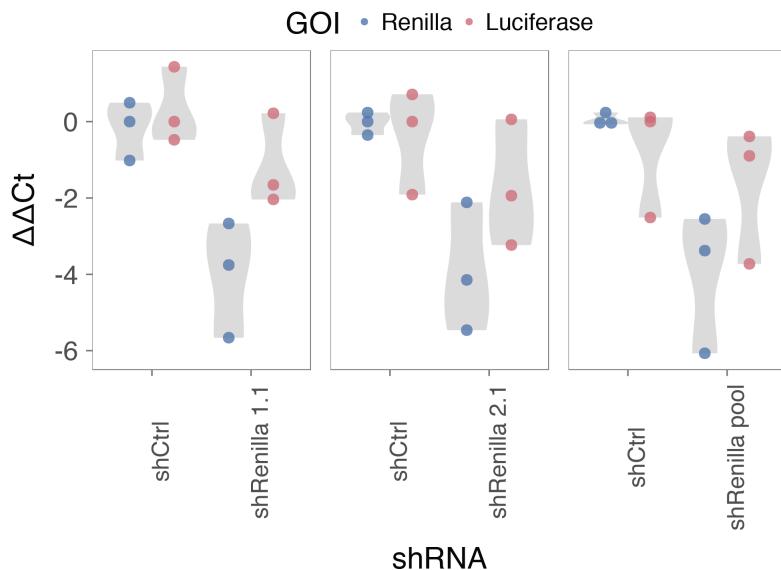
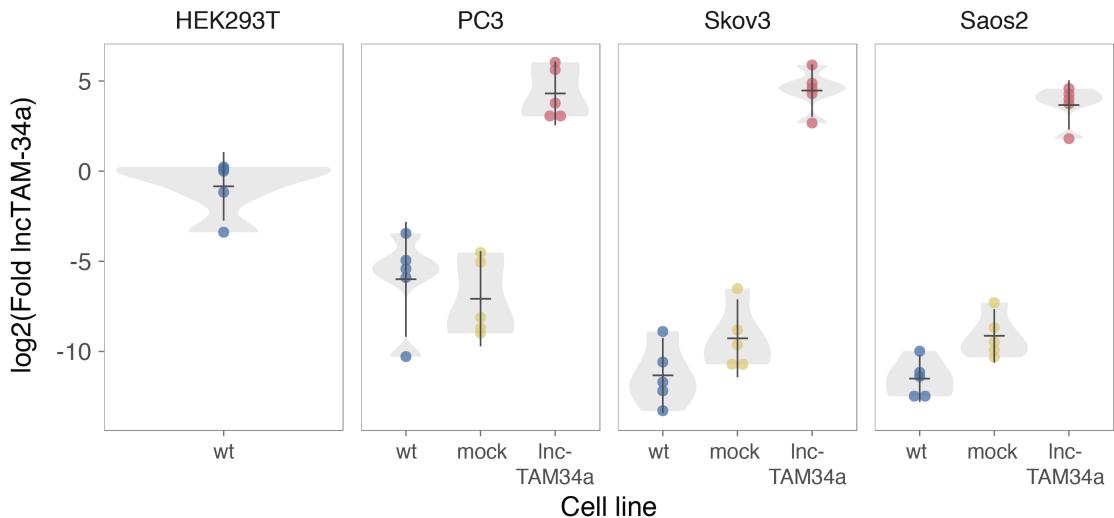


Figure 2 Supplement 2: Evaluating the effects of lncTAM34a down-regulation. HEK293T cells were co-transfected with the p1 construct and either shRenilla or shControl. Renilla and luciferase levels were measured with Q-PCR 48 hours after transfection. Individual points represent independent experiments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



875
876
877
878
879

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

880
881
882
883
884

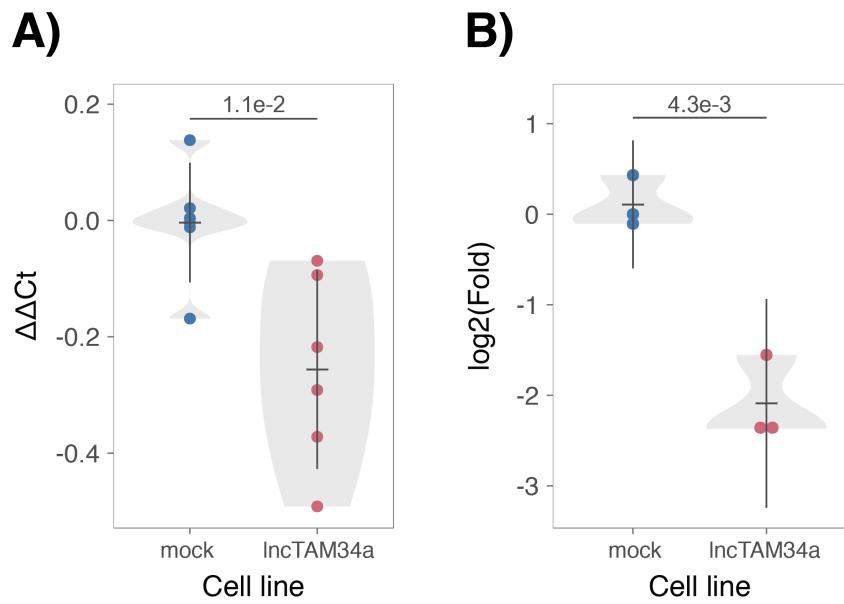
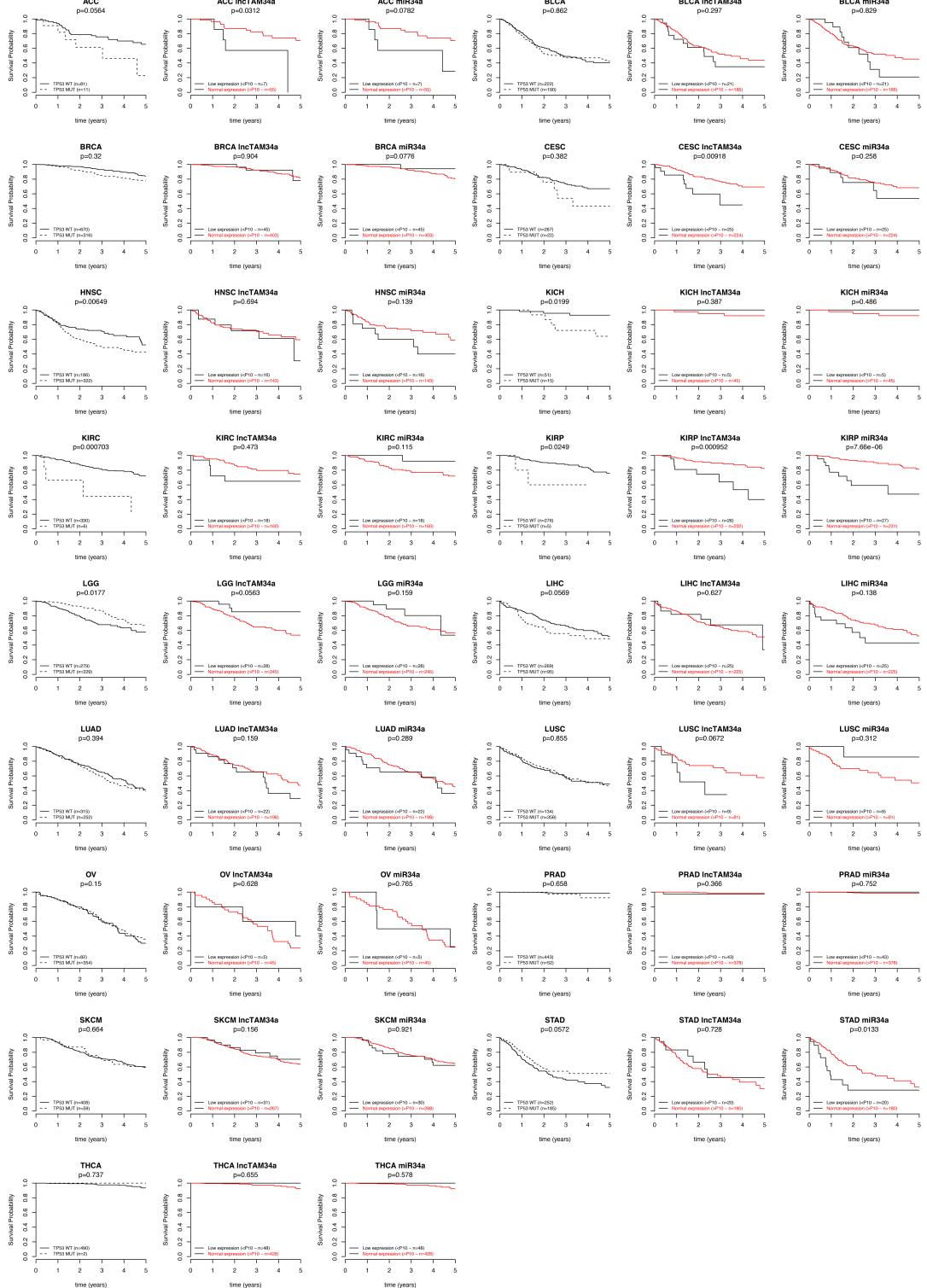


Figure 3 Supplement 2: Effects of *lncTAM34a* overexpression on cyclin D1. CCND1 expression (A) and western blot quantification of protein levels (B) in *lncTAM34a* over-expressing PC3 stable cell lines. Experiments were performed in biological sextuplets (A) or triplicates (B).



885

Figure 4-Supplement 1: Survival analysis in 17 cancers from TCGA. Kaplan-Meier survival curves comparing the effects of *TP53*-mutated samples (left), low *lncTAM34a* 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.

886

887

888

889

890

891

892 References

893

- 894 Agostini, M., P. Tucci, R. Killick, E. Candi, B. S. Sayan, P. Rivetti di Val Cervo, P.
895 Nicotera, F. McKeon, R. A. Knight, T. W. Mak and G. Melino (2011). "Neuronal
896 differentiation by TAp73 is mediated by microRNA-34a regulation of synaptic
897 protein targets." *Proc Natl Acad Sci U S A* **108**(52): 21093-21098. DOI:
898 10.1073/pnas.1112061109
899
- 900 Ahn, Y. H., D. L. Gibbons, D. Chakravarti, C. J. Creighton, Z. H. Rizvi, H. P. Adams, A.
901 Pertsemlidis, P. A. Gregory, J. A. Wright, G. J. Goodall, E. R. Flores and J. M. Kurie
902 (2012). "ZEB1 drives prometastatic actin cytoskeletal remodeling by
903 downregulating miR-34a expression." *J Clin Invest* **122**(9): 3170-3183. DOI:
904 10.1172/JCI63608
905
- 906 Allaire, J., Y. Xie, J. McPherson, J. Luraschi, K. Ushey, A. Atkins, H. Wickham, J.
907 Cheng and W. Chang (2017). rmarkdown: Dynamic Documents for R. R package
908 version 1.8. <https://CRAN.R-project.org/package=rmarkdown>
909
- 910 Amarzguioui, M., J. J. Rossi and D. Kim (2005). "Approaches for chemically
911 synthesized siRNA and vector-mediated RNAi." *FEBS Lett* **579**(26): 5974-5981.
912 DOI: 10.1016/j.febslet.2005.08.070
913
- 914 Arnold, J. B. (2017). ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'. R
915 package version 3.4.0. <https://CRAN.R-project.org/package=ggthemes>
916
- 917 Ashouri, A., V. I. Sayin, J. Van den Eynden, S. X. Singh, T. Papagiannakopoulos and
918 E. Larsson (2016). "Pan-cancer transcriptomic analysis associates long non-
919 coding RNAs with key mutational driver events." *Nature Communications* **7**:
920 13197. DOI: 10.1038/ncomms13197
921
- 922 Balbin, O. A., R. Malik, S. M. Dhanasekaran, J. R. Prensner, X. Cao, Y. M. Wu, D.
923 Robinson, R. Wang, G. Chen, D. G. Beer, A. I. Nesvizhskii and A. M. Chinnaiyan
924 (2015). "The landscape of antisense gene expression in human cancers." *Genome
925 Res* **25**(7): 1068-1079. DOI: 10.1101/gr.180596.114
926
- 927 Boekel, J., J. M. Chilton, I. R. Cooke, P. L. Horvatovich, P. D. Jagtap, L. Kall, J. Lehtio,
928 P. Lukasse, P. D. Moerland and T. J. Griffin (2015). "Multi-omic data analysis using
929 Galaxy." *Nat Biotechnol* **33**(2): 137-139. DOI: 10.1038/nbt.3134
930
- 931 Boque-Sastre, R., M. Soler, C. Oliveira-Mateos, A. Portela, C. Moutinho, S. Sayols, A.
932 Villanueva, M. Esteller and S. Guil (2015). "Head-to-head antisense transcription
933 and R-loop formation promotes transcriptional activation." *Proc Natl Acad Sci U
934 S A* **112**(18): 5785-5790. DOI: 10.1073/pnas.1421197112
935
- 936 Chang, T. C., D. Yu, Y. S. Lee, E. A. Wentzel, D. E. Arking, K. M. West, C. V. Dang, A.
937 Thomas-Tikhonenko and J. T. Mendell (2008). "Widespread microRNA
938 repression by Myc contributes to tumorigenesis." *Nat Genet* **40**(1): 43-50. DOI:
939 10.1038/ng.2007.30

- 940
941 Chang, W. (2014). extrafont: Tools for using fonts. R package version 0.17.
942 <https://CRAN.R-project.org/package=extrafont>
943
944 Chen, J., M. Sun, W. J. Kent, X. Huang, H. Xie, W. Wang, G. Zhou, R. Z. Shi and J. D.
945 Rowley (2004). "Over 20% of human transcripts might form sense-antisense
946 pairs." *Nucleic Acids Res* **32**(16): 4812-4820. DOI: 10.1093/nar/gkh818
947
948 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.
949 S. Zheng (2010). "The impact of miR-34a on protein output in hepatocellular
950 carcinoma HepG2 cells." *Proteomics* **10**(8): 1557-1572. DOI:
951 10.1002/pmic.200900646
952
953 Chim, C. S., K. Y. Wong, Y. Qi, F. Loong, W. L. Lam, L. G. Wong, D. Y. Jin, J. F. Costello
954 and R. Liang (2010). "Epigenetic inactivation of the miR-34a in hematological
955 malignancies." *Carcinogenesis* **31**(4): 745-750. DOI: 10.1093/carcin/bgq033
956
957 Cole, K. A., E. F. Attiyeh, Y. P. Mosse, M. J. Laquaglia, S. J. Diskin, G. M. Brodeur and
958 J. M. Maris (2008). "A functional screen identifies miR-34a as a candidate
959 neuroblastoma tumor suppressor gene." *Mol Cancer Res* **6**(5): 735-742. DOI:
960 10.1158/1541-7786.MCR-07-2102
961
962 Conley, A. B. and I. K. Jordan (2012). "Epigenetic regulation of human cis-natural
963 antisense transcripts." *Nucleic Acids Res* **40**(4): 1438-1445. DOI:
964 10.1093/nar/gkr1010
965
966 Consortium, E. P. (2012). "An integrated encyclopedia of DNA elements in the
967 human genome." *Nature* **489**(7414): 57-74. DOI: 10.1038/nature11247
968
969 Ding, N., H. Wu, T. Tao and E. Peng (2017). "NEAT1 regulates cell proliferation
970 and apoptosis of ovarian cancer by miR-34a-5p/BCL2." *Onco Targets Ther* **10**:
971 4905-4915. DOI: 10.2147/OTT.S142446
972
973 Djebali, S., C. A. Davis, A. Merkel, A. Dobin, T. Lassmann, A. Mortazavi, A. Tanzer, J.
974 Lagarde, W. Lin, F. Schlesinger, C. Xue, G. K. Marinov, J. Khatun, B. A. Williams, C.
975 Zaleski, J. Rozowsky, M. Roder, F. Kokocinski, R. F. Abdelhamid, T. Alioto, I.
976 Antoshechkin, M. T. Baer, N. S. Bar, P. Batut, K. Bell, I. Bell, S. Chakrabortty, X.
977 Chen, J. Chrast, J. Curado, T. Derrien, J. Drenkow, E. Dumais, J. Dumais, R.
978 Duttagupta, E. Falconnet, M. Fastuca, K. Fejes-Toth, P. Ferreira, S. Foissac, M. J.
979 Fullwood, H. Gao, D. Gonzalez, A. Gordon, H. Gunawardena, C. Howald, S. Jha, R.
980 Johnson, P. Kapranov, B. King, C. Kingswood, O. J. Luo, E. Park, K. Persaud, J. B.
981 Preall, P. Ribeca, B. Risk, D. Robyr, M. Sammeth, L. Schaffer, L. H. See, A. Shahab, J.
982 Skancke, A. M. Suzuki, H. Takahashi, H. Tilgner, D. Trout, N. Walters, H. Wang, J.
983 Wrobel, Y. Yu, X. Ruan, Y. Hayashizaki, J. Harrow, M. Gerstein, T. Hubbard, A.
984 Reymond, S. E. Antonarakis, G. Hannon, M. C. Giddings, Y. Ruan, B. Wold, P.
985 Carninci, R. Guigo and T. R. Gingeras (2012). "Landscape of transcription in
986 human cells." *Nature* **489**(7414): 101-108. DOI: 10.1038/nature11233
987

- 988 Gallardo, E., A. Navarro, N. Vinolas, R. M. Marrades, T. Diaz, B. Gel, A. Quera, E.
989 Bandres, J. Garcia-Foncillas, J. Ramirez and M. Monzo (2009). "miR-34a as a
990 prognostic marker of relapse in surgically resected non-small-cell lung cancer."
991 Carcinogenesis **30**(11): 1903-1909. DOI: 10.1093/carcin/bgp219
992
- 993 Geiger, T., A. Wehner, C. Schaab, J. Cox and M. Mann (2012). "Comparative
994 proteomic analysis of eleven common cell lines reveals ubiquitous but varying
995 expression of most proteins." Mol Cell Proteomics **11**(3): M111 014050. DOI:
996 10.1074/mcp.M111.014050
997
- 998 Granholm, V., S. Kim, J. C. Navarro, E. Sjolund, R. D. Smith and L. Kall (2014). "Fast
999 and accurate database searches with MS-GF+Percolator." J Proteome Res **13**(2):
1000 890-897. DOI: 10.1021/pr400937n
1001
- 1002 Grossman, R. L., A. P. Heath, V. Ferretti, H. E. Varmus, D. R. Lowy, W. A. Kibbe and
1003 L. M. Staudt (2016). "Toward a Shared Vision for Cancer Genomic Data." N Engl J
1004 Med **375**(12): 1109-1112. DOI: 10.1056/NEJMp1607591
1005
- 1006 Holman, J. D., D. L. Tabb and P. Mallick (2014). "Employing ProteoWizard to
1007 Convert Raw Mass Spectrometry Data." Curr Protoc Bioinformatics **46**: 13 24 11-
1008 19. DOI: 10.1002/0471250953.bi1324s46
1009
- 1010 Hunten, S., M. Kaller, F. Drepper, S. Oeljeklaus, T. Bonfert, F. Erhard, A. Dueck, N.
1011 Eichner, C. C. Friedel, G. Meister, R. Zimmer, B. Warscheid and H. Hermeking
1012 (2015). "p53-Regulated Networks of Protein, mRNA, miRNA, and lncRNA
1013 Expression Revealed by Integrated Pulsed Stable Isotope Labeling With Amino
1014 Acids in Cell Culture (pSILAC) and Next Generation Sequencing (NGS) Analyses."
1015 Mol Cell Proteomics **14**(10): 2609-2629. DOI: 10.1074/mcp.M115.050237
1016
- 1017 International Human Genome Sequencing, C. (2004). "Finishing the euchromatic
1018 sequence of the human genome." Nature **431**(7011): 931-945. DOI:
1019 10.1038/nature03001
1020
- 1021 Johnsson, P., A. Ackley, L. Vidarsdottir, W. O. Lui, M. Corcoran, D. Grander and K.
1022 V. Morris (2013). "A pseudogene long-noncoding-RNA network regulates PTEN
1023 transcription and translation in human cells." Nat Struct Mol Biol **20**(4): 440-
1024 446. DOI: 10.1038/nsmb.2516
1025
- 1026 Katayama, S., Y. Tomaru, T. Kasukawa, K. Waki, M. Nakanishi, M. Nakamura, H.
1027 Nishida, C. C. Yap, M. Suzuki, J. Kawai, H. Suzuki, P. Carninci, Y. Hayashizaki, C.
1028 Wells, M. Frith, T. Ravasi, K. C. Pang, J. Hallinan, J. Mattick, D. A. Hume, L. Lipovich,
1029 S. Batalov, P. G. Engstrom, Y. Mizuno, M. A. Faghihi, A. Sandelin, A. M. Chalk, S.
1030 Mottagui-Tabar, Z. Liang, B. Lenhard, C. Wahlestedt, R. G. E. R. Group, G. Genome
1031 Science and F. Consortium (2005). "Antisense transcription in the mammalian
1032 transcriptome." Science **309**(5740): 1564-1566. DOI: 10.1126/science.1112009
1033
- 1034 Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler and D.
1035 Haussler (2002). "The human genome browser at UCSC." Genome Res **12**(6):

1036 996-1006. DOI: 10.1101/gr.229102. Article published online before print in May
1037 2002

1038

1039 Kim, K. H., H. J. Kim and T. R. Lee (2017). "Epidermal long non-coding RNAs are
1040 regulated by ultraviolet irradiation." *Gene* **637**: 196-202. DOI:
1041 10.1016/j.gene.2017.09.043

1042

1043 Kim, S. and P. A. Pevzner (2014). "MS-GF+ makes progress towards a universal
1044 database search tool for proteomics." *Nature Communications* **5**: 5277. DOI:
1045 10.1038/ncomms6277

1046

1047 Kong, L., Y. Zhang, Z. Q. Ye, X. Q. Liu, S. Q. Zhao, L. Wei and G. Gao (2007). "CPC:
1048 assess the protein-coding potential of transcripts using sequence features and
1049 support vector machine." *Nucleic Acids Res* **35**(Web Server issue): W345-349.
1050 DOI: 10.1093/nar/gkm391

1051

1052 Lal, A., M. P. Thomas, G. Altschuler, F. Navarro, E. O'Day, X. L. Li, C. Concepcion, Y.
1053 C. Han, J. Thiery, D. K. Rajani, A. Deutsch, O. Hofmann, A. Ventura, W. Hide and J.
1054 Lieberman (2011). "Capture of microRNA-bound mRNAs identifies the tumor
1055 suppressor miR-34a as a regulator of growth factor signaling." *PLoS Genet* **7**(11):
1056 e1002363. DOI: 10.1371/journal.pgen.1002363

1057

1058 Leveille, N., C. A. Melo, K. Rooijers, A. Diaz-Lagares, S. A. Melo, G. Korkmaz, R.
1059 Lopes, F. Akbari Moqadam, A. R. Maia, P. J. Wijchers, G. Geeven, M. L. den Boer, R.
1060 Kalluri, W. de Laat, M. Esteller and R. Agami (2015). "Genome-wide profiling of
1061 p53-regulated enhancer RNAs uncovers a subset of enhancers controlled by a
1062 lncRNA." *Nature Communications* **6**: 6520. DOI: 10.1038/ncomms7520

1063

1064 Liu, C., K. Kelnar, B. Liu, X. Chen, T. Calhoun-Davis, H. Li, L. Patrawala, H. Yan, C.
1065 Jeter, S. Honorio, J. F. Wiggins, A. G. Bader, R. Fagin, D. Brown and D. G. Tang
1066 (2011). "The microRNA miR-34a inhibits prostate cancer stem cells and
1067 metastasis by directly repressing CD44." *Nat Med* **17**(2): 211-215. DOI:
1068 10.1038/nm.2284

1069

1070 Memczak, S., M. Jens, A. Elefsinioti, F. Torti, J. Krueger, A. Rybak, L. Maier, S. D.
1071 Mackowiak, L. H. Gregersen, M. Munschauer, A. Loewer, U. Ziebold, M.
1072 Landthaler, C. Kocks, F. le Noble and N. Rajewsky (2013). "Circular RNAs are a
1073 large class of animal RNAs with regulatory potency." *Nature* **495**(7441): 333-
1074 338. DOI: 10.1038/nature11928

1075

1076 O'Leary, N. A., M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad, R. McVeigh, B.
1077 Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y.
1078 Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C. M.
1079 Farrell, T. Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V. S. Joardar, V. K.
1080 Kodali, W. Li, D. Maglott, P. Masterson, K. M. McGarvey, M. R. Murphy, K. O'Neill,
1081 S. Pujar, S. H. Rangwala, D. Rausch, L. D. Riddick, C. Schoch, A. Shkeda, S. S. Storz,
1082 H. Sun, F. Thibaud-Nissen, I. Tolstoy, R. E. Tully, A. R. Vatsan, C. Wallin, D. Webb,
1083 W. Wu, M. J. Landrum, A. Kimchi, T. Tatusova, M. DiCuccio, P. Kitts, T. D. Murphy
1084 and K. D. Pruitt (2016). "Reference sequence (RefSeq) database at NCBI: current

1085 status, taxonomic expansion, and functional annotation." *Nucleic Acids Res*
1086 **44**(D1): D733-745. DOI: 10.1093/nar/gkv1189
1087
1088 Ozsolak, F., P. Kapranov, S. Foissac, S. W. Kim, E. Fishilevich, A. P. Monaghan, B.
1089 John and P. M. Milos (2010). "Comprehensive polyadenylation site maps in yeast
1090 and human reveal pervasive alternative polyadenylation." *Cell* **143**(6): 1018-
1091 1029. DOI: 10.1016/j.cell.2010.11.020
1092
1093 Polson, A., E. Durrett and D. Reisman (2011). "A bidirectional promoter reporter
1094 vector for the analysis of the p53/WDR79 dual regulatory element." *Plasmid*
1095 **66**(3): 169-179. DOI: 10.1016/j.plasmid.2011.08.004
1096
1097 Rashi-Elkeles, S., H. J. Warnatz, R. Elkon, A. Kupershtain, Y. Chobod, A. Paz, V.
1098 Amstislavskiy, M. Sultan, H. Safer, W. Nietfeld, H. Lehrach, R. Shamir, M. L. Yaspo
1099 and Y. Shiloh (2014). "Parallel profiling of the transcriptome, cistrome, and
1100 epigenome in the cellular response to ionizing radiation." *Sci Signal* **7**(325): rs3.
1101 DOI: 10.1126/scisignal.2005032
1102
1103 Raver-Shapira, N., E. Marciano, E. Meiri, Y. Spector, N. Rosenfeld, N. Moskovits, Z.
1104 Bentwich and M. Oren (2007). "Transcriptional activation of miR-34a
1105 contributes to p53-mediated apoptosis." *Mol Cell* **26**(5): 731-743. DOI:
1106 10.1016/j.molcel.2007.05.017
1107
1108 Rinn, J. L., M. Kertesz, J. K. Wang, S. L. Squazzo, X. Xu, S. A. Brugmann, L. H.
1109 Goodnough, J. A. Helms, P. J. Farnham, E. Segal and H. Y. Chang (2007).
1110 "Functional demarcation of active and silent chromatin domains in human HOX
1111 loci by noncoding RNAs." *Cell* **129**(7): 1311-1323. DOI:
1112 10.1016/j.cell.2007.05.022
1113
1114 Rokavec, M., M. G. Oner, H. Li, R. Jackstadt, L. Jiang, D. Lodygin, M. Kaller, D. Horst,
1115 P. K. Ziegler, S. Schwitalla, J. Slotta-Huspenina, F. G. Bader, F. R. Greten and H.
1116 Hermeking (2015). "Corrigendum. IL-6R/STAT3/miR-34a feedback loop
1117 promotes EMT-mediated colorectal cancer invasion and metastasis." *J Clin Invest*
1118 **125**(3): 1362. DOI: 10.1172/JCI81340
1119
1120 Schneider, C. A., W. S. Rasband and K. W. Eliceiri (2012). "NIH Image to ImageJ:
1121 25 years of image analysis." *Nat Methods* **9**(7): 671-675.
1122
1123 Serviss, J. T. (2017). miR34AasRNAproject.
1124 https://github.com/GranderLab/miR34a_asRNA_project
1125
1126 Serviss, J. T., P. Johnsson and D. Grander (2014). "An emerging role for long non-
1127 coding RNAs in cancer metastasis." *Front Genet* **5**: 234. DOI:
1128 10.3389/fgene.2014.00234
1129
1130 Slabakova, E., Z. Culig, J. Remsik and K. Soucek (2017). "Alternative mechanisms
1131 of miR-34a regulation in cancer." *Cell Death Dis* **8**(10): e3100. DOI:
1132 10.1038/cddis.2017.495
1133

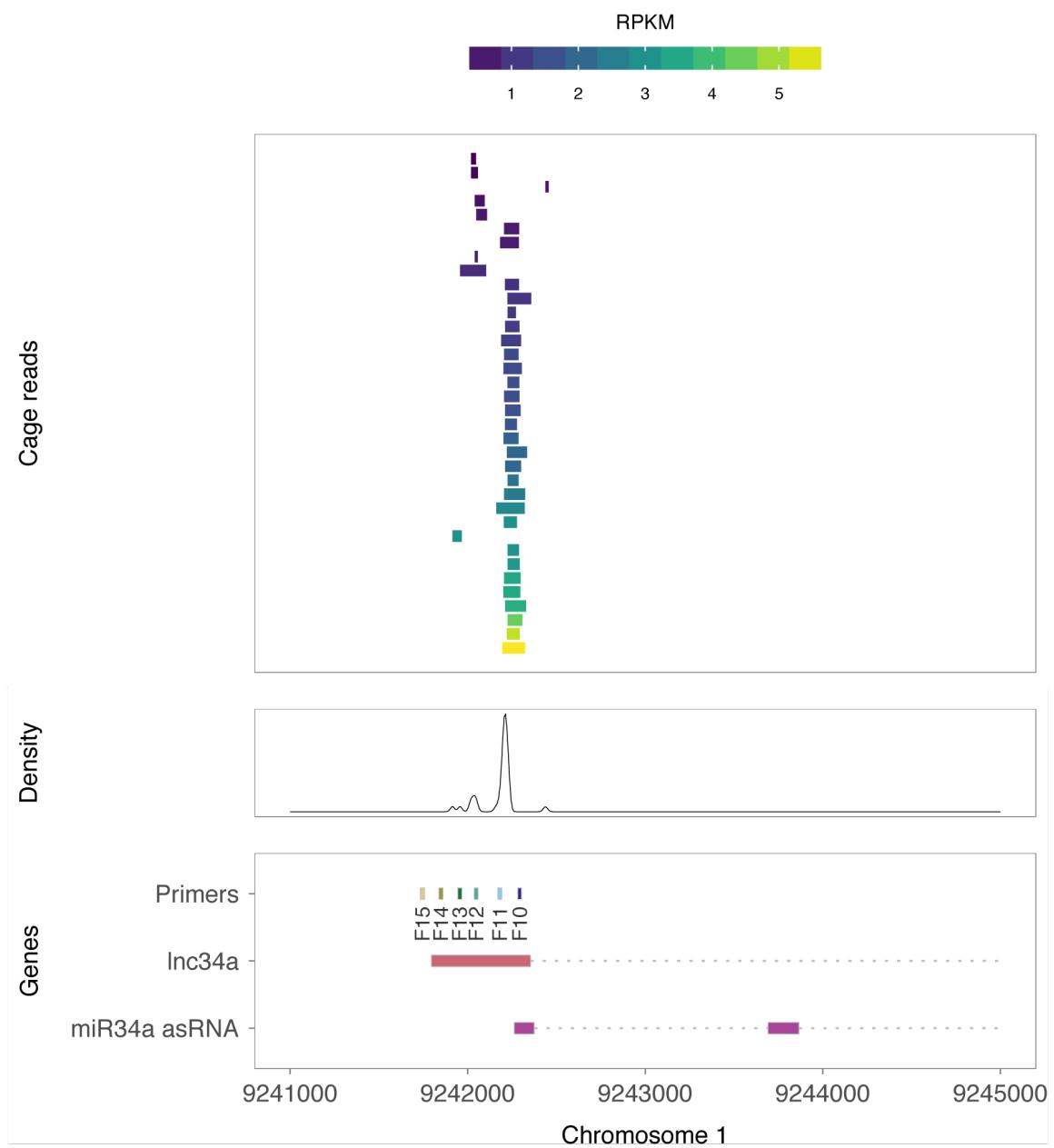
- 1134 Stahlhut, C. and F. J. Slack (2015). "Combinatorial Action of MicroRNAs let-7 and
1135 miR-34 Effectively Synergizes with Erlotinib to Suppress Non-small Cell Lung
1136 Cancer Cell Proliferation." *Cell Cycle* **14**(13): 2171-2180. DOI:
1137 10.1080/15384101.2014.1003008
1138
- 1139 Su, X., D. Chakravarti, M. S. Cho, L. Liu, Y. J. Gi, Y. L. Lin, M. L. Leung, A. El-Naggar,
1140 C. J. Creighton, M. B. Suraokar, I. Wistuba and E. R. Flores (2010). "TAp63
1141 suppresses metastasis through coordinate regulation of Dicer and miRNAs."
1142 *Nature* **467**(7318): 986-990. DOI: 10.1038/nature09459
1143
- 1144 Sun, F., H. Fu, Q. Liu, Y. Tie, J. Zhu, R. Xing, Z. Sun and X. Zheng (2008).
1145 "Downregulation of CCND1 and CDK6 by miR-34a induces cell cycle arrest."
1146 *FEBS Lett* **582**(10): 1564-1568. DOI: 10.1016/j.febslet.2008.03.057
1147
- 1148 Tarasov, V., P. Jung, B. Verdoodt, D. Lodygin, A. Epanchintsev, A. Menssen, G.
1149 Meister and H. Hermeking (2007). "Differential regulation of microRNAs by p53
1150 revealed by massively parallel sequencing: miR-34a is a p53 target that induces
1151 apoptosis and G1-arrest." *Cell Cycle* **6**(13): 1586-1593. DOI:
1152 10.4161/cc.6.13.4436
1153
- 1154 Team, R. C. (2017). "R: A Language and Environment for Statistical Computing."
1155 from <https://www.R-project.org/>.
1156
- 1157 Therneau, T. M. (2015). A Package for Survival Analysis in S. version 2.38.
1158 <https://CRAN.R-project.org/package=survival>
1159
- 1160 Turner, A. M., A. M. Ackley, M. A. Matrone and K. V. Morris (2012).
1161 "Characterization of an HIV-targeted transcriptional gene-silencing RNA in
1162 primary cells." *Hum Gene Ther* **23**(5): 473-483. DOI: 10.1089/hum.2011.165
1163
- 1164 Vogt, M., J. Mundig, M. Gruner, S. T. Liffers, B. Verdoodt, J. Hauk, L.
1165 Steinstraesser, A. Tannapfel and H. Hermeking (2011). "Frequent concomitant
1166 inactivation of miR-34a and miR-34b/c by CpG methylation in colorectal,
1167 pancreatic, mammary, ovarian, urothelial, and renal cell carcinomas and soft
1168 tissue sarcomas." *Virchows Arch* **458**(3): 313-322. DOI: 10.1007/s00428-010-
1169 1030-5
1170
- 1171 Wang, L., P. Bu, Y. Ai, T. Srinivasan, H. J. Chen, K. Xiang, S. M. Lipkin and X. Shen
1172 (2016). "A long non-coding RNA targets microRNA miR-34a to regulate colon
1173 cancer stem cell asymmetric division." *eLife* **5**. DOI: 10.7554/eLife.14620
1174
- 1175 Wang, L., H. J. Park, S. Dasari, S. Wang, J. P. Kocher and W. Li (2013). "CPAT:
1176 Coding-Potential Assessment Tool using an alignment-free logistic regression
1177 model." *Nucleic Acids Res* **41**(6): e74. DOI: 10.1093/nar/gkt006
1178
- 1179 Wang, X., J. Li, K. Dong, F. Lin, M. Long, Y. Ouyang, J. Wei, X. Chen, Y. Weng, T. He
1180 and H. Zhang (2015). "Tumor suppressor miR-34a targets PD-L1 and functions
1181 as a potential immunotherapeutic target in acute myeloid leukemia." *Cell Signal*
1182 **27**(3): 443-452. DOI: 10.1016/j.cellsig.2014.12.003

1183
1184 Wickham, H. (2016). gtable: Arrange 'Grobs' in Tables. R package version 0.2.0.
1185 <https://CRAN.R-project.org/package=gtable>
1186
1187 Wickham, H. (2017). scales: Scale Functions for Visualization. R package version
1188 0.5.0. <https://CRAN.R-project.org/package=scales>
1189
1190 Wickham, H. (2017). tidyverse: Easily Install and Load the 'Tidyverse'. R package
1191 version 1.2.1. <https://CRAN.R-project.org/package=tidyverse>
1192
1193 Wickham, L. H. a. H. (2017). rlang: Functions for Base Types and Core R and
1194 'Tidyverse' Features. R package version 0.1.4. [https://CRAN.R-
1195 project.org/package=rlang](https://CRAN.R-project.org/package=rlang)
1196
1197 Wickham, S. M. B. a. H. (2014). magrittr: A Forward-Pipe Operator for R. R
1198 package version 1.5. <https://CRAN.R-project.org/package=magrittr>
1199
1200 Wilkins, D. ggenes: Draw Gene Arrow Maps in 'ggplot2'. R package version
1201 0.2.0.9003. <https://github.com/wilcox/ggenes>
1202
1203 Xiao, N. (2017). liftr: Containerize R Markdown Documents. R package version
1204 0.7. <https://CRAN.R-project.org/package=liftr>
1205
1206 Xie, Y. (2017). knitr: A General-Purpose Package for Dynamic Report Generation
1207 in R. R package version 1.17. <https://yihui.name/knitr/>
1208
1209 Yang, P., Q. J. Li, Y. Feng, Y. Zhang, G. J. Markowitz, S. Ning, Y. Deng, J. Zhao, S.
1210 Jiang, Y. Yuan, H. Y. Wang, S. Q. Cheng, D. Xie and X. F. Wang (2012). "TGF-beta-
1211 miR-34a-CCL22 signaling-induced Treg cell recruitment promotes venous
1212 metastases of HBV-positive hepatocellular carcinoma." *Cancer Cell* **22**(3): 291-
1213 303. DOI: 10.1016/j.ccr.2012.07.023
1214
1215 Yap, K. L., S. Li, A. M. Munoz-Cabello, S. Raguz, L. Zeng, S. Mujtaba, J. Gil, M. J.
1216 Walsh and M. M. Zhou (2010). "Molecular interplay of the noncoding RNA ANRIL
1217 and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional
1218 silencing of INK4a." *Mol Cell* **38**(5): 662-674. DOI: 10.1016/j.molcel.2010.03.021
1219
1220 Yu, W., D. Gius, P. Onyango, K. Muldoon-Jacobs, J. Karp, A. P. Feinberg and H. Cui
1221 (2008). "Epigenetic silencing of tumour suppressor gene p15 by its antisense
1222 RNA." *Nature* **451**(7175): 202-206. DOI: 10.1038/nature06468
1223
1224 Zenz, T., J. Mohr, E. Eldering, A. P. Kater, A. Buhler, D. Kienle, D. Winkler, J. Durig,
1225 M. H. van Oers, D. Mertens, H. Dohner and S. Stilgenbauer (2009). "miR-34a as
1226 part of the resistance network in chronic lymphocytic leukemia." *Blood* **113**(16):
1227 3801-3808. DOI: 10.1182/blood-2008-08-172254
1228
1229
1230
1231

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 miR34a asRNAs 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 miR34a asRNA 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 miR34a asRNA 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. 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.



CAGE analysis at the miR34a locus: 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 miR34a asRNA 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, 15 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 miR34a asRNA 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 miR34a asRNA start site. Several additional peaks, albeit with a much lower average expression, aligns slightly upstream of the annotated miR34a asRNA 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 the transcription start site 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 claimed by the authors. An alternative interpretation could be that lnc34a expression is present in a subset of the examined cell lines although at too low levels to be detected. Finally, lnc34a may not be 5'-capped precluding its detection by CAGE.

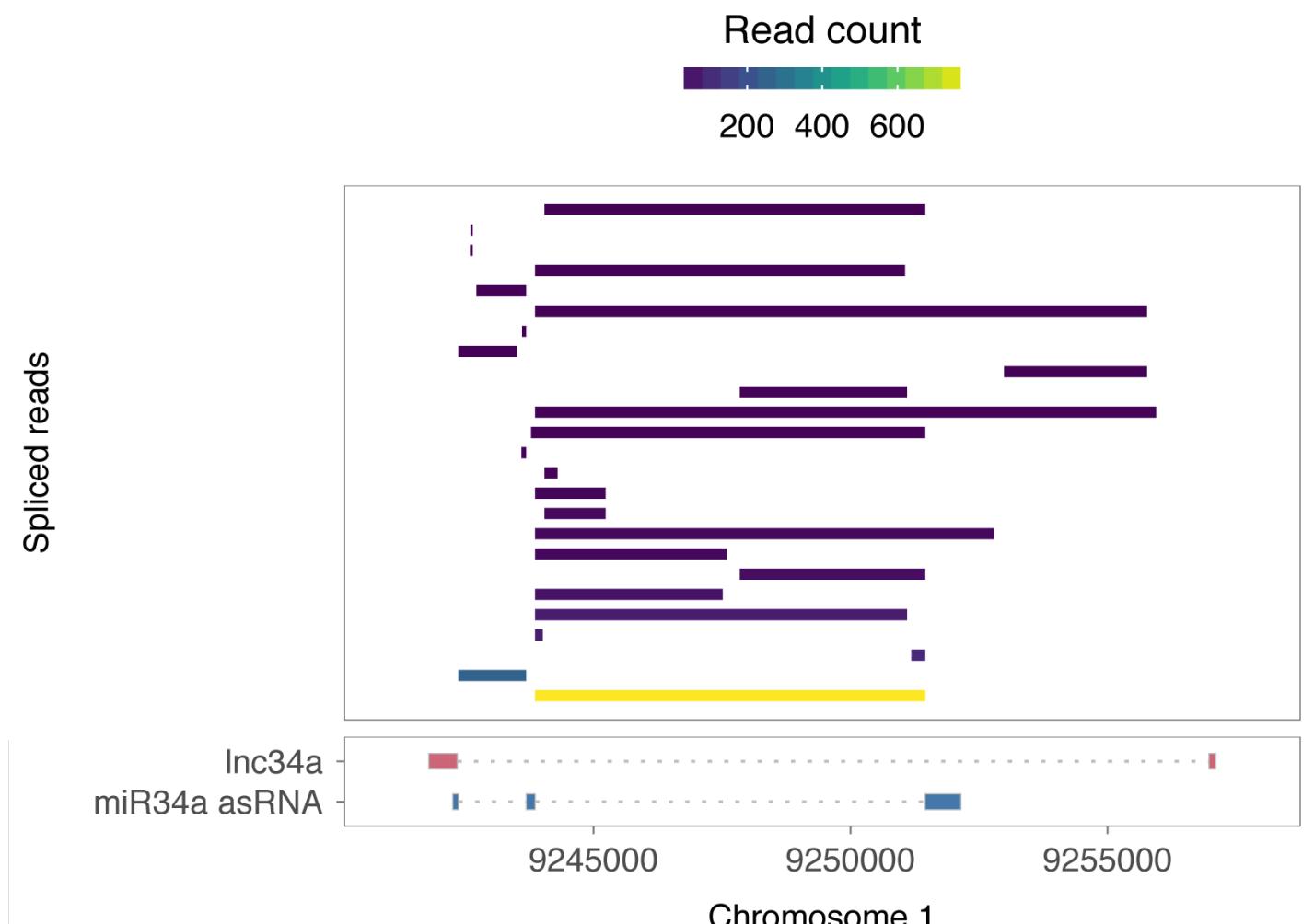
In order to lnc34a expression in a manner that is not dependant on 5'-capping, we proceeded to examine spliced RNA sequencing reads from 36 cell lines, taking advantage of the fact that lnc34a has an exon which is not present in any miR34a asRNA isoforms.

These results indicate that, although splice junctions corresponding to multiple isoforms detected by PCR cloning of the miR34a asRNA 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 expression may be confined to colon cancer stem cells as evidence for its expression in the examined cell lines is not forthcoming.

In addition, we note several 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, miR34a asRNA is strongly regulated by TP53 as the evidence shows in this and others (Léveillé 2015, Rashi-Elkeles 2014, Hünten 2015, Ashouri 2016, Kim 2017) findings.

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 miR34a asRNA 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 miR34a asRNA. For these reasons, we believe miR34a asRNA and lnc34a to be individual antisense RNA transcripts.



Splice junction analysis at the miR34a locus: 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 miR34a asRNA 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, Hmescat, Hmscbm, Hmscuc, Hob, Hpcpl, Hpiepc, Hsavec, Hsomm, Huvec, Hvmaf, Hwp, Imr90, Mcf7, Monocd14, Nhdf, Nhek, Nhemfm2, Nhemmm2, 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. miR34a asRNA and Lnc34a transcripts are represented for reference (bottom panel).

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