

# An antisense RNA capable of modulating the expression of the tumor suppressor microRNA-34a

Jason T. Serviss<sup>1\*</sup>, Felix Clemens Richter<sup>1,2</sup>, Jimmy Van den Eynden<sup>3</sup>, Nathanael Johansson Andrews<sup>1</sup>, Miranda Houtman<sup>1,4</sup>, Mattias Vesterlund<sup>1</sup>, Laura Schwarzmueller<sup>1,5</sup>, Per Johnsson<sup>6,7</sup>, Erik Larsson<sup>3</sup>, Dan Grandér<sup>1</sup>, Katja Pokrovskaja Tamm<sup>1</sup>

<sup>1</sup> Department of Oncology and Pathology, Karolinska Institutet, Stockholm, Sweden, SE-17177

<sup>2</sup> Kennedy Institute of Rheumatology, University of Oxford, Roosevelt Drive, Oxford OX3 7FY, UK

<sup>3</sup> Department of Medical Biochemistry and Cell Biology, Institute of Biomedicine, The Sahlgrenska Academy, University of Gothenburg, SE-405 30 Gothenburg, Sweden

<sup>4</sup> Rheumatology Unit, Department of Medicine, Karolinska University Hospital Solna, Karolinska Institutet, Stockholm, Sweden

<sup>5</sup> Laboratory for Experimental Oncology and Radiobiology (LEXOR), Center for Experimental Medicine, Institute of Biomedicine and Cell Technology, University of Szeged, Szeged, Hungary.

<sup>6</sup> Ludwig Institute for Cancer Research, Stockholm, Sweden

<sup>7</sup> Department of Cell and Molecular Biology, Karolinska Institutet

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

**\* Correspondence:**

Jason T. Serviss [jason.serviss@ki.se](mailto:jason.serviss@ki.se)

## Abstract

The microRNA-34a is a well-studied tumor suppressor microRNA (miRNA) and a direct downstream target of TP53 and has roles in several pathways associated with oncogenesis, such as proliferation, cellular growth, and differentiation. Due to its broad tumor suppressive activity, it is not surprising that *miR34a* expression is altered in a wide variety of solid tumors and hematological malignancies. However, the mechanisms by which *miR34a* is regulated in these cancers is largely unknown. In this study, we find that a long non-coding RNA transcribed antisense to the *miR34a* host gene, is critical for *miR34a* expression and mediation of its cellular functions in multiple types of human cancer. We name this long non-coding RNA *lncTAM34a*, and characterize its ability to facilitate *miR34a* expression under different types of 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 both indicates, its tumor suppressive effect and makes it a  
78 valuable prognostic marker (Cole et al. 2008, Gallardo et al. 2009, Zenz et al.  
79 2009, Cheng et al. 2010, Liu et al. 2011). Reduced *miR34a* transcription is  
80 mediated 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

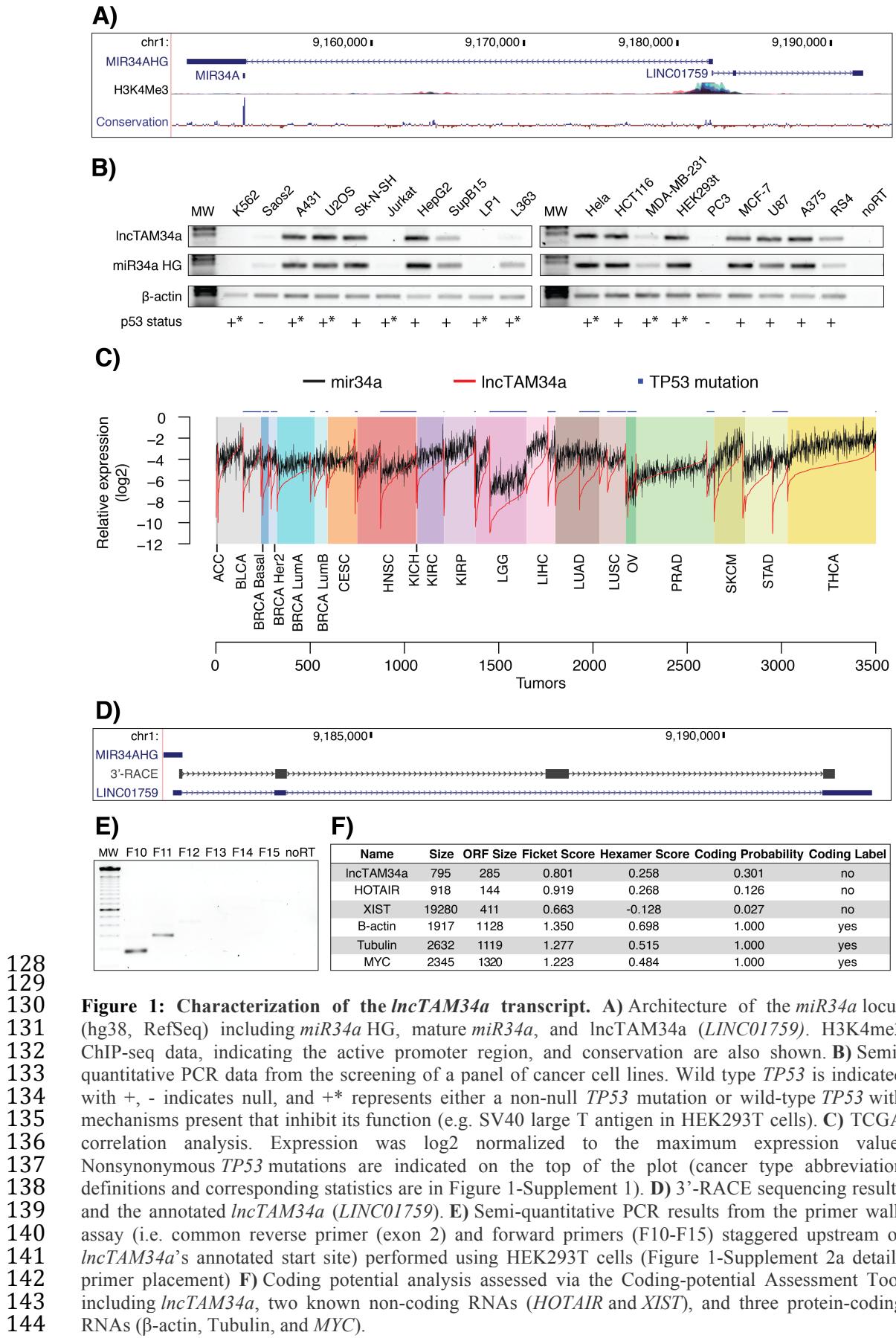
## 111 **Results**

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*<sup>-/-</sup> cells tending to have decreased or undetectable expression of both transcripts.

127



**Figure 1: Characterization of the *lncTAM34a* transcript.** **A)** Architecture of the *miR34a* locus (hg38, RefSeq) including *miR34a* HG, mature *miR34a*, and *lncTAM34a* (*LINC01759*). H3K4me3 ChIP-seq data, indicating the active promoter region, and conservation are also shown. **B)** Semi-quantitative PCR data from the screening of a panel of cancer cell lines. Wild type *TP53* is indicated with +, - indicates null, and +\* represents either a non-null *TP53* mutation or wild-type *TP53* with mechanisms present that inhibit its function (e.g. SV40 large T antigen in HEK293T cells). **C)** TCGA correlation analysis. Expression was log<sub>2</sub> normalized to the maximum expression value. Nonsynonymous *TP53* mutations are indicated on the top of the plot (cancer type abbreviation definitions and corresponding statistics are in Figure 1-Supplement 1). **D)** 3'-RACE sequencing results and the annotated *lncTAM34a* (*LINC01759*). **E)** Semi-quantitative PCR results from the primer walk assay (i.e. common reverse primer (exon 2) and forward primers (F10-F15) staggered upstream of *lncTAM34a*'s annotated start site) performed using HEK293T cells (Figure 1-Supplement 2a details primer placement) **F)** Coding potential analysis assessed via the Coding-potential Assessment Tool including *lncTAM34a*, two known non-coding RNAs (*HOTAIR* and *XIST*), and three protein-coding RNAs ( $\beta$ -actin, Tubulin, and *MYC*).

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 in a polyadenylation negative state (**Figure 1-Figure**  
175 **Supplement 2b**). Furthermore, we investigated the propensity of *lncTAM34a*  
176 to be alternatively spliced in U2OS cells, using PCR cloning followed by  
177 sequencing and found that the transcript is post-transcriptionally spliced to  
178 form multiple isoforms (**Figure 1-Figure Supplement 2c**). In order to evaluate  
179 the subcellular localization of *lncTAM34a*, we made use of RNA sequencing  
180 data from five cancer cell lines included in the ENCODE (Consortium 2012)  
181 project that had been fractionated into cytosolic and nuclear fractions. The  
182 analysis revealed that the *lncTAM34a* transcript primarily localizes to the  
183 nucleus with only a minor fraction in the cytosol (**Figure 1-Figure**  
184 **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). \*\*\* We hope to  
196 be able to scan for peptides matching to *IncTAM34a* in CPTAC and Geiger et  
197 al., 2012 before submission and will mention results here.\*\*\*

198

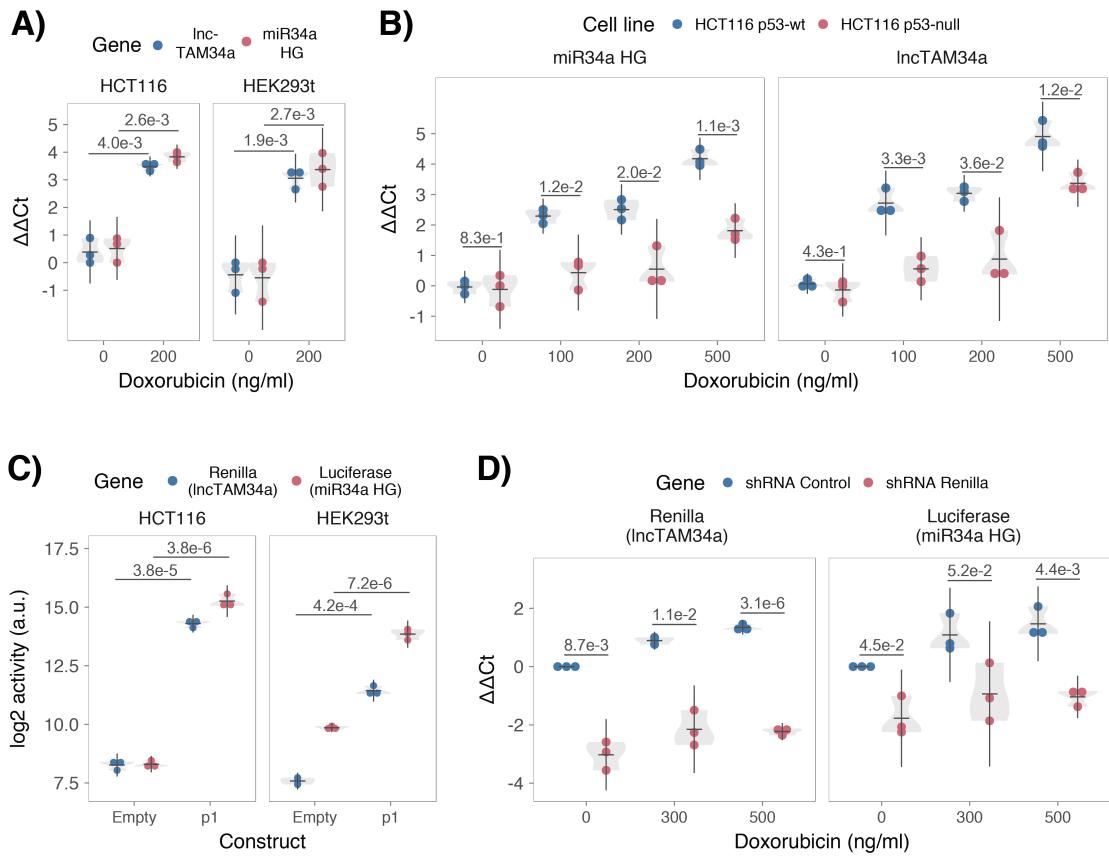
199 **TP53-mediated regulation of *IncTAM34a* expression**

200 *miR34a* is a known downstream target of TP53 and has been previously  
201 shown to exhibit increased expression within multiple contexts of cellular  
202 stress. *IncTAM34a* has also been shown to be induced upon TP53 activation  
203 in several global analyses of TP53-regulated lncRNAs (Rashi-Elkeles et al.  
204 2014, Hunten et al. 2015, Leveille et al. 2015, Ashouri et al. 2016, Kim et al.  
205 2017). To confirm these results in our biological systems, we treated  
206 HEK293T, embryonic kidney cells, and HCT116, colorectal cancer cells, with  
207 the DNA damaging agent doxorubicin to activate TP53. QPCR-mediated  
208 measurements of both *miR34a* HG and asRNA indicated that their expression  
209 levels were increased in response to doxorubicin treatment in both cell lines  
210 (**Fig. 2a**). To assess whether TP53 was responsible for the increase  
211 in *IncTAM34a* expression upon DNA damage, we treated *TP53<sup>+/+</sup>* and *TP53<sup>-/-</sup>*  
212 HCT116 cells with increasing concentrations of doxorubicin and monitored the  
213 expression of both *miR34a* HG and *IncTAM34a*. We observed a dose-  
214 dependent increase in both *miR34a* HG and *IncTAM34a* expression levels  
215 with increasing amounts of doxorubicin, revealing that these two transcripts  
216 are co-regulated, although, this effect was largely abrogated in *TP53<sup>-/-</sup>* cells  
217 (**Fig. 2b**). These results indicate that TP53 activation increases *IncTAM34a*  
218 expression upon DNA damage. Nevertheless, *TP53<sup>-/-</sup>* cells also showed a  
219 dose-dependent increase in both *miR34a* HG and asRNA, suggesting that

220 additional factors, other than *TP53* are capable of initiating an increase in  
221 expression of both of these transcripts upon DNA damage.

222

223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236



**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*<sup>+/+</sup> and *TP53*<sup>-/-</sup> 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.

237 The head-to-head orientation of *miR34a* HG and *lncTAM34a*, suggests that  
238 transcription is initiated from a single promoter in a bi-directional manner (**Fig**  
239 **1a**). To investigate whether *miR34a* HG and *lncTAM34a* are transcribed from  
240 the same promoter as divergent transcripts, we cloned the previously reported  
241 *miR34a* HG promoter, including the TP53 binding site, into a luciferase/renilla  
242 dual reporter vector which we hereafter refer to as p1 (**Figure 2-Figure**  
243 **Supplement 1a-b**) (Raver-Shapira et al. 2007). Upon transfection of p1 into  
244 HCT116 and HEK293T cell lines we observed increases in both luciferase  
245 and renilla indicating that *miR34a* HG and *lncTAM34a* expression can be  
246 regulated by a single promoter contained within the p1 construct (**Fig. 2c**).  
247

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

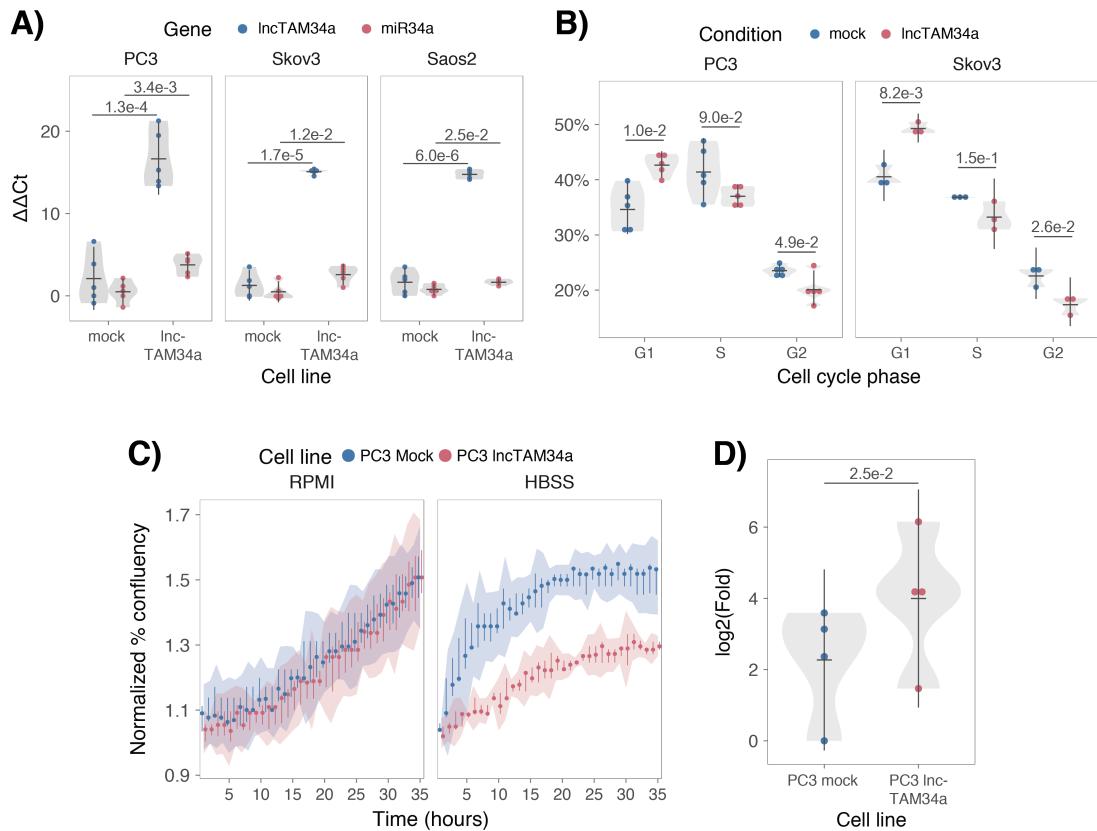
262 levels of *miR34a* HG and that the transcriptional product of the *lncTAM34a*  
263 within the p1 construct contributes to inducing a *miR34a* response. To further  
264 support these conclusions and better understand the role of *lncTAM34a*  
265 during TP53 activation, *TP53<sup>+/+</sup>* HCT116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

275

#### 276 ***lncTAM34a* can regulate *miR34a* host gene independently of TP53**

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

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



293

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

305 *miR34a* has been previously shown to regulate cell cycle progression, with  
306 *miR34a* induction causing G1 arrest (Raver-Shapira et al. 2007, Tarasov et al.  
307 2007). Cell cycle analysis via determination of DNA content showed a  
308 significant increase in G1 phase cells and a concomitant decrease in G2  
309 phase cells in the PC3 and Skov3 *lncTAM34a* over-expressing cell lines,  
310 indicating G1 arrest (**Fig. 3b**). The effects of *miR34a* on the cell cycle are  
311 mediated by its ability to target cell cycle regulators such as cyclin D1  
312 (*CCND1*) (Sun et al. 2008). Quantification of both *CCND1* RNA expression  
313 (**Figure 3-Figure Supplement 2a**) and protein levels (**Figure 3-Figure**  
314 **Supplement 2b**) in the PC3 *lncTAM34a* over-expressing cell line showed a  
315 significant decrease of *CCND1* levels compared to the mock control.  
316 Collectively, these results indicate that *lncTAM34a*-mediated induction of  
317 *miR34a* is sufficient to result in the corresponding *miR34a*-directed effects on  
318 cell cycle.

319  
320 *miR34a* is also a well-known inhibitor of cellular growth via its ability to  
321 negatively regulate growth factor signaling. Furthermore, starvation has been  
322 shown to induce *miR34a* expression causing down-regulation of numerous  
323 pro-survival growth factors (Lal et al. 2011). We further interrogated the  
324 effects of *lncTAM34a* over-expression by monitoring the growth of the PC3  
325 stable cell lines in both normal and starvation conditions via confluency  
326 measurements over a 35-hour period. Under normal growth conditions there  
327 is a small but significant reduction ( $P = 3.0\text{e-}8$ ; linear regression, **Fig. 3c**) in  
328 confluency in the *lncTAM34a* over-expressing cell lines compared to mock  
329 control. However, these effects on cell growth are drastically increased in

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

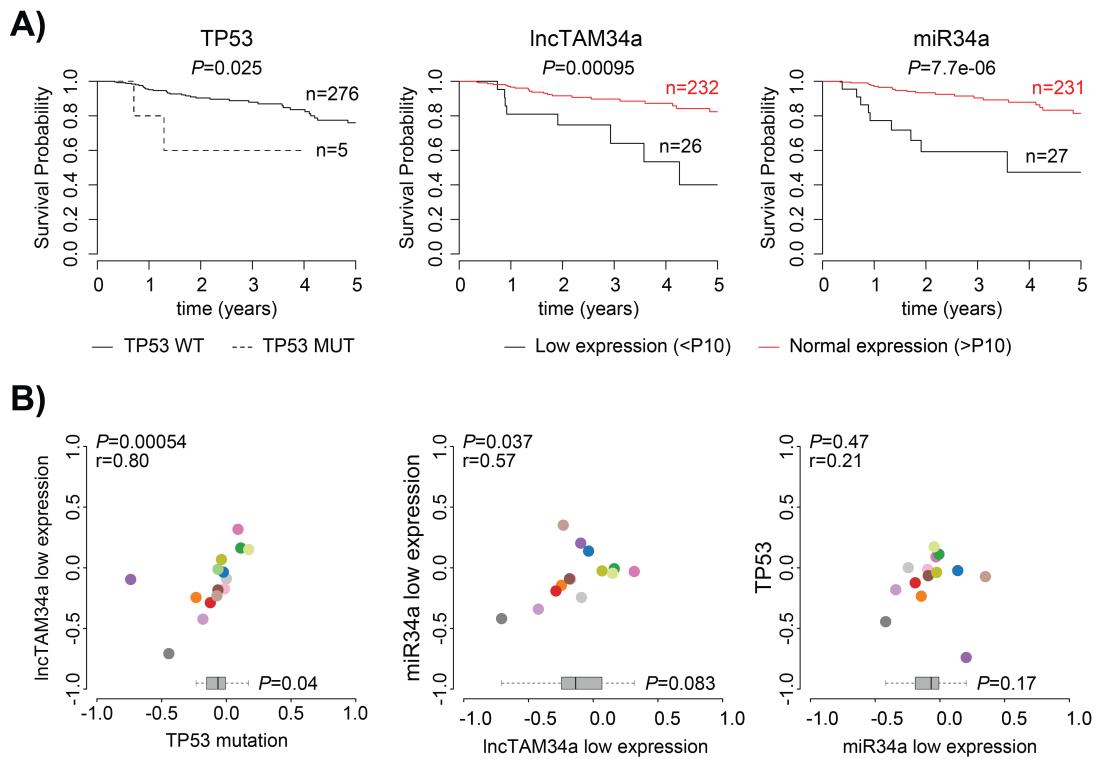
337

338 ***lncTAM34a* transcriptionally activates miR34a host gene**

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

351

352



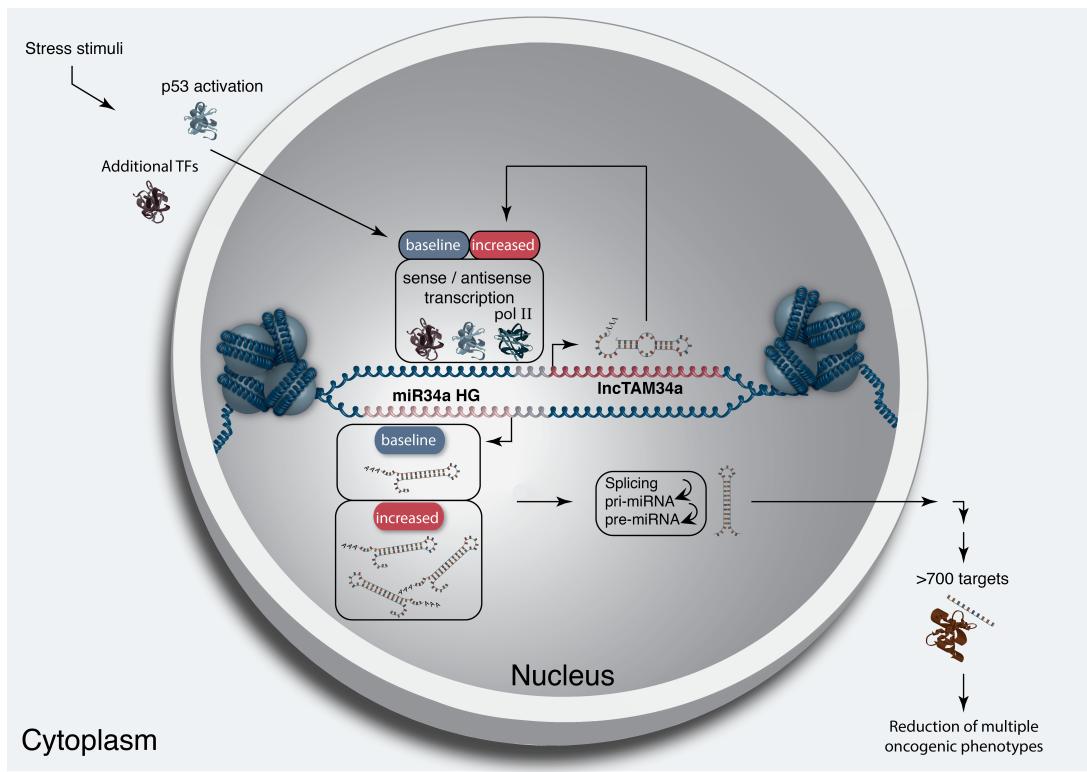
353 **Figure 4: Survival analysis.** A) Kaplan-Meier survival curves comparing the effects of *TP53*-mutated samples (left), low *lncTAM34a* expression (middle) and low *miR34a* expression (right) to control samples in papillary kidney cancer (results for other cancers in Figure 4-Supplement 1). B) Correlation 354 analysis between the effects on the 5-year survival probability of *TP53*-mutated samples, low 355 *lncTAM34a* expression, and low *miR34a* expression as indicated. For each variable the 5-year survival 356 probability was compared to the control group (negative value indicates lower survival, positive value 357 indicates higher survival). Spearman correlation coefficients are given on top left of each plot. Each dot 358 indicates one cancer type (see Fig.1c for legend). Boxplots on the bottom summarize the effects for the 359 parameter on the x-axis, with indication of p-values, as calculated using paired Wilcoxon signed rank 360 test. Low expression was defined as *TP53* non-mutated samples having expression values in the bottom 361 10th percentile.

362

363

364  
365 **Low *lncTAM34a* expression levels are associated with decreased**  
366 **survival**

367  
368 As *TP53* mutations and low expression of *miR34a* have been associated with  
369 worse prognosis in cancer, we compared survival rates of samples with low  
370 expression of *lncTAM34a* (bottom 10th percentile) to control samples in 17  
371 cancer types from TCGA (**Figure 4-Supplement 1**) (Gallardo et al. 2009,  
372 Zenz et al. 2009, Liu et al. 2011). To correct for the effect of *TP53* mutations  
373 we focused on non-*TP53* mutated samples, and noted a worse survival for the  
374 low expression group in several cancers. This effect was most pronounced in  
375 papillary kidney cancer (unadjusted  $P=0.00095$ ; **Fig. 4a**). By systematically  
376 comparing 5-year survival probabilities between the low expression group and  
377 the control group for each cancer we found a median reduction of 5-year  
378 survival probability of 9.6% ( $P=0.083$ ; Wilcoxon signed rank test; **Fig. 4b**).  
379 Furthermore, we found that *lncTAM34a* expression showed similar patterns in  
380 terms of direction and strength of association with 5-year survival probability  
381 as *miR34a* expression ( $r=0.57$ ,  $P=0.037$ ) and *TP53* mutations ( $r=0.80$ ,  
382  $P=0.00054$ ) across the different cancer types (**Fig. 4b**). Although these results  
383 do not implicate any causal relationship, they do indicate a striking similarity  
384 between the association of worse prognosis and *TP53* mutations, low  
385 *miR34a*, and low *lncTAM34a* expression.



386

387 **Figure 5: A graphical summary of the proposed *IncTAM34a* function.** Stress stimuli, originating in  
 388 the cytoplasm or nucleus, activate TP53 as well as additional factors. These factors then bind to  
 389 the *miR34a* promoter and drive baseline transcription levels of the sense and antisense  
 390 strands. *IncTAM34a* serves to further increases *miR34a* HG transcription levels resulting in enrichment  
 391 of polymerase II at the *miR34a* promoter and a positive feed-forward loop. *IncTAM34a*-mediated  
 392 increases in *miR34a* HG potentially occur via direct DNA binding alone, by direct DNA binding and  
 393 recruitment of additional factors, or through a yet unknown mechanism. *miR34a* HG then, in turn,  
 394 is spliced and processed before being exported to the cytoplasm. The *miR34a* pre-miRNA then undergoes  
 395 further processing before the mature *miR34a* binds to the RISC complex allowing it to bind and repress  
 396 its targets and exert its tumor suppressive effects.

397 **Discussion**

398  
399 Multiple studies have previously shown asRNAs to be crucial for the  
400 appropriate regulation of cancer-associated protein-coding genes and that  
401 their dysregulation can lead to a perturbation of tumor suppressive and  
402 oncogenic pathways, as well as, cancer-related phenotypes (Yu et al. 2008,  
403 Yap et al. 2010, Serviss et al. 2014, Balbin et al. 2015). Here we show that  
404 asRNAs are also capable of regulating cancer-associated miRNAs resulting in  
405 similar consequences as protein-coding gene dysregulation (**Fig. 4**).  
406 Interestingly, we show that, both in the presence and absence of  
407 *TP53*, *lncTAM34a* provides an additional regulatory level to control *miR34a*  
408 expression in both homeostasis and upon encountering various forms of  
409 cellular stress. Furthermore, we find that a *lncTAM34a*-mediated increase in  
410 *miR34a* expression is sufficient to drive the appropriate cellular responses to  
411 these stress stimuli (**Fig. 2d and Fig. 3c**). Previous studies have exploited  
412 various molecular biology methods to up-regulate *miR34a* expression in a  
413 *TP53*-deficient background showing similar phenotypic outcomes although,  
414 here we show a novel mechanism by which this can be achieved in an  
415 endogenous manner (Liu et al. 2011, Ahn et al. 2012, Yang et al. 2012,  
416 Stahlhut et al. 2015, Wang et al. 2015).

417

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

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

444  
445 An antisense transcript arising from the *miR34a* locus, *Lnc34a*, has been  
446 previously reported to negatively regulate the expression of *miR34a* (Wang et  
447 al. 2016). Although the *Lnc34a* and *lncTAM34a* transcripts share some

448 sequence similarity, we believe them to be separate RNAs that are,  
449 potentially, different isoforms of the same gene. We utilized CAGE and  
450 RNAseq data from the ENCODE project to evaluate the presence of  
451 *IncTAM34a* and *Lnc34a* in 28 and 36 commonly used cancer cell lines,  
452 respectively. Although the results show the presence of *IncTAM34a* in these  
453 cell lines, we find no evidence for *Lnc34a* transcription (**Supplementary**  
454 **Document 1**). These results are in line with the findings of Wang et al.  
455 indicating that *Lnc34a* is highly expressed in colon cancer stem cell spheres  
456 compared to other cell types used in their study and, furthermore suggests,  
457 that these two transcripts are not commonly co-expressed. The fact  
458 that *IncTAM34a* and *Lnc34a* would appear to have opposing roles in their  
459 regulation of *miR34a*, further underlines the complexity of the regulation at  
460 this locus.

461

462 Clinical trials utilizing *miR34a* replacement therapy have previously been  
463 conducted but, disappointingly, were terminated after adverse side effects of  
464 an immunological nature were observed in several of the patients (Slabakova  
465 et al. 2017). Although it is not presently clear if these side effects were caused  
466 by *miR34a* or the liposomal carrier used to deliver the miRNA, the multitude of  
467 evidence indicating *miR34a*'s crucial role in oncogenesis still makes its  
468 therapeutic induction an interesting strategy and needs further investigation.  
469 Our results indicate an association between survival probability and low  
470 *IncTAM34a* expression making it an attractive candidate for controlled  
471 preclinical studies. Due to *IncTAM34a*-mediated positive feedback on *miR34a*  
472 expression, initiation of this feedback mechanism may be able to provide a

473 sustained *miR34a* induction in a relatively more robust manner than *miR34a*  
474 replacement alone. In summary, our results have identified *lncTAM34a* as a  
475 vital player in the regulation of *miR34a* and its particular importance in typical  
476 examples of cellular stress encountered in cancer. The conclusions drawn in  
477 this study provide essential insight regarding asRNA-mediated regulation of  
478 cancer-associated miRNAs and, contribute to fundamental knowledge  
479 concerning *miR34a* regulation necessary for its efficient induction in clinical  
480 settings.

481

## 482 **Materials and Methods**

### 483 **Cell Culture**

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

496

### 497 **Bioinformatics, Data Availability, and Statistical Testing**

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

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

501

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

513

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

517

### 518 **Coding Potential**

519 Protein-coding capacity was evaluated using the Coding-potential  
520 Assessment Tool (Wang et al. 2013) and Coding-potential Calculator (Kong et  
521 al. 2007) with default settings. Transcript sequences for use with Coding-  
522 potential Assessment Tool were downloaded from the UCSC genome  
523 browser using the Ensembl

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

529

### 530 **shRNAs**

531 shRNA-expressing constructs were cloned into the U6M2 construct using the  
532 BgIII and KpnI restriction sites as previously described (Amarzguioui et al.  
533 2005). shRNA constructs were transfected using Lipofectamine 2000 or 3000  
534 (ThermoFisher Scientific, Cat# 12566014 and L3000015). The sequences  
535 targeting renilla is as follows: shRenilla 1.1 (AAT ACA CCG CGC TAC TGG  
536 C), shRenilla 2.1 (TAA CGG GAT TTC ACG AGG C).

537

### 538 **Bi-directional Promoter Cloning**

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

543

### 544 **Promoter Activity**

545 Cells were co-transfected with the p1 renilla/firefly bidirectional promoter  
546 construct (Polson et al. 2011) and GFP by using Lipofectamine 2000 (Life  
547 Technologies, Cat# 12566014). The expression of GFP and luminescence  
548 was measured 24 h post transfection by using the Dual-Glo Luciferase Assay

549 System (Promega, Cat# E2920) and detected by the GloMax-Multi+ Detection  
550 System (Promega, Cat# SA3030). The expression of luminescence was  
551 normalized to GFP.

552

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

554 The U6 promoter was amplified from the U6M2 cloning plasmid (Amarzguioui  
555 et al. 2005) and ligated into the Not1 restriction site of the pHIV7-IMPDH2  
556 vector (Turner et al. 2012). *IncTAM34a* was PCR amplified and subsequently  
557 cloned into the Nhe1 and Pac1 restriction sites in the pHIV7-IMPDH2-U6  
558 plasmid.

559

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

561 Lentivirus production was performed as previously described in (Turner et al.  
562 2012). Briefly, HEK293T cells were transfected with viral and expression  
563 constructs using Lipofectamine 2000 (ThermoFisher Scientific, Cat#  
564 12566014), after which viral supernatants were harvested 48 and 72 hours  
565 post-transfection. Viral particles were concentrated using PEG-IT solution  
566 (Systems Biosciences, Palo Alto, CA, USA. Cat# LV825A-1) according to the  
567 manufacturer's recommendations. HEK293T cells were used for virus titration  
568 and GFP expression was evaluated 72hrs post-infection via flow cytometry  
569 (LSRII, BD Biosciences, San Jose, CA, USA) after which TU/ml was  
570 calculated.

571

572 Stable lines were generated by infecting cells with a multiplicity of infection of  
573 1 after which 1-2 µM mycophenolic acid (Merck, Kenilworth, NJ, USA. Cat#

574 M5255) selection was initiated 48-72 hours post-infection. Cells were  
575 expanded as the selection process was monitored via flow cytometry analysis  
576 (LSRII, BD Biosciences) of GFP and selection was terminated once > 90% of  
577 the cells were GFP positive. Quantification of *lncTAM34a* over-expression and  
578 *miR34a* was performed in biological quintuplet for all cell lines.

579

#### 580 **Western Blotting**

581 Samples were lysed in 50 mM Tris-HCl (Sigma Aldrich, St. Louis, MO, USA.  
582 Cat# T2663), pH 7.4, 1% NP-40 (Sigma Aldrich, Cat# I8896), 150 mM NaCl  
583 (Sigma Aldrich, Cat# S5886), 1 mM EDTA (Promega, Madison, WI, USA.  
584 Cat# V4231), 1% glycerol (Sigma Aldrich, Cat# G5516), 100 µM vanadate  
585 (Sigma Aldrich, Cat# S6508), protease inhibitor cocktail (Roche Diagnostics,  
586 Basel, Switzerland, Cat# 004693159001) and PhosSTOP (Roche  
587 Diagnostics, Cat# 04906837001). Lysates were subjected to SDS-PAGE and  
588 transferred to PVDF membranes. The proteins were detected by western blot  
589 analysis by using an enhanced chemiluminescence system (Western  
590 Lightning-ECL, PerkinElmer, Waltham, MA, USA. Cat# NEL103001EA).  
591 Antibodies used were specific for CCND1 1:1000 (Cell Signaling, Danvers,  
592 MA, USA. Cat# 2926), and GAPDH 1:5000 (Abcam, Cambridge, UK, Cat#  
593 ab9485). All western blot quantifications were performed using ImageJ  
594 (Schneider et al. 2012).

595

#### 596 **RNA Extraction and cDNA Synthesis**

597 For downstream SYBR green applications, RNA was extracted using the  
598 RNeasy mini kit (Qiagen, Venlo, Netherlands, Cat# 74106) and subsequently

599 treated with DNase (Ambion Turbo DNA-free, ThermoFisher Scientific, Cat#  
600 AM1907). 500ng RNA was used for cDNA synthesis using MuMLV  
601 (ThermoFisher Scientific, Cat# 28025013) and a 1:1 mix of oligo(dT) and  
602 random nanomers.

603

604 For analysis of miRNA expression with Taqman, samples were isolated with  
605 TRIzol reagent (ThermoFisher Scientific, Cat# 15596018) and further  
606 processed with the miRNeasy kit (Qiagen, Cat# 74106). cDNA synthesis was  
607 performed using the TaqMan MicroRNA Reverse Transcription Kit  
608 (ThermoFisher Scientific, Cat# 4366597) using the corresponding oligos  
609 according to the manufacturer's recommendations.

610

### 611 **QPCR and PCR**

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

617

618 QPCR for miRNA expression analysis was performed according to the primer  
619 probe set manufacturers recommendations (ThermoFisher Scientific) and  
620 using the TaqMan Universal PCR Master Mix (ThermoFisher Scientific, Cat#  
621 4304437) with the same cycling scheme as above. Primer and probe sets for  
622 TaqMan were also purchased from ThermoFisher Scientific (Life  
623 Technologies at time of purchase, TaqMan® MicroRNA Assay, hsa-miR-34a,

624 human, Cat# 4440887, Assay ID: 000426 and Control miRNA Assay, RNU48,  
625 human, Cat# 4440887, Assay ID: 001006).

626

627 The  $\Delta\Delta Ct$  method was used to quantify gene expression. All QPCR-based  
628 experiments were performed in at least technical duplicate. Primers for all  
629 PCR-based experiments are listed in **Supplementary Document 2** and  
630 arranged by figure.

631

### 632 **Cell Cycle Distribution**

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

642

### 643 **3' Rapid Amplification of cDNA Ends**

644 3'-RACE was performed as described as previously in (Johnsson et al. 2013).  
645 Briefly, U2OS cell RNA was polyA-tailed using yeast polyA polymerase  
646 (ThermoFisher Scientific, Cat# 74225Z25KU) after which cDNA was  
647 synthesized using oligo(dT) primers. Nested-PCR was performed first using a  
648 forward primer in *lncTAM34a* exon 1 and a tailed oligo(dT) primer followed by

649 a second PCR using an alternate *lncTAM34a* exon 1 primer and a reverse  
650 primer binding to the tail of the previously used oligo(dT) primer. PCR  
651 products were gel purified and cloned the Strata Clone Kit (Agilent  
652 Technologies, Santa Clara, CA, USA. Cat# 240205), and sequenced.

653

654 **Chromatin Immunoprecipitation**

655 The ChIP was performed as previously described in (Johnsson et al. 2013)  
656 with the following modifications. Cells were crosslinked in 1% formaldehyde  
657 (Merck, Cat# 1040039025), quenched with 0.125M glycine (Sigma Aldrich,  
658 Cat# G7126), and lysed in cell lysis buffer comprised of: 5mM PIPES (Sigma  
659 Aldrich, Cat# 80635), 85mM KCL (Merck, Cat# 4936), 0.5% NP40 (Sigma  
660 Aldrich, Cat# I8896), protease inhibitor (Roche Diagnostics, Cat#  
661 004693159001). Samples were then sonicated in 50mM TRIS-HCL pH 8.0  
662 (Sigma Aldrich, MO, USA, Cat# T2663) 10mM EDTA (Promega, WI, USA,  
663 Cat# V4231), 1% SDS (ThermoFisher Scientific, Cat# AM9822), and protease  
664 inhibitor (Roche Diagnostics, Cat# 004693159001) using a Bioruptor  
665 Sonicator (Diagenode, Denville, NJ, USA). Samples were incubated over  
666 night at 4°C with the polII antibody (Abcam, Cat# ab5095) and subsequently  
667 pulled down with Salmon Sperm DNA/Protein A Agarose (Millipore, Cat# 16-  
668 157) beads. DNA was eluted in an elution buffer of 1% SDS (ThermoFisher  
669 Scientific, Cat# AM9822) 100mM NaHCO3 (Sigma Aldrich, Cat# 71631),  
670 followed by reverse crosslinking, RNaseA (ThermoFisher Scientific, Cat#  
671 1692412) and protease K (New England Biolabs, Ipswich, MA, USA, Cat#  
672 P8107S) treatment. The DNA was eluted using Qiagen PCR purification kit  
673 (Cat# 28106) and quantified via QPCR. QPCR was performed in technical

674 duplicate using the standard curve method and reported absolute values. The  
675 fraction of input was subsequently calculated using the mean of the technical  
676 replicates followed by calculating the fold over the control condition. Statistical  
677 testing was performed using 4 biological replicates with the null hypothesis  
678 that the true log 2 fold change values were equal to zero.

679

## 680 **Confluency Analysis**

681 Cells were incubated in the Spark Multimode Microplate (Tecan, Männedorf,  
682 Switzerland) reader for 48 hours at 37°C with 5% CO<sub>2</sub> in a humidity chamber.  
683 Confluency was measured every hour using bright-field microscopy and the  
684 percentage of confluency was reported via the plate reader's inbuilt algorithm.  
685 Percentage of confluency was normalized to the control sample in each  
686 condition (shown in figure) and then ranked to move the data to a linear scale.  
687 Using the mean of the technical duplicates in three biological replicates, the  
688 rank was then used to construct a linear model, of the dependency of the rank  
689 on the time and cell lines variables for each growth condition. Reported p-  
690 values are derived from the t-test, testing the null hypothesis that the  
691 coefficient estimate of the cell line variable is equal to 0.

692

## 693 **Pharmacological Compounds**

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

695

## 696 **Cellular Localization Analysis**

697 Quantified RNAseq data from 11 cell lines from the GRCh38 assembly was  
698 downloaded from the ENCODE project database and quantifications for

699 *lncTAM34a* (ENSG00000234546), GAPDH (ENSG00000111640), and  
700 MALAT1 (ENSG00000251562) were extracted. Cell lines for which data was  
701 downloaded include: A549, GM12878, HeLa-S3, HepG2, HT1080, K562  
702 MCF-7, NCI-H460, SK-MEL-5, SK-N-DZ, SK-N-SH. Initial exploratory analysis  
703 revealed that several cell lines should be removed from the analysis due to a)  
704 a larger proportion of GAPDH in the nucleus than cytoplasm or b) variation of  
705 *lncTAM34a* expression is too large to draw conclusions, or c) they have no or  
706 low (<6 TPM) *lncTAM34a* expression. Furthermore, only polyadenylated  
707 libraries were used in the final analysis, due to the fact that the cellular  
708 compartment enrichment was improved in these samples. All analyzed genes  
709 are reported to be polyadenylated. In addition, only samples with 2 biological  
710 replicates were retained. For each cell type, gene, and biological replicate the  
711 fraction of transcripts per million (TPM) in each cellular compartment was  
712 calculated as the fraction of TPM in the specific compartment by the total  
713 TPM. The mean and standard deviation for the fraction was subsequently  
714 calculated for each cell type and cellular compartment and this information  
715 was represented in the final figure.

716

## 717 **CAGE Analysis**

718 All available CAGE data from the ENCODE project (Consortium 2012) for 36  
719 cell lines was downloaded from the UCSC genome browser (Kent et al. 2002)  
720 for genome version hg19. Of these, 28 cell lines had CAGE transcription start  
721 sites (TSS) mapping to the plus strand of chromosome 1 and in regions  
722 corresponding to 200 base pairs upstream of the *lnc34a* start site (9241796 -  
723 200) and 200 base pairs upstream of the GENCODE annotated *lncTAM34a*

724 start site (9242263 + 200). These cell lines included: HFDPC, H1-hESC,  
725 HMEpC, HAoEC, HPIEpC, HSaVEC, GM12878, hMSC-BM, HUVEC,  
726 AG04450, hMSC-UC, IMR90, NHDF, SK-N-SH\_RA, BJ, HOB, HPC-PL,  
727 HAoAF, NHEK, HVMF, HWP, MCF-7, HepG2, hMSC-AT, NHEM.f\_M2,  
728 SkMC, NHEM\_M2, and HCH. In total 74 samples were included. 17 samples  
729 were polyA-, 47 samples were polyA+, and 10 samples were total RNA. In  
730 addition, 34 samples were whole cell, 15 enriched for the cytosolic fraction, 15  
731 enriched for the nucleolus, and 15 enriched for the nucleus. All CAGE  
732 transcription start sites were plotted and the RPKM of the individual reads was  
733 used to color each read to indicate their relative abundance. In cases where  
734 CAGE TSS spanned identical regions, the RPMKs of the regions were  
735 summed and represented as one CAGE TSS in the figure. In addition, a  
736 density plot shows the distribution of the CAGE reads in the specified  
737 interval.

738

### 739 Splice Junction Analysis

740 All available whole cell (i.e. non-fractionated) spliced read data originating  
741 from the Cold Spring Harbor Lab in the ENCODE project (Consortium 2012)  
742 for 38 cell lines was downloaded from the UCSC genome browser (Kent et al.  
743 2002). Of these cell lines, 36 had spliced reads mapping to the plus strand of  
744 chromosome 1 and in the region between the *lnc34a* start (9241796) and  
745 transcription termination (9257102) site (note that *lncTAM34a* resides totally  
746 within this region). Splice junctions from the following cell lines were included  
747 in the final figure: A549, Ag04450, Bj, CD20, CD34 mobilized, Gm12878,  
748 H1hesc, Haoaf, Haoec, Hch, Helas3, Hepg2, Hfdpc, Hmec, Hmepc, Hmscat,

749 Hmscbm, Hmscuc, Hob, Hpcpl, Hpiepc, Hsavec, Hsmm, Huvec, Hvmf, Hwp,  
750 Imr90, Mcf7, Monocd14, Nhdf, Nhek, Nhjemfm2, Nhjemm2, Nhlf, Skmc, and  
751 Sknsh. All splice junctions were included in the figure and colored according  
752 to the number of reads corresponding to each. In cases where identical reads  
753 were detected multiple times, the read count was summed and represented  
754 as one read in the figure.

755

## 756 **TCGA Data Analysis**

757 RNA-Seq data and copy number data were downloaded from TCGA and  
758 processed as described previously (Ashouri et al. 2016). Briefly, RNA-Seq  
759 data were aligned to the human hg19 assembly and quantified using  
760 GENCODE (v19) annotated HTSeq-counts and FPKM normalizations.  
761 Expression data from *miR34a* and *lncTAM34a* (identified as RP3-510D11.2)  
762 were used for further analysis. Copy number amplitudes for GENCODE genes  
763 were determined from segmented copy-number data. Samples that were  
764 diploid for *lncTAM34a* were identified as those samples that had copy number  
765 amplitudes between -0.1 and 0.1.

766

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

770

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

774

## 775 **Acknowledgments**

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

778

779 **Competing Interests**

780  
781 The authors declare no competing interests.

782

783  
784 **Funding**

785  
786 This work has been supported by the Swedish Research Council [521-2012-  
787 2037], Swedish Cancer Society [150768], Cancer Research Foundations of  
788 Radiumhemmet [144063] and the Swedish Childhood Cancer Foundation  
789 [PR2015-0009].

790

791

792 **Figure Supplements**

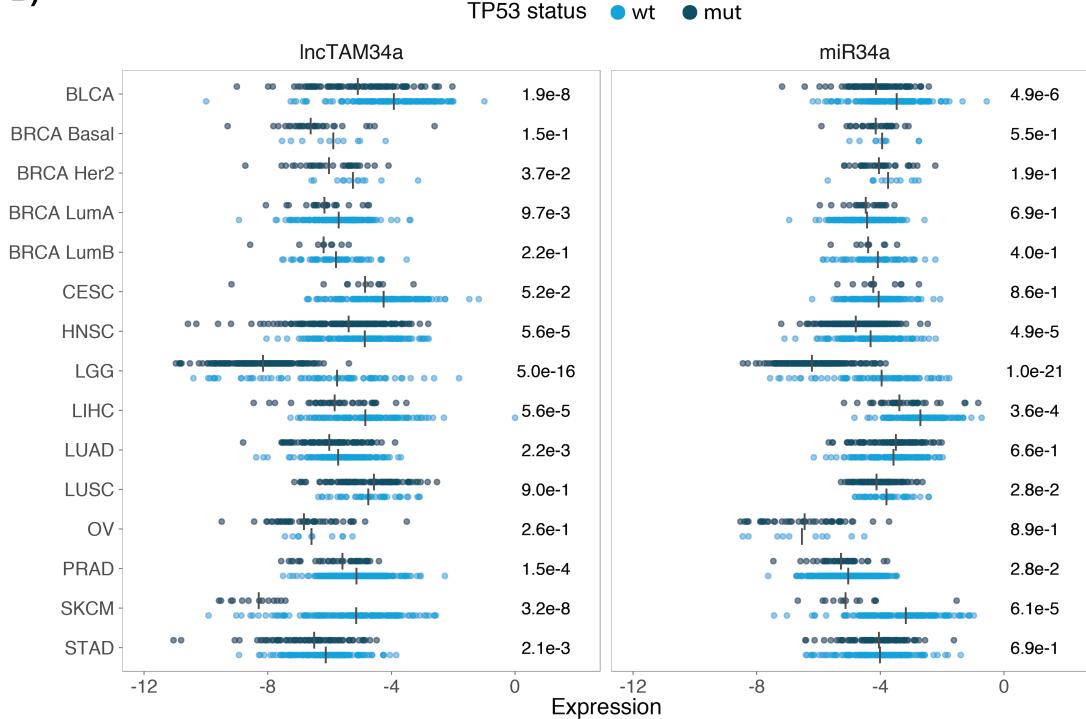
793  
794 Figure 1-Supplement 1: TCAG expression levels and correlation analysis  
795 statistics.  
796  
797 Figure 1-Supplement 2: Molecular characteristics of *IncTAM34a*.  
798  
799 Figure 2-Supplement 1: A schematic representation of the p1 construct.  
800  
801 Figure 2-Supplement 2: Evaluating the effects of *IncTAM34a* down-regulation.  
802  
803 Figure 3-Supplement 1: Physiological relevance of *IncTAM34a*  
804 overexpression.  
805  
806 Figure 3-Supplement 2: Effects of *IncTAM34a* overexpression on cyclin D1.  
807  
808 Figure 4-Supplement 1: Survival analysis in 17 cancers from TCGA.  
809  
810 Supplementary Document 1: Evaluating the relationship between *IncTAM34a*  
811 and *Inc34a*.  
812  
813 Supplementary Document 2: A table of primers used in this study.

814 **Supplementary Figures**

815 **A)**

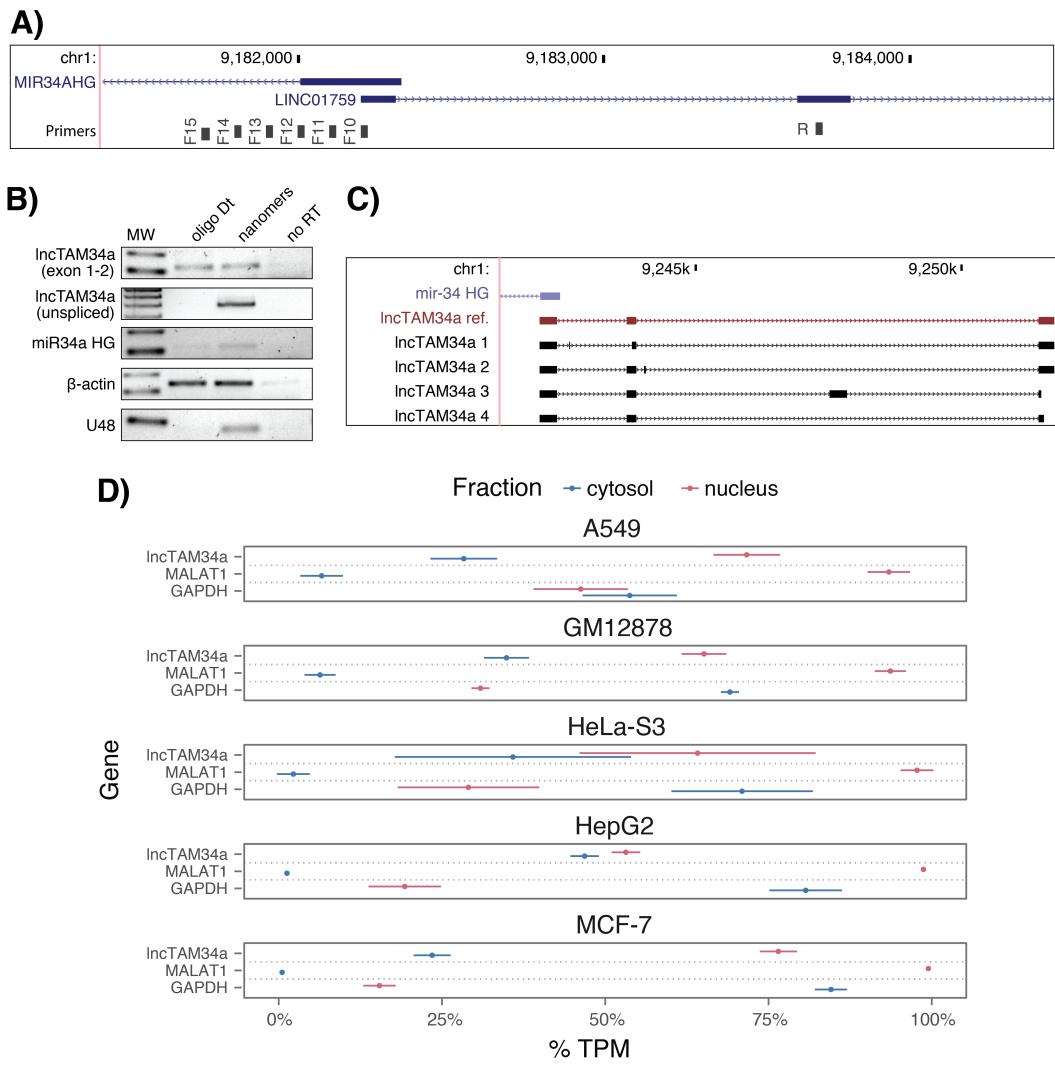
cancer	all n	all rho	all p	TP53wt n	TP53wt rho	TP53wt p	TP53mut n	TP53mut rho	TP53mut p
Adrenocortical carcinoma ( ACC )	10	0.55	1.04e-01	10	0.55	1.04e-01	NA	NA	NA
Bladder Urothelial Carcinoma ( BLCA )	228	0.51	7.89e-17	134	0.45	3.86e-08	94	0.43	1.73e-05
Breast invasive carcinoma ( BRCA ) Basal	42	0.57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

816 **B)**



817 **Figure 1 Supplement 1: TCGA normalized expression levels and correlation analysis statistics.**

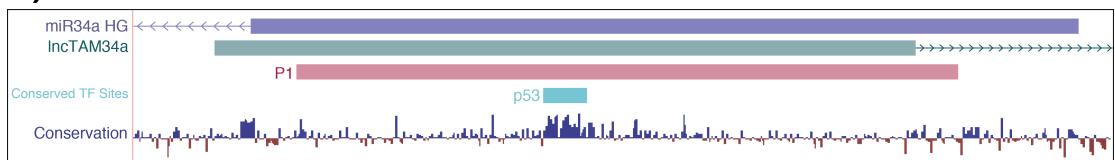
818 A) Spearman's rho and p-values (p) from the correlation analysis in Figure 1a between miR34a and  
819 IncTAM34a expression in TP53 wild type (wt) and mutated (mut) samples within TCGA cancer types.  
820 NA indicates not applicable, due to a lack of data for the specific group. B) Expression levels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



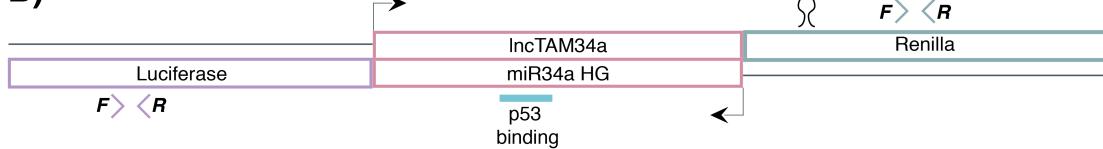
828  
829

830 **Figure 1 Supplement 2: Molecular characteristics of lncTAM34a.** A) A schematic representation of  
831 the primer placement in the primer walk assay. B) Polyadenylation status of spliced and  
832 unspliced lncTAM34a in HEK293T cells. C) Sequencing results from the analysis of lncTAM34a  
833 isoforms in U2OS cells. lncTAM34a ref. refers to the full-length transcript as defined by the 3'-RACE  
834 and primer walk assay. D) Analysis of coding potential of the lncTAM34a transcript using the Coding-  
835 potential Calculator. E) RNAseq data from five fractionated cell lines in the ENCODE project showing  
836 the percentage of transcripts per million (TPM) for lncTAM34a. MALAT1 (nuclear localization) and  
837 GAPDH (cytoplasmic localization) are included as fractionation controls. Points represent the mean  
838 and horizontal lines represent the standard deviation from two biological replicates.  
839

**A)**

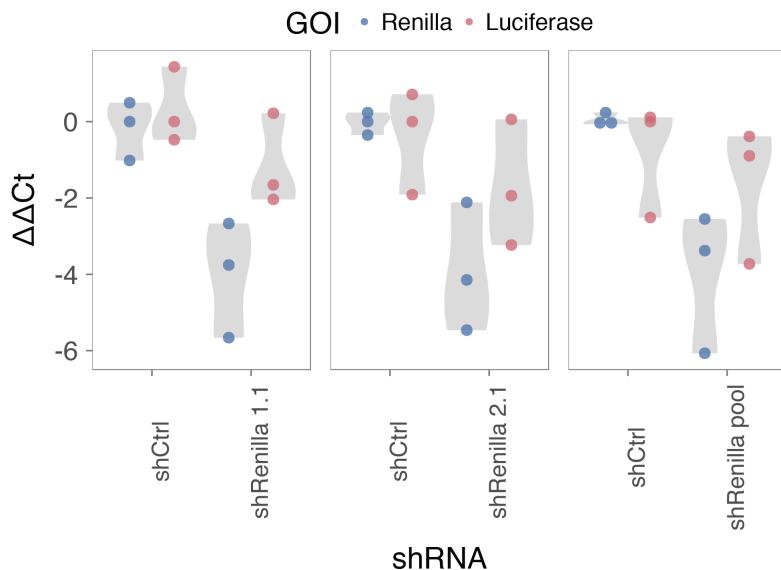


**B)**



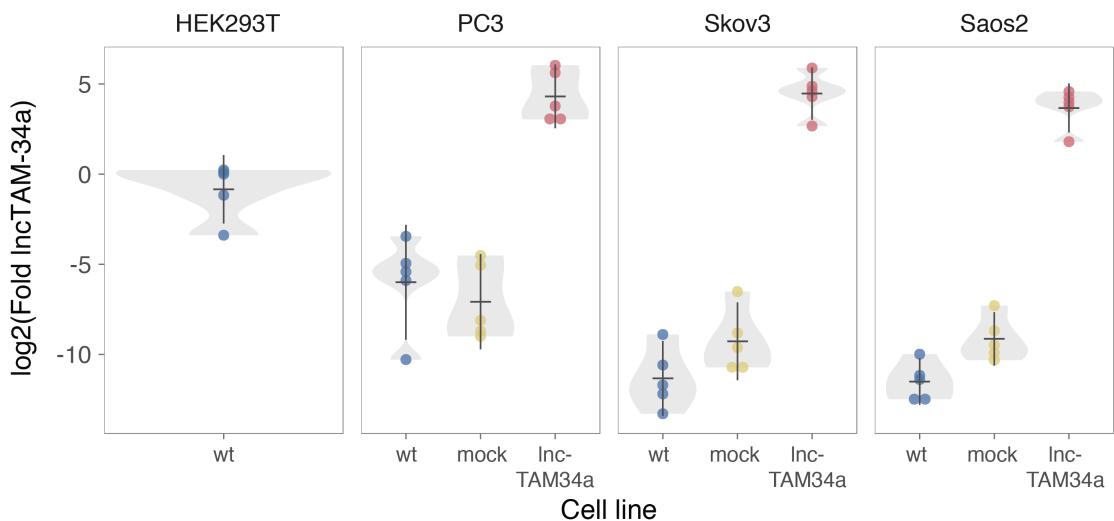
840  
841  
842  
843  
844  
845

**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.



846  
847  
848  
849  
850  
851  
852

**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.

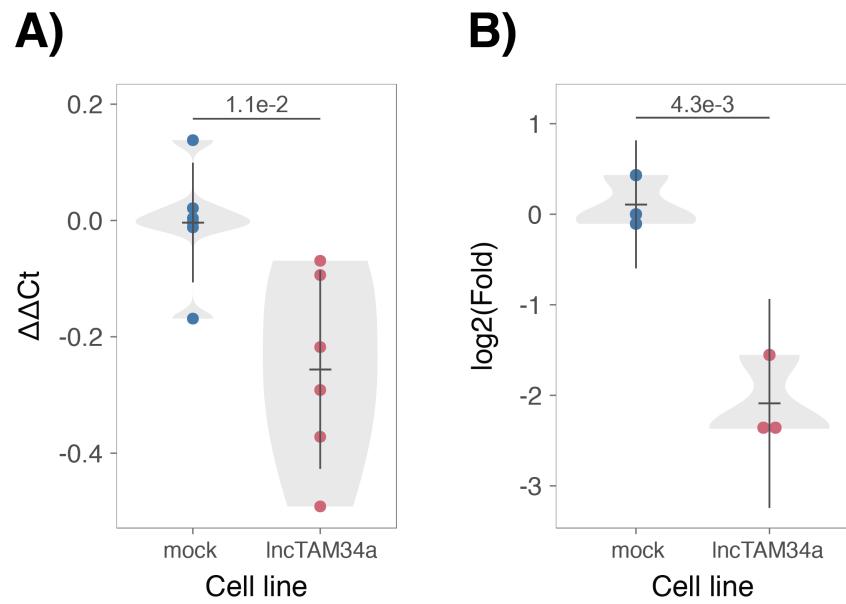


853  
854  
855  
856  
857

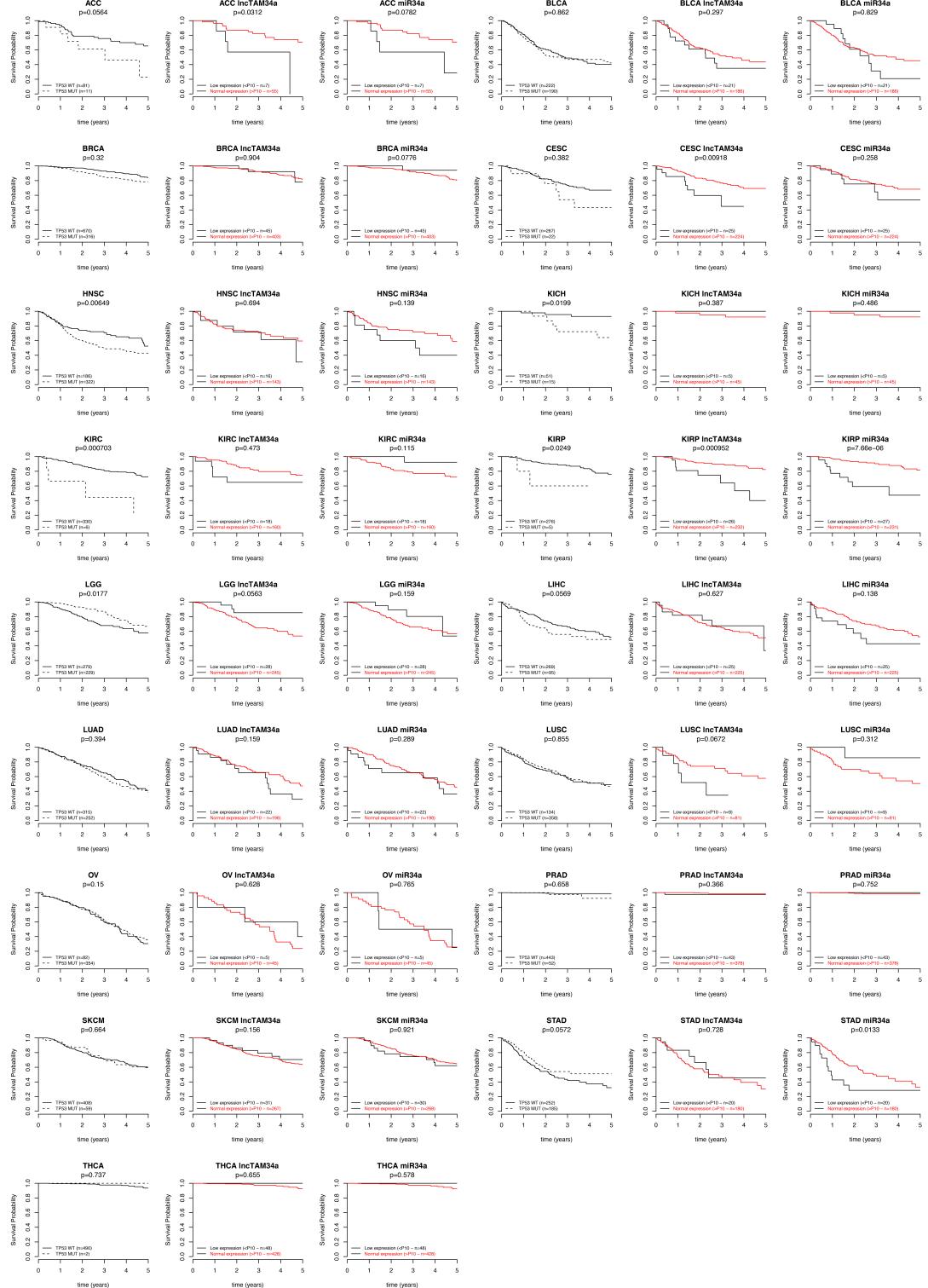
**Figure 3 Supplement 1: Physiological relevance of *lncTAM34a* overexpression.** Comparison of *lncTAM34a* expression in HEK293T cells (high endogenous *lncTAM34a*), and the wild-type (wt), mock, and *lncTAM34a* over-expressing stable cell lines.

858  
859

860  
861  
862



**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).



**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.

869

870 **References**

871

872 Agostini, M., P. Tucci, R. Killick, E. Candi, B. S. Sayan, P. Rivetti di Val Cervo, P.  
873 Nicotera, F. McKeon, R. A. Knight, T. W. Mak and G. Melino (2011). "Neuronal  
874 differentiation by TAp73 is mediated by microRNA-34a regulation of synaptic  
875 protein targets." *Proc Natl Acad Sci U S A* **108**(52): 21093-21098. DOI:  
876 10.1073/pnas.1112061109

877

878 Ahn, Y. H., D. L. Gibbons, D. Chakravarti, C. J. Creighton, Z. H. Rizvi, H. P. Adams, A.  
879 Pertsemlidis, P. A. Gregory, J. A. Wright, G. J. Goodall, E. R. Flores and J. M. Kurie  
880 (2012). "ZEB1 drives prometastatic actin cytoskeletal remodeling by  
881 downregulating miR-34a expression." *J Clin Invest* **122**(9): 3170-3183. DOI:  
882 10.1172/JCI63608

883

884 Allaire, J., Y. Xie, J. McPherson, J. Luraschi, K. Ushey, A. Atkins, H. Wickham, J.  
885 Cheng and W. Chang (2017). rmarkdown: Dynamic Documents for R. R package  
886 version 1.8. <https://CRAN.R-project.org/package=rmarkdown>

887

888 Amarzguioui, M., J. J. Rossi and D. Kim (2005). "Approaches for chemically  
889 synthesized siRNA and vector-mediated RNAi." *FEBS Lett* **579**(26): 5974-5981.  
890 DOI: 10.1016/j.febslet.2005.08.070

891

892 Arnold, J. B. (2017). ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'. R  
893 package version 3.4.0. <https://CRAN.R-project.org/package=ggthemes>

894

895 Ashouri, A., V. I. Sayin, J. Van den Eynden, S. X. Singh, T. Papagiannakopoulos and  
896 E. Larsson (2016). "Pan-cancer transcriptomic analysis associates long non-  
897 coding RNAs with key mutational driver events." *Nature Communications* **7**:  
898 13197. DOI: 10.1038/ncomms13197

899

900 Balbin, O. A., R. Malik, S. M. Dhanasekaran, J. R. Prensner, X. Cao, Y. M. Wu, D.  
901 Robinson, R. Wang, G. Chen, D. G. Beer, A. I. Nesvizhskii and A. M. Chinnaiyan  
902 (2015). "The landscape of antisense gene expression in human cancers." *Genome*  
903 *Res* **25**(7): 1068-1079. DOI: 10.1101/gr.180596.114

904

905 Boque-Sastre, R., M. Soler, C. Oliveira-Mateos, A. Portela, C. Moutinho, S. Sayols, A.  
906 Villanueva, M. Esteller and S. Guil (2015). "Head-to-head antisense transcription  
907 and R-loop formation promotes transcriptional activation." *Proc Natl Acad Sci U*  
908 *SA* **112**(18): 5785-5790. DOI: 10.1073/pnas.1421197112

909

910 Chang, T. C., D. Yu, Y. S. Lee, E. A. Wentzel, D. E. Arking, K. M. West, C. V. Dang, A.  
911 Thomas-Tikhonenko and J. T. Mendell (2008). "Widespread microRNA  
912 repression by Myc contributes to tumorigenesis." *Nat Genet* **40**(1): 43-50. DOI:  
913 10.1038/ng.2007.30

914

915 Chang, W. (2014). extrafont: Tools for using fonts. R package version 0.17.  
916 <https://CRAN.R-project.org/package=extrafont>

917

- 918 Chen, J., M. Sun, W. J. Kent, X. Huang, H. Xie, W. Wang, G. Zhou, R. Z. Shi and J. D.  
919 Rowley (2004). "Over 20% of human transcripts might form sense-antisense  
920 pairs." *Nucleic Acids Res* **32**(16): 4812-4820. DOI: 10.1093/nar/gkh818
- 921
- 922 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.  
923 S. Zheng (2010). "The impact of miR-34a on protein output in hepatocellular  
924 carcinoma HepG2 cells." *Proteomics* **10**(8): 1557-1572. DOI:  
925 10.1002/pmic.200900646
- 926
- 927 Chim, C. S., K. Y. Wong, Y. Qi, F. Loong, W. L. Lam, L. G. Wong, D. Y. Jin, J. F. Costello  
928 and R. Liang (2010). "Epigenetic inactivation of the miR-34a in hematological  
929 malignancies." *Carcinogenesis* **31**(4): 745-750. DOI: 10.1093/carcin/bgq033
- 930
- 931 Cole, K. A., E. F. Attiyeh, Y. P. Mosse, M. J. Laquaglia, S. J. Diskin, G. M. Brodeur and  
932 J. M. Maris (2008). "A functional screen identifies miR-34a as a candidate  
933 neuroblastoma tumor suppressor gene." *Mol Cancer Res* **6**(5): 735-742. DOI:  
934 10.1158/1541-7786.MCR-07-2102
- 935
- 936 Conley, A. B. and I. K. Jordan (2012). "Epigenetic regulation of human cis-natural  
937 antisense transcripts." *Nucleic Acids Res* **40**(4): 1438-1445. DOI:  
938 10.1093/nar/gkr1010
- 939
- 940 Consortium, E. P. (2012). "An integrated encyclopedia of DNA elements in the  
941 human genome." *Nature* **489**(7414): 57-74. DOI: 10.1038/nature11247
- 942
- 943 Ding, N., H. Wu, T. Tao and E. Peng (2017). "NEAT1 regulates cell proliferation  
944 and apoptosis of ovarian cancer by miR-34a-5p/BCL2." *Onco Targets Ther* **10**:  
945 4905-4915. DOI: 10.2147/OTT.S142446
- 946
- 947 Djebali, S., C. A. Davis, A. Merkel, A. Dobin, T. Lassmann, A. Mortazavi, A. Tanzer, J.  
948 Lagarde, W. Lin, F. Schlesinger, C. Xue, G. K. Marinov, J. Khatun, B. A. Williams, C.  
949 Zaleski, J. Rozowsky, M. Roder, F. Kokocinski, R. F. Abdelhamid, T. Alioto, I.  
950 Antoshechkin, M. T. Baer, N. S. Bar, P. Batut, K. Bell, I. Bell, S. Chakrabortty, X.  
951 Chen, J. Chrast, J. Curado, T. Derrien, J. Drenkow, E. Dumais, J. Dumais, R.  
952 Duttagupta, E. Falconnet, M. Fastuca, K. Fejes-Toth, P. Ferreira, S. Foissac, M. J.  
953 Fullwood, H. Gao, D. Gonzalez, A. Gordon, H. Gunawardena, C. Howald, S. Jha, R.  
954 Johnson, P. Kapranov, B. King, C. Kingswood, O. J. Luo, E. Park, K. Persaud, J. B.  
955 Preall, P. Ribeca, B. Risk, D. Robyr, M. Sammeth, L. Schaffer, L. H. See, A. Shahab, J.  
956 Skancke, A. M. Suzuki, H. Takahashi, H. Tilgner, D. Trout, N. Walters, H. Wang, J.  
957 Wrobel, Y. Yu, X. Ruan, Y. Hayashizaki, J. Harrow, M. Gerstein, T. Hubbard, A.  
958 Reymond, S. E. Antonarakis, G. Hannon, M. C. Giddings, Y. Ruan, B. Wold, P.  
959 Carninci, R. Guigo and T. R. Gingeras (2012). "Landscape of transcription in  
960 human cells." *Nature* **489**(7414): 101-108. DOI: 10.1038/nature11233
- 961
- 962 Gallardo, E., A. Navarro, N. Vinolas, R. M. Marrades, T. Diaz, B. Gel, A. Quera, E.  
963 Bandres, J. Garcia-Foncillas, J. Ramirez and M. Monzo (2009). "miR-34a as a  
964 prognostic marker of relapse in surgically resected non-small-cell lung cancer."  
965 *Carcinogenesis* **30**(11): 1903-1909. DOI: 10.1093/carcin/bgp219
- 966

- 967 Grossman, R. L., A. P. Heath, V. Ferretti, H. E. Varmus, D. R. Lowy, W. A. Kibbe and  
968 L. M. Staudt (2016). "Toward a Shared Vision for Cancer Genomic Data." N Engl J  
969 Med **375**(12): 1109-1112. DOI: 10.1056/NEJMmp1607591
- 970
- 971 Hunten, S., M. Kaller, F. Drepper, S. Oeljeklaus, T. Bonfert, F. Erhard, A. Dueck, N.  
972 Eichner, C. C. Friedel, G. Meister, R. Zimmer, B. Warscheid and H. Hermeking  
973 (2015). "p53-Regulated Networks of Protein, mRNA, miRNA, and lncRNA  
974 Expression Revealed by Integrated Pulsed Stable Isotope Labeling With Amino  
975 Acids in Cell Culture (pSILAC) and Next Generation Sequencing (NGS) Analyses."  
976 Mol Cell Proteomics **14**(10): 2609-2629. DOI: 10.1074/mcp.M115.050237
- 977
- 978 International Human Genome Sequencing, C. (2004). "Finishing the euchromatic  
979 sequence of the human genome." Nature **431**(7011): 931-945. DOI:  
980 10.1038/nature03001
- 981
- 982 Johnsson, P., A. Ackley, L. Vidarsdottir, W. O. Lui, M. Corcoran, D. Grander and K.  
983 V. Morris (2013). "A pseudogene long-noncoding-RNA network regulates PTEN  
984 transcription and translation in human cells." Nat Struct Mol Biol **20**(4): 440-  
985 446. DOI: 10.1038/nsmb.2516
- 986
- 987 Katayama, S., Y. Tomaru, T. Kasukawa, K. Waki, M. Nakanishi, M. Nakamura, H.  
988 Nishida, C. C. Yap, M. Suzuki, J. Kawai, H. Suzuki, P. Carninci, Y. Hayashizaki, C.  
989 Wells, M. Frith, T. Ravasi, K. C. Pang, J. Hallinan, J. Mattick, D. A. Hume, L. Lipovich,  
990 S. Batalov, P. G. Engstrom, Y. Mizuno, M. A. Faghihi, A. Sandelin, A. M. Chalk, S.  
991 Mottagui-Tabar, Z. Liang, B. Lenhard, C. Wahlestedt, R. G. E. R. Group, G. Genome  
992 Science and F. Consortium (2005). "Antisense transcription in the mammalian  
993 transcriptome." Science **309**(5740): 1564-1566. DOI: 10.1126/science.1112009
- 994
- 995 Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler and D.  
996 Haussler (2002). "The human genome browser at UCSC." Genome Res **12**(6):  
997 996-1006. DOI: 10.1101/gr.229102. Article published online before print in May  
998 2002
- 999
- 1000 Kim, K. H., H. J. Kim and T. R. Lee (2017). "Epidermal long non-coding RNAs are  
1001 regulated by ultraviolet irradiation." Gene **637**: 196-202. DOI:  
1002 10.1016/j.gene.2017.09.043
- 1003
- 1004 Kong, L., Y. Zhang, Z. Q. Ye, X. Q. Liu, S. Q. Zhao, L. Wei and G. Gao (2007). "CPC:  
1005 assess the protein-coding potential of transcripts using sequence features and  
1006 support vector machine." Nucleic Acids Res **35**(Web Server issue): W345-349.  
1007 DOI: 10.1093/nar/gkm391
- 1008
- 1009 Lal, A., M. P. Thomas, G. Altschuler, F. Navarro, E. O'Day, X. L. Li, C. Concepcion, Y.  
1010 C. Han, J. Thiery, D. K. Rajani, A. Deutsch, O. Hofmann, A. Ventura, W. Hide and J.  
1011 Lieberman (2011). "Capture of microRNA-bound mRNAs identifies the tumor  
1012 suppressor miR-34a as a regulator of growth factor signaling." PLoS Genet **7**(11):  
1013 e1002363. DOI: 10.1371/journal.pgen.1002363
- 1014

- 1015 Leveille, N., C. A. Melo, K. Rooijers, A. Diaz-Lagares, S. A. Melo, G. Korkmaz, R.  
1016 Lopes, F. Akbari Moqadam, A. R. Maia, P. J. Wijchers, G. Geeven, M. L. den Boer, R.  
1017 Kalluri, W. de Laat, M. Esteller and R. Agami (2015). "Genome-wide profiling of  
1018 p53-regulated enhancer RNAs uncovers a subset of enhancers controlled by a  
1019 lncRNA." *Nature Communications* **6**: 6520. DOI: 10.1038/ncomms7520
- 1020
- 1021 Liu, C., K. Kelnar, B. Liu, X. Chen, T. Calhoun-Davis, H. Li, L. Patrawala, H. Yan, C.  
1022 Jeter, S. Honorio, J. F. Wiggins, A. G. Bader, R. Fagin, D. Brown and D. G. Tang  
1023 (2011). "The microRNA miR-34a inhibits prostate cancer stem cells and  
1024 metastasis by directly repressing CD44." *Nat Med* **17**(2): 211-215. DOI:  
1025 10.1038/nm.2284
- 1026
- 1027 Memczak, S., M. Jens, A. Elefsinioti, F. Torti, J. Krueger, A. Rybak, L. Maier, S. D.  
1028 Mackowiak, L. H. Gregersen, M. Munschauer, A. Loewer, U. Ziebold, M.  
1029 Landthaler, C. Kocks, F. le Noble and N. Rajewsky (2013). "Circular RNAs are a  
1030 large class of animal RNAs with regulatory potency." *Nature* **495**(7441): 333-  
1031 338. DOI: 10.1038/nature11928
- 1032
- 1033 O'Leary, N. A., M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad, R. McVeigh, B.  
1034 Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y.  
1035 Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C. M.  
1036 Farrell, T. Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V. S. Joardar, V. K.  
1037 Kodali, W. Li, D. Maglott, P. Masterson, K. M. McGarvey, M. R. Murphy, K. O'Neill,  
1038 S. Pujar, S. H. Rangwala, D. Rausch, L. D. Riddick, C. Schoch, A. Shkeda, S. S. Storz,  
1039 H. Sun, F. Thibaud-Nissen, I. Tolstoy, R. E. Tully, A. R. Vatsan, C. Wallin, D. Webb,  
1040 W. Wu, M. J. Landrum, A. Kimchi, T. Tatusova, M. DiCuccio, P. Kitts, T. D. Murphy  
1041 and K. D. Pruitt (2016). "Reference sequence (RefSeq) database at NCBI: current  
1042 status, taxonomic expansion, and functional annotation." *Nucleic Acids Res*  
1043 **44**(D1): D733-745. DOI: 10.1093/nar/gkv1189
- 1044
- 1045 Ozsolak, F., P. Kapranov, S. Foissac, S. W. Kim, E. Fishilevich, A. P. Monaghan, B.  
1046 John and P. M. Milos (2010). "Comprehensive polyadenylation site maps in yeast  
1047 and human reveal pervasive alternative polyadenylation." *Cell* **143**(6): 1018-  
1048 1029. DOI: 10.1016/j.cell.2010.11.020
- 1049
- 1050 Polson, A., E. Durrett and D. Reisman (2011). "A bidirectional promoter reporter  
1051 vector for the analysis of the p53/WDR79 dual regulatory element." *Plasmid*  
1052 **66**(3): 169-179. DOI: 10.1016/j.plasmid.2011.08.004
- 1053
- 1054 Rashi-Elkeles, S., H. J. Warnatz, R. Elkon, A. Kupershtain, Y. Chobod, A. Paz, V.  
1055 Amstislavskiy, M. Sultan, H. Safer, W. Nietfeld, H. Lehrach, R. Shamir, M. L. Yaspo  
1056 and Y. Shiloh (2014). "Parallel profiling of the transcriptome, cistrome, and  
1057 epigenome in the cellular response to ionizing radiation." *Sci Signal* **7**(325): rs3.  
1058 DOI: 10.1126/scisignal.2005032
- 1059
- 1060 Raver-Shapira, N., E. Marciano, E. Meiri, Y. Spector, N. Rosenfeld, N. Moskovits, Z.  
1061 Bentwich and M. Oren (2007). "Transcriptional activation of miR-34a  
1062 contributes to p53-mediated apoptosis." *Mol Cell* **26**(5): 731-743. DOI:  
1063 10.1016/j.molcel.2007.05.017

1064  
1065 Rinn, J. L., M. Kertesz, J. K. Wang, S. L. Squazzo, X. Xu, S. A. Brugmann, L. H.  
1066 Goodnough, J. A. Helms, P. J. Farnham, E. Segal and H. Y. Chang (2007).  
1067 "Functional demarcation of active and silent chromatin domains in human HOX  
1068 loci by noncoding RNAs." *Cell* **129**(7): 1311-1323. DOI:  
1069 10.1016/j.cell.2007.05.022  
1070  
1071 Rokavec, M., M. G. Oner, H. Li, R. Jackstadt, L. Jiang, D. Lodygin, M. Kaller, D. Horst,  
1072 P. K. Ziegler, S. Schwitalla, J. Slotta-Huspenina, F. G. Bader, F. R. Greten and H.  
1073 Hermeking (2015). "Corrigendum. IL-6R/STAT3/miR-34a feedback loop  
1074 promotes EMT-mediated colorectal cancer invasion and metastasis." *J Clin Invest*  
1075 **125**(3): 1362. DOI: 10.1172/JCI81340  
1076  
1077 Schneider, C. A., W. S. Rasband and K. W. Eliceiri (2012). "NIH Image to ImageJ:  
1078 25 years of image analysis." *Nat Methods* **9**(7): 671-675.  
1079  
1080 Serviss, J. T. (2017). miR34AasRNAproject.  
1081 [https://github.com/GranderLab/miR34a\\_asRNA\\_project](https://github.com/GranderLab/miR34a_asRNA_project)  
1082  
1083 Serviss, J. T., P. Johnsson and D. Grander (2014). "An emerging role for long non-  
1084 coding RNAs in cancer metastasis." *Front Genet* **5**: 234. DOI:  
1085 10.3389/fgene.2014.00234  
1086  
1087 Slabakova, E., Z. Culig, J. Remsik and K. Soucek (2017). "Alternative mechanisms  
1088 of miR-34a regulation in cancer." *Cell Death Dis* **8**(10): e3100. DOI:  
1089 10.1038/cddis.2017.495  
1090  
1091 Stahlhut, C. and F. J. Slack (2015). "Combinatorial Action of MicroRNAs let-7 and  
1092 miR-34 Effectively Synergizes with Erlotinib to Suppress Non-small Cell Lung  
1093 Cancer Cell Proliferation." *Cell Cycle* **14**(13): 2171-2180. DOI:  
1094 10.1080/15384101.2014.1003008  
1095  
1096 Su, X., D. Chakravarti, M. S. Cho, L. Liu, Y. J. Gi, Y. L. Lin, M. L. Leung, A. El-Naggar,  
1097 C. J. Creighton, M. B. Suraokar, I. Wistuba and E. R. Flores (2010). "TAp63  
1098 suppresses metastasis through coordinate regulation of Dicer and miRNAs."  
1099 *Nature* **467**(7318): 986-990. DOI: 10.1038/nature09459  
1100  
1101 Sun, F., H. Fu, Q. Liu, Y. Tie, J. Zhu, R. Xing, Z. Sun and X. Zheng (2008).  
1102 "Downregulation of CCND1 and CDK6 by miR-34a induces cell cycle arrest."  
1103 *FEBS Lett* **582**(10): 1564-1568. DOI: 10.1016/j.febslet.2008.03.057  
1104  
1105 Tarasov, V., P. Jung, B. Verdoodt, D. Lodygin, A. Epanchintsev, A. Menssen, G.  
1106 Meister and H. Hermeking (2007). "Differential regulation of microRNAs by p53  
1107 revealed by massively parallel sequencing: miR-34a is a p53 target that induces  
1108 apoptosis and G1-arrest." *Cell Cycle* **6**(13): 1586-1593. DOI:  
1109 10.4161/cc.6.13.4436  
1110  
1111 Team, R. C. (2017). "R: A Language and Environment for Statistical Computing."  
1112 from <https://www.R-project.org/>.

- 1113  
1114 Therneau, T. M. (2015). A Package for Survival Analysis in S. version 2.38.  
1115 <https://CRAN.R-project.org/package=survival>  
1116  
1117 Turner, A. M., A. M. Ackley, M. A. Matrone and K. V. Morris (2012).  
1118 "Characterization of an HIV-targeted transcriptional gene-silencing RNA in  
1119 primary cells." *Hum Gene Ther* **23**(5): 473-483. DOI: 10.1089/hum.2011.165  
1120  
1121 Vogt, M., J. Munding, M. Gruner, S. T. Liffers, B. Verdoodt, J. Hauk, L.  
1122 Steinstraesser, A. Tannapfel and H. Hermeking (2011). "Frequent concomitant  
1123 inactivation of miR-34a and miR-34b/c by CpG methylation in colorectal,  
1124 pancreatic, mammary, ovarian, urothelial, and renal cell carcinomas and soft  
1125 tissue sarcomas." *Virchows Arch* **458**(3): 313-322. DOI: 10.1007/s00428-010-  
1126 1030-5  
1127  
1128 Wang, L., P. Bu, Y. Ai, T. Srinivasan, H. J. Chen, K. Xiang, S. M. Lipkin and X. Shen  
1129 (2016). "A long non-coding RNA targets microRNA miR-34a to regulate colon  
1130 cancer stem cell asymmetric division." *eLife* **5**. DOI: 10.7554/eLife.14620  
1131  
1132 Wang, L., H. J. Park, S. Dasari, S. Wang, J. P. Kocher and W. Li (2013). "CPAT:  
1133 Coding-Potential Assessment Tool using an alignment-free logistic regression  
1134 model." *Nucleic Acids Res* **41**(6): e74. DOI: 10.1093/nar/gkt006  
1135  
1136 Wang, X., J. Li, K. Dong, F. Lin, M. Long, Y. Ouyang, J. Wei, X. Chen, Y. Weng, T. He  
1137 and H. Zhang (2015). "Tumor suppressor miR-34a targets PD-L1 and functions  
1138 as a potential immunotherapeutic target in acute myeloid leukemia." *Cell Signal*  
1139 **27**(3): 443-452. DOI: 10.1016/j.cellsig.2014.12.003  
1140  
1141 Wickham, H. (2016). gtable: Arrange 'Grobs' in Tables. R package version 0.2.0.  
1142 <https://CRAN.R-project.org/package=gtable>  
1143  
1144 Wickham, H. (2017). scales: Scale Functions for Visualization. R package version  
1145 0.5.0. <https://CRAN.R-project.org/package=scales>  
1146  
1147 Wickham, H. (2017). tidyverse: Easily Install and Load the 'Tidyverse'. R package  
1148 version 1.2.1. <https://CRAN.R-project.org/package=tidyverse>  
1149  
1150 Wickham, L. H. a. H. (2017). rlang: Functions for Base Types and Core R and  
1151 'Tidyverse' Features. R package version 0.1.4. [https://CRAN.R-  
1152 project.org/package=rlang](https://CRAN.R-project.org/package=rlang)  
1153  
1154 Wickham, S. M. B. a. H. (2014). magrittr: A Forward-Pipe Operator for R. R  
1155 package version 1.5. <https://CRAN.R-project.org/package=magrittr>  
1156  
1157 Wilkins, D. gggenes: Draw Gene Arrow Maps in 'ggplot2'. R package version  
1158 0.2.0.9003. <https://github.com/wilkox/gggenes>  
1159  
1160 Xiao, N. (2017). liftr: Containerize R Markdown Documents. R package version  
1161 0.7. <https://CRAN.R-project.org/package=liftr>

- 1162  
1163 Xie, Y. (2017). knitr: A General-Purpose Package for Dynamic Report Generation  
1164 in R. R package version 1.17. <https://yihui.name/knitr/>  
1165  
1166 Yang, P., Q. J. Li, Y. Feng, Y. Zhang, G. J. Markowitz, S. Ning, Y. Deng, J. Zhao, S.  
1167 Jiang, Y. Yuan, H. Y. Wang, S. Q. Cheng, D. Xie and X. F. Wang (2012). "TGF-beta-  
1168 miR-34a-CCL22 signaling-induced Treg cell recruitment promotes venous  
1169 metastases of HBV-positive hepatocellular carcinoma." *Cancer Cell* **22**(3): 291-  
1170 303. DOI: 10.1016/j.ccr.2012.07.023  
1171  
1172 Yap, K. L., S. Li, A. M. Munoz-Cabello, S. Raguz, L. Zeng, S. Mujtaba, J. Gil, M. J.  
1173 Walsh and M. M. Zhou (2010). "Molecular interplay of the noncoding RNA ANRIL  
1174 and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional  
1175 silencing of INK4a." *Mol Cell* **38**(5): 662-674. DOI: 10.1016/j.molcel.2010.03.021  
1176  
1177 Yu, W., D. Gius, P. Onyango, K. Muldoon-Jacobs, J. Karp, A. P. Feinberg and H. Cui  
1178 (2008). "Epigenetic silencing of tumour suppressor gene p15 by its antisense  
1179 RNA." *Nature* **451**(7175): 202-206. DOI: 10.1038/nature06468  
1180  
1181 Zenz, T., J. Mohr, E. Eldering, A. P. Kater, A. Buhler, D. Kienle, D. Winkler, J. Durig,  
1182 M. H. van Oers, D. Mertens, H. Dohner and S. Stilgenbauer (2009). "miR-34a as  
1183 part of the resistance network in chronic lymphocytic leukemia." *Blood* **113**(16):  
1184 3801-3808. DOI: 10.1182/blood-2008-08-172254  
1185  
1186  
1187  
1188