

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

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

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

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

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

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

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

*** Correspondence:**

Jason T. Serviss 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 *miR34a* host gene, is critical for *miR34a* expression and mediation of its cellular functions in multiple types of human cancer. In addition, we characterize *miR34a* antisense RNA's ability to facilitate *miR34a* expression under different types of cellular stress in both *TP53* deficient and wildtype settings.

38 **Introduction**

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

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

60

61 Responses to cellular stress, e.g. DNA damage, sustained oncogene
62 expression, and nutrient deprivation, are all tightly controlled cellular pathways

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

85

86 Studies across multiple cancer types have reported a decrease in oncogenic
87 phenotypes when *miR34a* expression is induced in a *TP53*-null background,

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

105

106 **Results**

107

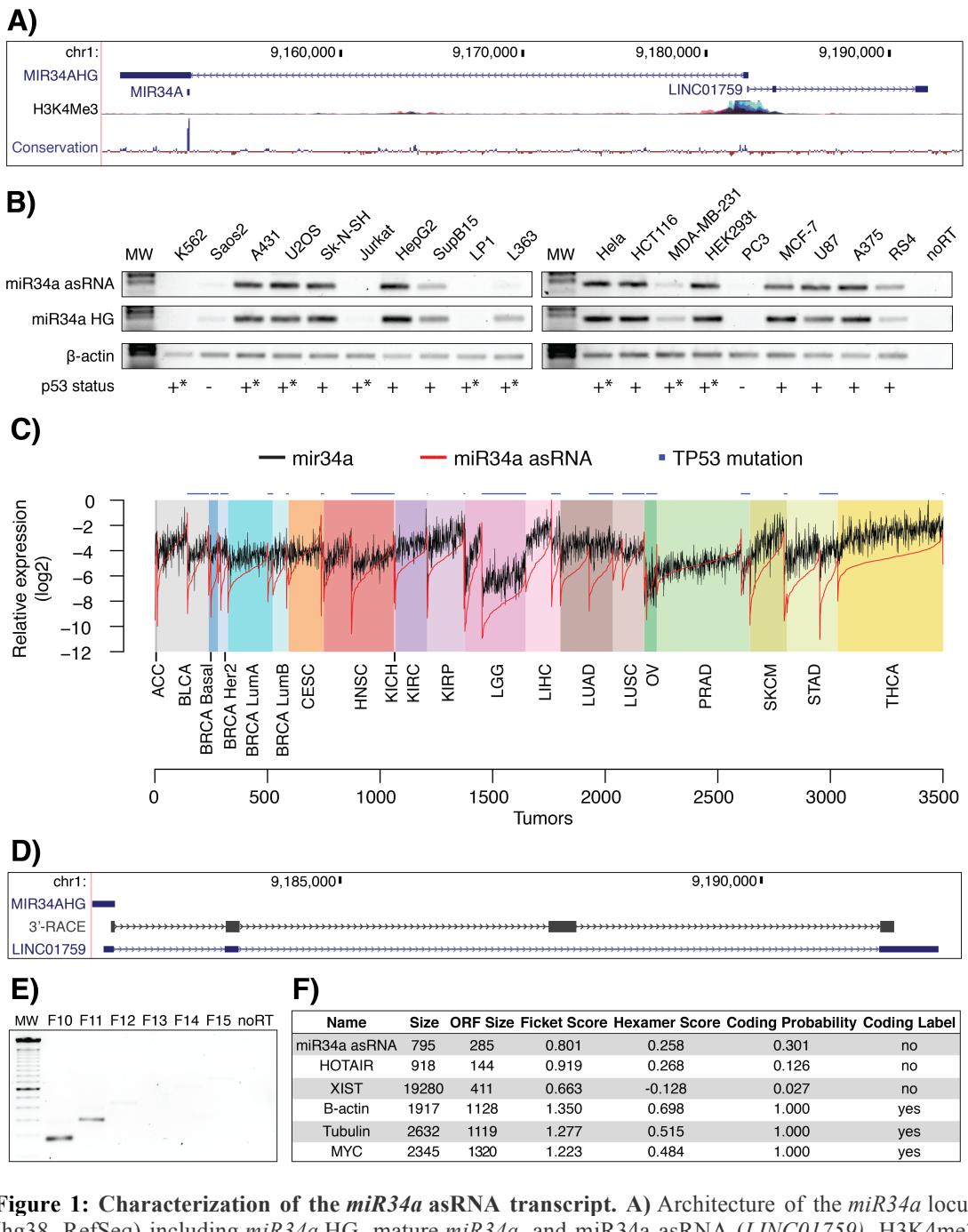
108 ***miR34a* asRNA is a broadly expressed, non-coding transcript whose**
109 **levels correlate with *miR34a* expression**

110

111 *miR34a* asRNA is transcribed in a “head-to-head” orientation with
112 approximately 100 base pair overlap with the *miR34a* host gene (HG) (**Fig.**
113 **1a**). Due to the fact that sense/antisense pairs can be both concordantly and
114 discordantly expressed, we sought to evaluate this relationship in the case of
115 *miR34a* HG and its asRNA. Using a diverse panel of cancer cell lines, we

116 detected co-expression of both the *miR34a* HG and *miR34a* asRNA (**Fig. 1b**).
117 We used cell lines with a known *TP53* status in the panel due to previous
118 reports that *miR34a* is a known downstream target of TP53. These results
119 indicate that *miR34a* HG and *miR34a* asRNA are co-expressed and that their
120 expression levels correlate with *TP53* status, with *TP53*^{-/-} cells tending to have
121 decreased or undetectable expression of both transcripts.

122



125 **Figure 1: Characterization of the *miR34a* asRNA transcript.** A) Architecture of the *miR34a* locus
126 (hg38, RefSeq) including *miR34a* HG, mature *miR34a*, and *miR34a* asRNA (*LINC01759*). H3K4me3
127 ChIP-seq data, indicating the active promoter region, and conservation are also shown. B) Semi-
128 quantitative PCR data from the screening of a panel of cancer cell lines. *Indicates either a non-null
129 *TP53* mutation or wild-type *TP53* with mechanisms present that inhibit *TP53* function (e.g. SV40 large
130 T antigen in HEK293T cells). C) TCGA correlation analysis. Expression was log2 normalized to the
131 maximum expression value. Nonsynonymous *TP53* mutations are indicated on the top of the plot
132 (cancer type abbreviation definitions and corresponding statistics are in Figure 1-Supplement 1). D) 3'-
133 RACE sequencing results and the annotated *miR34a* asRNA (*LINC01759*). E) Semi-quantitative PCR
134 results from the primer walk assay (i.e. forward primers, F10-F15, staggered upstream of the transcripts
135 annotated start site) performed using HEK293T cells (Figure 1 Supplement 2a details primer
136 placement) F) Coding potential analysis assessed via the Coding-potential Assessment Tool
137 including *miR34a* asRNA, two known non-coding RNAs (*HOTAIR* and *XIST*), and three protein-
138 coding RNAs (*β-actin*, *tubulin*, and *MYC*).

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

150

151 Next, we aimed to gain a thorough understanding of *miR34a* asRNA's
152 molecular characteristics and cellular localization. To experimentally
153 determine the 3' termination site for the *miR34a* asRNA transcript we
154 performed 3' rapid amplification of cDNA ends (RACE) using the U2OS
155 osteosarcoma cell line that exhibited high endogenous levels
156 of *miR34a* asRNA in the cell panel screening. Sequencing the cloned cDNA
157 indicated that the transcripts 3' transcription termination site is 525 base pairs
158 upstream of the *miR34a* asRNA's transcript's annotated termination site (**Fig.**
159 **1d**). Next, we characterized the *miR34a* asRNA 5' transcription start site by
160 carrying out a primer walk assay, i.e. a common reverse primer was placed in
161 exon 2 and forward primers were gradually staggered upstream of the
162 transcripts annotated start site (**Figure 1-Figure Supplement 2a**). Our results
163 indicated that the 5' start site for *miR34a* asRNA is in fact approximately 90bp

164 (F11 primer) to 220bp (F12 primer) upstream of the annotated start site (**Fig.**
165 **1e**). Polyadenylation status was evaluated via cDNA synthesis with either
166 random nanomers or oligoDT primers followed by semi-quantitative PCR
167 which indicated that the *miR34a* asRNA is polyadenylated although the
168 unspliced form seems to only be in a polyadenylation negative state (**Figure**
169 **1-Figure Supplement 2b**). Furthermore, we investigated the propensity
170 of *miR34a* asRNA to be alternatively spliced in U2OS cells, using PCR
171 cloning and sequencing and found that the transcript is post-transcriptionally
172 spliced to form multiple isoforms (**Figure 1-Figure Supplement 2c**). In order
173 to evaluate the subcellular localization of *miR34a* asRNA, we made use of
174 RNA sequencing data from five cancer cell lines included in the ENCODE
175 (Consortium 2012) project that had been fractionated into cytosolic and
176 nuclear fractions. The analysis revealed that the *miR34a* asRNA transcript
177 primarily localizes the nucleus with only a minor fraction in the cytosol (**Figure**
178 **1-Figure Supplement 2d**).

179
180 Lastly, we utilized several approaches to evaluate the coding potential of
181 the *miR34a* asRNA transcript. The Coding-Potential Assessment Tool is a
182 bioinformatics-based tool that uses a logistic regression model to evaluate
183 coding-potential by examining ORF length, ORF coverage, Fickett score and
184 hexamer score (Wang et al. 2013). Results indicated that *miR34a* asRNA has
185 a similar low coding capacity to known non-coding transcripts such as
186 *HOTAIR* and *XIST* (**Fig. 1F**). We further confirmed these results using the
187 Coding-Potential Calculator that uses a support based machine-based
188 classifier and accesses an alternate set of discriminatory features (**Figure 1-**

189 **Figure Supplement 2e)** (Kong et al. 2007). *** We hope to be able to scan
190 for peptides matching to miR34a asRNA in CPTAC and Geiger et al., 2012
191 before submission and will mention results here.***

192

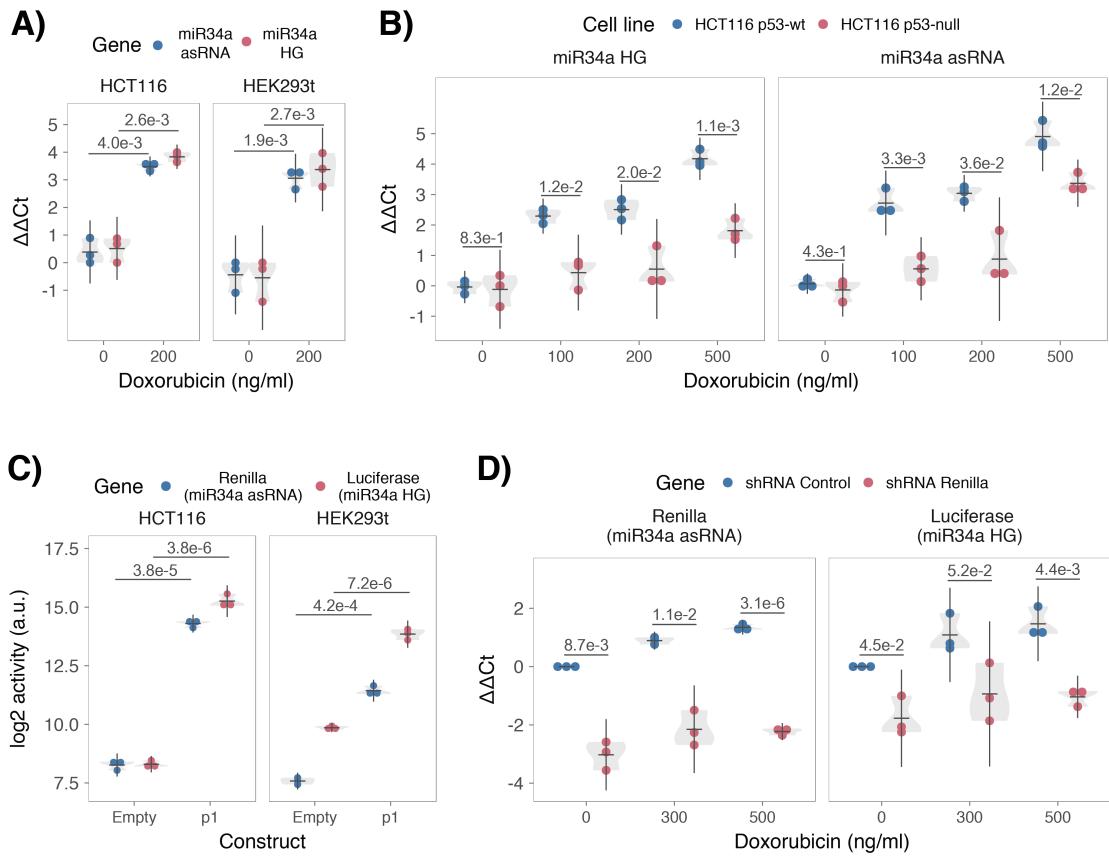
193 **TP53-mediated regulation of *miR34a* asRNA expression**

194 *miR34a* is a known downstream target of TP53 and has been previously
195 shown to exhibit increased expression within multiple contexts of cellular
196 stress. *miR34a* asRNA has also been shown to be induced upon TP53
197 activation in several global analyses of p53-regulated lncRNAs (Rashi-Elkeles
198 et al. 2014, Hunten et al. 2015, Leveille et al. 2015, Ashouri et al. 2016, Kim et
199 al. 2017). To confirm these results in our biological systems, we treated
200 HEK293T, embryonic kidney cells, and HCT116, colorectal cancer cells, with
201 the DNA damaging agent doxorubicin to activate TP53. QPCR-mediated
202 measurements of both *miR34a* HG and asRNA indicated that their expression
203 levels were increased in response to doxorubicin treatment in both cell lines
204 (**Fig. 2a**). To assess whether TP53 was responsible for the increase
205 in *miR34a* asRNA expression upon DNA damage, we
206 treated *TP53^{+/+}* and *TP53^{-/-}* HCT116 cells with increasing concentrations of
207 doxorubicin and monitored the expression of both *miR34a* HG and asRNA.
208 We observed a dose-dependent increase in both *miR34a* HG and asRNA
209 expression levels with increasing amounts of doxorubicin, indicating that
210 these two transcripts are co-regulated, although, this effect was largely
211 abrogated in *TP53^{-/-}* cells (**Fig. 2b**). These results indicate
212 that TP53 activation increases *miR34a* asRNA expression upon DNA
213 damage. Nevertheless, *TP53^{-/-}* cells also showed a dose-dependent increase

214 in both *miR34a* HG and asRNA, indicating that additional factors, other
215 than *TP53* are capable of initiating an increase in expression of both of these
216 transcripts upon DNA damage.

217

218



233

Figure 2: TP53-mediated regulation of the *miR34a* locus. **A)** Evaluating the effects of 24 hours of treatment with 200 ng/ml doxorubicin on *miR34a* asRNA and HG in HCT116 and HEK293T cells.* **B)** Monitoring *miR34a* HG and asRNA expression levels during 24 hours 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. Resulting p-values from statistical testing are shown above the shRenilla samples which were compared to the shRNA control using the respective treatment condition.* 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 biological triplicate.

234 The head-to-head orientation of *miR34a* HG and asRNA, suggests that
235 transcription is initiated from a single promoter in a bi-directional manner (**Fig**
236 **1a**). To investigate whether *miR34a* HG and asRNA are transcribed from the
237 same promoter as divergent transcripts, we cloned the previously reported
238 *miR34a* HG promoter, including the TP53 binding site, into a luciferase/renilla
239 dual reporter vector which we hereafter refer to as p1 (**Figure 2-Figure**
240 **Supplement 1a-b**) (Raver-Shapira et al. 2007). Upon transfection of p1 into
241 HCT116 and HEK293T cell lines we observed increases in both luciferase
242 and renilla indicating that *miR34a* HG and asRNA expression can be
243 regulated by a single promoter contained within the p1 construct (**Fig. 2c**).

244

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

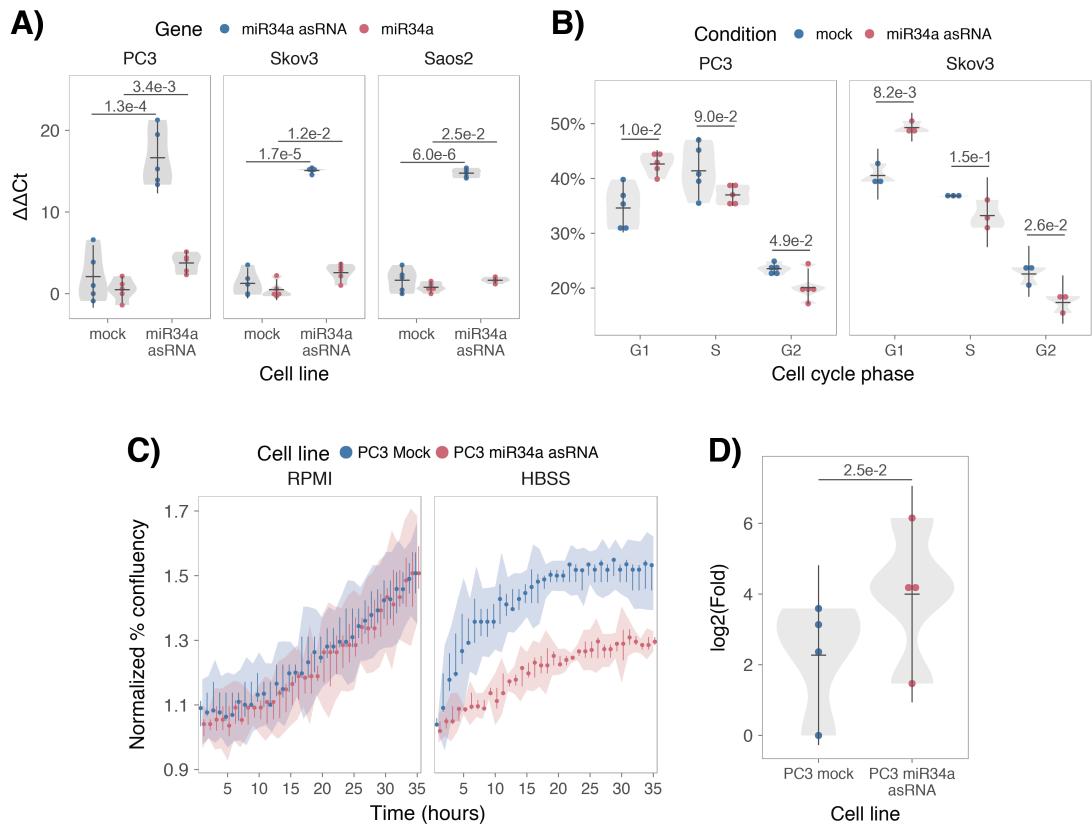
259 levels of *miR34a* HG and that the transcriptional product of the *miR34a*
260 asRNA within in the p1 construct contributes to inducing a miR34a response.
261 To further support these conclusions and better understand the role of
262 *miR34a* asRNA during TP53 activation, *TP53^{+/+}* HCT116 cells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

272

273 ***miR34a* asRNA can regulate *miR34a* host gene independently of *TP53***

274 Despite the fact that TP53 regulates *miR34a* HG and asRNA expression, our
275 results indicated that other factors are also able to regulate this locus (**Fig.**
276 **2b**). Utilizing a lentiviral system, we stably over-expressed the *miR34a* asRNA
277 transcript in three *TP53*-null cell lines, PC3 (prostate cancer), Saos2
278 (osteogenic sarcoma), and Skov3 (ovarian adenocarcinoma). We first
279 analyzed the levels of *miR34a* asRNA in these stable cell lines, compared to
280 HEK293T cells, which have high endogenous levels of *miR34a* asRNA. On
281 average, the over-expression was approximately 30-fold higher in the over-
282 expression cell lines than in HEK293T cells, roughly corresponding to
283 physiologically relevant levels in cells encountering a stress stimulus, such as

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



290

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

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

317
318 *miR34a* is also a well-known inhibitor of cellular growth via its ability to
319 negatively regulate growth factor signaling. Furthermore, starvation has been
320 shown to induce *miR34a* expression causing down-regulation of numerous
321 pro-survival growth factors (Lal et al. 2011). We further interrogated the
322 effects of *miR34a* asRNA over-expression by monitoring the growth of the
323 cells in both normal and starvation conditions via confluence measurements
324 over a 35-hour period. Under normal growth conditions there is a small but
325 significant reduction ($P = 3.0\text{e-}8$; linear regression, **Fig. 3c**) in confluence in
326 the *miR34a* asRNA over-expressing cell lines compared to mock control.
327 However, these effects on cell growth are drastically increased in starvation

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

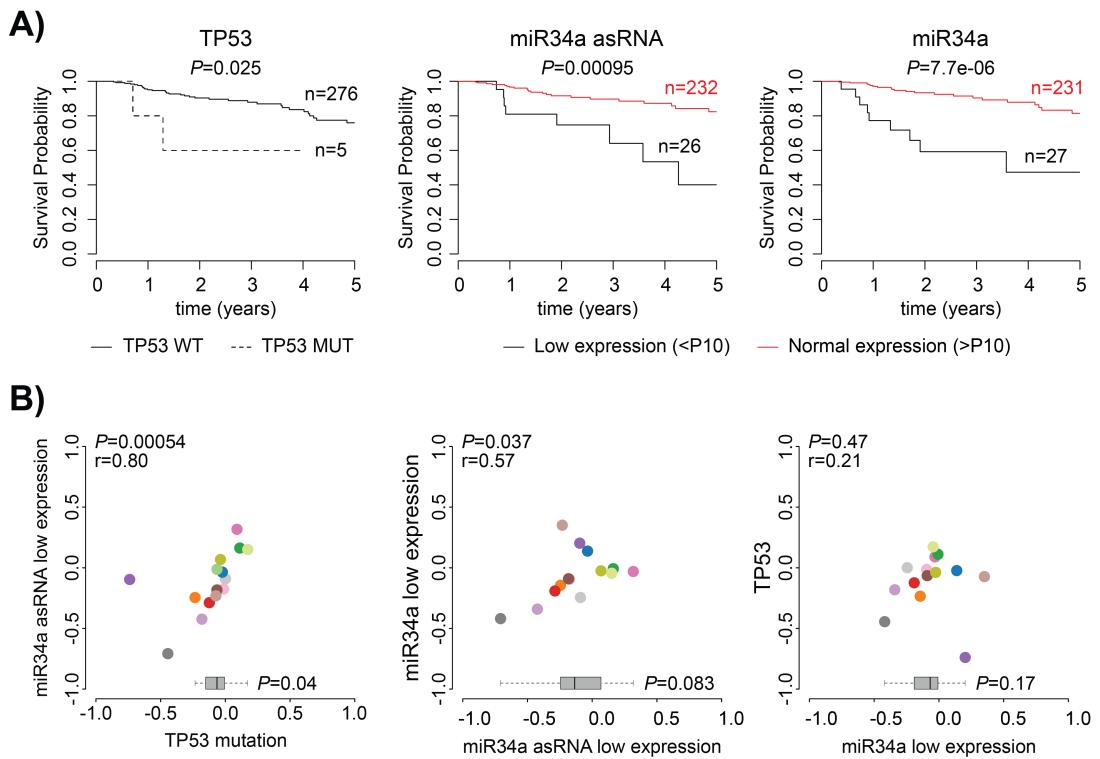
335

336 **miR34a asRNA transcriptionally activates miR34a HG**

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

349

350



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

352

353

354

355

356

357

358

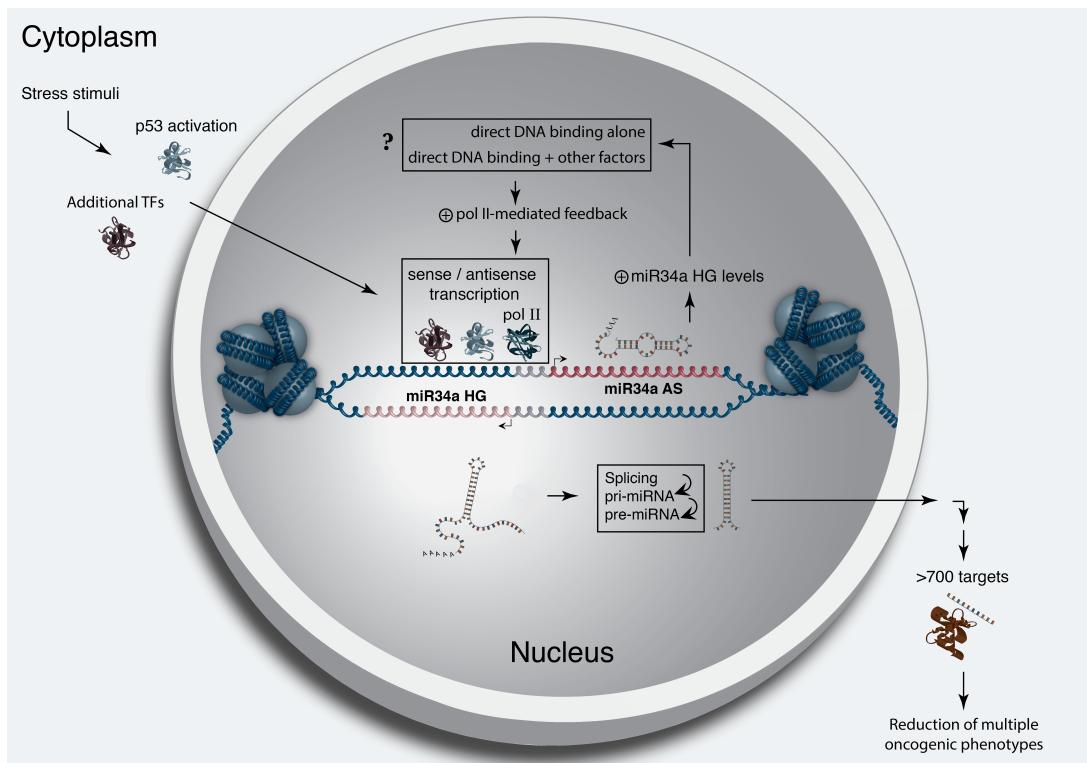
359

360

361
362 **Low miR34a asRNA expression levels are associated with decreased**
363 **survival**
364
365 As TP53 mutations and low expression of miR34a have been associated with
366 worse prognosis in cancer, we compared survival rates of samples with low
367 expression of miR34a asRNA (bottom 10th percentile) to control samples in
368 17 cancer types from TCGA (**Figure 4-Supplement 1**) (Gallardo et al. 2009,
369 Zenz et al. 2009, Liu et al. 2011). To correct for the effect of TP53 mutations
370 we focused on non-TP53 mutated samples, and noted a worse survival for the
371 low expression group in several cancers. This effect was most pronounced in
372 papillary kidney cancer (unadjusted $P=0.00095$; **Fig. 4a**). By systematically
373 comparing 5-year survival probabilities between the low expression group and
374 the control group for each cancer we found a median reduction of 5-year
375 survival probability of 9.6% ($P=0.083$; Wilcoxon signed rank test; **Fig. 4b**).
376 Furthermore, we found that miR34a asRNA expression showed similar
377 patterns in terms of direction and strength of association with 5-year survival
378 probability as miR34a expression ($r=0.57$, $P=0.037$) and TP53 mutations
379 ($r=0.80$, $P=0.00054$) across the different cancer types (**Fig. 4b**). Although
380 these results do not implicate any causal relationship, they do indicate a
381 striking similarity between the association of worse prognosis and TP53
382 mutations, low miR34a and low miR34a asRNA expression.

383

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



394 **Discussion**

395
396 Multiple studies have previously shown asRNAs to be crucial for the
397 appropriate regulation of cancer-associated protein-coding genes and that
398 their dysregulation can lead to a perturbation of tumor suppressive and
399 oncogenic pathways, as well as, cancer-related phenotypes (Yu et al. 2008,
400 Yap et al. 2010, Serviss et al. 2014, Balbin et al. 2015). Here we show that
401 asRNAs are also capable of regulating cancer-associated miRNAs resulting in
402 similar consequences as protein-coding gene dysregulation (**Fig. 4**).
403 Interestingly, we show that, both in the presence and absence of
404 *TP53*, *miR34a* asRNA provides an additional regulatory level to control
405 *miR34a* expression in both homeostasis and upon encountering various forms
406 of cellular stress. Furthermore, we find that a *miR34a* asRNA-mediated
407 increase in *miR34a* expression is sufficient to drive the appropriate cellular
408 responses to these stress stimuli (**Fig. 2d and Fig. 3c**). Previous studies have
409 exploited various molecular biology methods to up-regulate *miR34a*
410 expression in a *TP53*-deficient background showing similar phenotypic
411 outcomes although, here we show a novel mechanism by which this can be
412 achieved in an endogenous manner (Liu et al. 2011, Ahn et al. 2012, Yang et
413 al. 2012, Stahlhut et al. 2015, Wang et al. 2015).

414

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

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

442

443 An antisense transcript arising from the *miR34a* locus, *Lnc34a*, has been
444 previously reported to negatively regulate the expression of *miR34a* (Wang et

445 al. 2016). Although the *Lnc34a* and *miR34a* asRNA transcripts share some
446 sequence similarity, we believe them to be separate RNAs that are,
447 potentially, different isoforms of the same gene. We utilized CAGE and
448 RNAseq data from the ENCODE project to evaluate the presence of *miR34a*
449 and *Lnc34a* in 28 and 36 commonly used cancer cell lines, respectively.
450 Although the results show the presence of *miR34a* asRNA in these cell lines,
451 we find no evidence for *Lnc34a* transcription. These results are in line with the
452 findings of Wang et al. We thoroughly address our reasons for these beliefs
453 and give appropriate supporting evidence in (**Supplementary Document 1**).
454 The fact that *Lnc34a* and *miR34a* asRNA would appear to have opposing
455 roles in their regulation of *miR34a* further underlines the complexity of the
456 regulation at this locus.

457

458 Clinical trials utilizing *miR34a* replacement therapy have previously been
459 conducted but, disappointingly, were terminated after adverse side effects of
460 an immunological nature were observed in several of the patients (Slabakova
461 et al. 2017). Although it is not presently clear if these side effects were caused
462 by *miR34a* or the liposomal carrier used to deliver the miRNA, the multitude of
463 evidence indicating *miR34a*'s crucial role in oncogenesis still makes its
464 therapeutic induction an interesting strategy and needs further investigation.
465 In summary, our results have identified *miR34a* asRNA as a vital player in the
466 regulation of *miR34a* and its particular importance in typical examples of
467 cellular stress encountered in cancer. **We believe the conclusions drawn in
468 this study to be essential in the progress towards developing a better
469 understanding of the regulation of cancer-associated miRNAs and,**

470 specifically, the tumor suppressor *miR34a*.

471

472 Materials and Methods

473 Cell Culture

474 All cell lines were cultured at 5% CO₂ and 37°C with HEK293T, Saos2, and
475 Skov3 cells cultured in DMEM high glucose (GE Healthcare Life Sciences,
476 Hyclone, Amersham. UK, Cat# SH30081), HCT116 and U2OS cells in
477 McCoy's 5a (ThermoFisher Scientific, Pittsburgh, MA, USA. Cat# SH30200),
478 and PC3 cells in RPMI (GE Healthcare Life Sciences, Hyclone, Cat#
479 SH3009602) and 2 mM L-glutamine (GE Healthcare Life Sciences, Hyclone,
480 Cat# SH3003402). All growth mediums were supplemented with 10% heat-
481 inactivated FBS (ThermoFisher Scientific, Gibco, Cat# 12657029) and 50
482 µg/ml of streptomycin (ThermoFisher Scientific, Gibco, Cat# 15140122) and
483 50 µg/ml of penicillin (ThermoFisher Scientific, Gibco, Cat# 15140122). All cell
484 lines were purchased from ATCC, tested negative for mycoplasma, and their
485 identity was verified via STR profiling.

486

487 Bioinformatics, Data Availability, and Statistical Testing

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

491

492 All raw experimental data, code used for analysis, and supplementary
493 methods are available for review at ([Serviss 2017](#)) and are provided as an R
494 package. All analysis took place using the R statistical programming language
495 (Team 2017) using external packages that are documented in the package

496 associated with this article (Wilkins , Chang 2014, Wickham 2014, Therneau
497 2015, Wickham 2016, Allaire et al. 2017, Arnold 2017, Wickham 2017,
498 Wickham 2017, Wickham 2017, Xiao 2017, Xie 2017). The package facilitates
499 replication of the operating system and package versions used for the original
500 analysis, reproduction of each individual figure and figure supplement
501 included in the article, and easy review of the code used for all steps of the
502 analysis, from raw-data to figure.

503

504 The significance threshold (alpha) in this study was set to 0.05. Statistical
505 testing was performed using a unpaired two sample Student's t-test unless
506 otherwise specified.

507

508 **Coding Potential**

509 Protein-coding capacity was evaluated using the Coding-potential
510 Assessment Tool (Wang et al. 2013) and Coding-potential Calculator (Kong et
511 al. 2007) with default settings. Transcript sequences for use with Coding-
512 potential Assessment Tool were downloaded from the UCSC genome
513 browser using the Ensembl
514 accessions: *HOTAIR* (ENST00000455246), *XIST* (ENST00000429829), β-
515 actin (ENST00000331789), Tubulin (ENST00000427480),
516 and *MYC* (ENST00000377970). Transcript sequences for use with Coding-
517 potential Calculator were downloaded from the UCSC genome browser using
518 the following IDs: *HOTAIR* (uc031qho.1), β-actin (uc003soq.4).

519

520 **shRNAs**

521 shRNA-expressing constructs were cloned into the U6M2 construct using the
522 BgIII and KpnI restriction sites as previously described (Amarzguioui et al.
523 2005). shRNA constructs were transfected using Lipofectamine 2000 or 3000
524 (ThermoFisher Scientific, Cat# 12566014 and L3000015). The sequences
525 targeting renilla is as follows: shRenilla 1.1 (AAT ACA CCG CGC TAC TGG
526 C), shRenilla 2.1 (TAA CGG GAT TTC ACG AGG C).

527

528 **Bi-directional Promoter Cloning**

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

533

534 **Promoter Activity**

535 Cells were co-transfected with the renilla/firefly bidirectional promoter
536 construct (Polson et al. 2011) and GFP by using Lipofectamine 2000 (Life
537 Technologies, Cat# 12566014). The expression of GFP and luminescence
538 was measured 24 h post transfection by using the Dual-Glo Luciferase Assay
539 System (Promega, Cat# E2920) and detected by the GloMax-Multi+ Detection
540 System (Promega, Cat# SA3030). The expression of luminescence was
541 normalized to GFP.

542

543 **Generation of U6-expressed miR34a AS Lentiviral Constructs**

544 The U6 promoter was amplified from the U6M2 cloning plasmid (Amarzguioui
545 et al. 2005) and ligated into the Not1 restriction site of the pHIV7-IMPDH2

546 vector (Turner et al. 2012). *miR43a* asRNA was PCR amplified and
547 subsequently cloned into the Nhe1 and Pac1 restriction sites in the pHIV7-
548 IMPDH2-U6 plasmid.

549

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

551 Lentivirus production was performed as previously described in (Turner et al.
552 2012). Briefly, HEK293T cells were transfected with viral and expression
553 constructs using Lipofectamine 2000 (ThermoFisher Scientific, Cat#
554 12566014), after which viral supernatants were harvested 48 and 72 hours
555 post-transfection. Viral particles were concentrated using PEG-IT solution
556 (Systems Biosciences, Palo Alto, CA, USA. Cat# LV825A-1) according to the
557 manufacturer's recommendations. HEK293T cells were used for virus titration
558 and GFP expression was evaluated 72hrs post-infection via flow cytometry
559 (LSRII, BD Biosciences, San Jose, CA, USA) after which TU/ml was
560 calculated.

561

562 Stable lines were generated by infecting cells with a multiplicity of infection of
563 1 after which 1-2 µM mycophenolic acid (Merck, Kenilworth, NJ, USA. Cat#
564 M5255) selection was initiated 48 hours post-infection. Cells were expanded
565 as the selection process was monitored via flow cytometry analysis (LSRII,
566 BD Biosciences) of GFP and selection was terminated once > 90% of the
567 cells were GFP positive. Quantification of *miR34a* asRNA over-expression
568 and *miR34a* was performed in biological quintuplet for all cell lines.

569

570 **Western Blotting**

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

585

586 **RNA Extraction and cDNA Synthesis**

587 For downstream SYBR green applications, RNA was extracted using the
588 RNeasy mini kit (Qiagen, Venlo, Netherlands, Cat# 74106) and subsequently
589 treated with DNase (Ambion Turbo DNA-free, ThermoFisher Scientific, Cat#
590 AM1907). 500ng RNA was used for cDNA synthesis using MuMLV
591 (ThermoFisher Scientific, Cat# 28025013) and a 1:1 mix of oligo(dT) and
592 random nanomers.

593

594 For analysis of miRNA expression with Taqman, samples were isolated with
595 TRIzol reagent (ThermoFisher Scientific, Cat# 15596018) and further

596 processed with the miRNeasy kit (Qiagen, Cat# 74106). cDNA synthesis was
597 performed using the TaqMan MicroRNA Reverse Transcription Kit
598 (ThermoFisher Scientific, Cat# 4366597) using the corresponding oligos
599 according to the manufacturer's recommendations.

600

601 **QPCR and PCR**

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

607

608 QPCR for miRNA expression analysis was performed according to the primer
609 probe set manufacturers recommendations (ThermoFisher Scientific) and
610 using the TaqMan Universal PCR Master Mix (ThermoFisher Scientific, Cat#
611 4304437) with the same cycling scheme as above. Primer and probe sets for
612 TaqMan were also purchased from ThermoFisher Scientific (Life
613 Technologies at time of purchase, TaqMan® MicroRNA Assay, hsa-miR-34a,
614 human, Cat# 4440887, Assay ID: 000426 and Control miRNA Assay, RNU48,
615 human, Cat# 4440887, Assay ID: 001006).

616

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

621

622 **Cell Cycle Distribution**

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

632

633 **3' Rapid Amplification of cDNA Ends**

634 3'-RACE was performed as described as previously in (Johnsson et al. 2013).
635 Briefly, U2OS cell RNA was polyA-tailed using yeast polyA polymerase
636 (ThermoFisher Scientific, Cat# 74225Z25KU) after which cDNA was
637 synthesized using oligo(dT) primers. Nested-PCR was performed first using a
638 forward primer in *miR34a* asRNA exon 1 and a tailed oligo(dT) primer
639 followed by a second PCR using an alternate *miR34a* asRNA exon 1 primer
640 and a reverse primer binding to the tail of the previously used oligo(dT)
641 primer. PCR products were gel purified and cloned the Strata Clone Kit
642 (Agilent Technologies, Santa Clara, CA, USA. Cat# 240205), and sequenced.

643

644 **Chromatin Immunoprecipitation**

645 The ChIP was performed as previously described in (Johnsson et al. 2013)

646 with the following modifications. Cells were crosslinked in 1% formaldehyde
647 (Merck, Cat# 1040039025), quenched with 0.125M glycine (Sigma Aldrich,
648 Cat# G7126), and lysed in cell lysis buffer comprised of: 5mM PIPES (Sigma
649 Aldrich, Cat# 80635), 85mM KCL (Merck, Cat# 4936), 0.5% NP40 (Sigma
650 Aldrich, Cat# I8896), protease inhibitor (Roche Diagnostics, Cat#
651 004693159001). Samples were then sonicated in 50mM TRIS-HCL pH 8.0
652 (Sigma Aldrich, MO, USA, Cat# T2663) 10mM EDTA (Promega, WI, USA,
653 Cat# V4231), 1% SDS (ThermoFisher Scientific, Cat# AM9822), and protease
654 inhibitor (Roche Diagnostics, Cat# 004693159001) using a Bioruptor
655 Sonicator (Diagenode, Denville, NJ, USA). Samples were incubated over
656 night at 4°C with the polII antibody (Abcam, Cat# ab5095) and subsequently
657 pulled down with Salmon Sperm DNA/Protein A Agarose (Millipore, Cat# 16-
658 157) beads. DNA was eluted in an elution buffer of 1% SDS (ThermoFisher
659 Scientific, Cat# AM9822) 100mM NaHCO3 (Sigma Aldrich, Cat# 71631),
660 followed by reverse crosslinking, RNaseA (ThermoFisher Scientific, Cat#
661 1692412) and protease K (New England Biolabs, Ipswich, MA, USA, Cat#
662 P8107S) treatment. The DNA was eluted using Qiagen PCR purification kit
663 (Cat# 28106) and quantified via QPCR. QPCR was performed in technical
664 duplicate using the standard curve method and reported absolute values. The
665 fraction of input was subsequently calculated using the mean of the technical
666 replicates followed by calculating the fold over the control condition. Statistical
667 testing was performed using 4 biological replicates with the null hypothesis
668 that the true log 2 fold change values were equal to zero.

669

670 **Confluency Analysis**

671 Cells were incubated in the Spark Multimode Microplate (Tecan, Männedorf,
672 Switzerland) reader for 48 hours at 37°C with 5% CO₂ in a humidity chamber.
673 Confluency was measured every hour using bright-field microscopy and the
674 percentage of confluence was reported via the plate reader's inbuilt algorithm.
675 Percentage of confluence was normalized to the control sample in each
676 condition (shown in figure) and then ranked to move the data to a linear scale.
677 Using the mean of the technical duplicates in three biological replicates, the
678 rank was then used to construct a linear model, of the dependency of the rank
679 on the time and cell lines variables for each growth condition. Reported p-
680 values are derived from the t-test, testing the null hypothesis that the
681 coefficient estimate of the cell line variable is equal to 0.

682

683 **Pharmacological Compounds**

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

685

686 **Cellular Localization Analysis**

687 Quantified RNAseq data from 11 cell lines from the GRCh38 assembly was
688 downloaded from the ENCODE project database and quantifications for
689 *miR34a* asRNA (ENSG00000234546), GAPDH (ENSG00000111640), and
690 MALAT1 (ENSG00000251562) were extracted. Cell lines for which data was
691 downloaded include: A549, GM12878, HeLa-S3, HepG2, HT1080, K562
692 MCF-7, NCI-H460, SK-MEL-5, SK-N-DZ, SK-N-SH. Initial exploratory analysis
693 revealed that several cell lines should be removed from the analysis due to a)
694 a larger proportion of GAPDH in the nucleus than cytoplasm or b) variation of
695 *miR34a* asRNA expression is too large to draw conclusions, or c) they have

696 no or low (<6 TPM) *miR34a* asRNA expression. Furthermore, only
697 polyadenylated libraries were used in the final analysis, due to the fact that
698 the cellular compartment enrichment was improved in these samples. All
699 analyzed genes are reported to be polyadenylated. In addition, only samples
700 with 2 biological replicates were retained. For each cell type, gene, and
701 biological replicate the fraction of transcripts per million (TPM) in each cellular
702 compartment was calculated as the fraction of TPM in the specific
703 compartment by the total TPM. The mean and standard deviation for the
704 fraction was subsequently calculated for each cell type and cellular
705 compartment and this information was represented in the final figure.

706

707 **CAGE Analysis**

708 All available CAGE data from the ENCODE project (Consortium 2012) for 36
709 cell lines was downloaded from the UCSC genome browser (Kent et al. 2002)
710 for genome version hg19. Of these, 28 cell lines had CAGE transcription start
711 sites (TSS) mapping to the plus strand of chromosome 1 and in regions
712 corresponding to 200 base pairs upstream of the *lnc34a* start site (9241796 -
713 200) and 200 base pairs upstream of the GENCODE
714 annotated *miR34a* asRNA start site (9242263 + 200). These cell lines
715 included: HFDPC, H1-hESC, HMEpC, HAoEC, HPIEpC, HSaVEC, GM12878,
716 hMSC-BM, HUVEC, AG04450, hMSC-UC, IMR90, NHDF, SK-N-SH_RA, BJ,
717 HOB, HPC-PL, HAoAF, NHEK, HVMF, HWP, MCF-7, HepG2, hMSC-AT,
718 NHEM.f_M2, SkMC, NHEM_M2, and HCH. In total 74 samples were included.
719 17 samples were polyA-, 47 samples were polyA+, and 10 samples were total
720 RNA. In addition, 34 samples were whole cell, 15 enriched for the cytosolic

721 fraction, 15 enriched for the nucleolus, and 15 enriched for the nucleus. All
722 CAGE transcription start sites were plotted and the RPKM of the individual
723 reads was used to color each read to indicate their relative abundance. In
724 cases where CAGE TSS spanned identical regions, the RPMKs of the regions
725 were summed and represented as one CAGE TSS in the figure. In addition, a
726 density plot shows the distribution of the CAGE reads in the specified
727 interval.

728

729 **Splice Junction Analysis**

730 All available whole cell (i.e. non-fractionated) spliced read data originating
731 from the Cold Spring Harbor Lab in the ENCODE project (Consortium 2012)
732 for 38 cell lines was downloaded from the UCSC genome browser (Kent et al.
733 2002). Of these cell lines, 36 had spliced reads mapping to the plus strand of
734 chromosome 1 and in the region between the *lnc34a* start (9241796) and
735 transcription termination (9257102) site (note that *miR34a* asRNA resides
736 totally within this region). Splice junctions from the following cell lines were
737 included in the final figure: A549, Ag04450, Bj, CD20, CD34 mobilized,
738 Gm12878, H1hesc, Haoaf, Haoec, Hch, Helas3, Hepg2, Hfdpc, Hmec,
739 Hmepc, Hmescat, Hmscbm, Hmscuc, Hob, Hpcpl, Hpiepc, Hsavec, Hsmm,
740 Huvec, Hvmf, Hwp, Imr90, Mcf7, Monocd14, Nhdf, Nhek, Nhemfm2,
741 Nhemm2, Nhlf, Skmc, and Sknsh. All splice junctions were included in the
742 figure and colored according to the number of reads corresponding to each. In
743 cases where identical reads were detected multiple times, the read count was
744 summed and represented as one read in the figure.

745

746 **TCGA Data Analysis**

747 RNA-Seq data and copy number data were downloaded from TCGA and
748 processed as described previously (Ashouri et al. 2016). Briefly, RNA-Seq
749 data were aligned to the human hg19 assembly and quantified using
750 GENCODE (v19) annotated HTSeq-counts and FPKM normalizations.
751 Expression data from miR34a and miR34 asRNA (identified as RP3-
752 510D11.2) were used for further analysis. Copy number amplitudes for
753 GENCODE genes were determined from segmented copy-number data.
754 Samples that were diploid for miR34 asRNA were identified as those samples
755 that had copy number amplitudes between -0.1 and 0.1.

756

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

760

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

764

765 **Acknowledgments**

766 The authors would like to kindly thank Martin Enge for his critical review of the
767 manuscript and fruitful discussions. Mattias(?)

768

769 **Competing Interests**

770
771 The authors declare no competing interests.

772

773

774 **Funding**

775
776 This work has been supported by the Swedish Research Council [521-2012-
777 2037], Swedish Cancer Society [150768], Cancer Research Foundations of
778 Radiumhemmet [144063] and the Swedish Childhood Cancer Foundation
779 [PR2015-0009].

780
781

782 **Figure Supplements**

783

784 Figure 1-Supplement 1: TCAG expression levels and correlation analysis
785 statistics.

786

787 Figure 1-Supplement 2: Molecular characteristics of miR34a asRNA.

788

789 Figure 2-Supplement 1: A schematic representation of the p1 construct.

790

791 Figure 2-Supplement 2: Evaluating the effects of miR34a asRNA down-
792 regulation.

793

794 Figure 3-Supplement 1: Physiological relevance of miR34a asRNA
795 overexpression.

796

797 Figure 3-Supplement 2: Effects of miR34a asRNA overexpression on cyclin
798 D1.

799

800 Figure 4-Supplement 1: Survival analysis in 17 cancers from TCGA.

801

802 Supplementary Document 1: Evaluating the relationship between miR34a
803 asRNA and lnc34a.

804

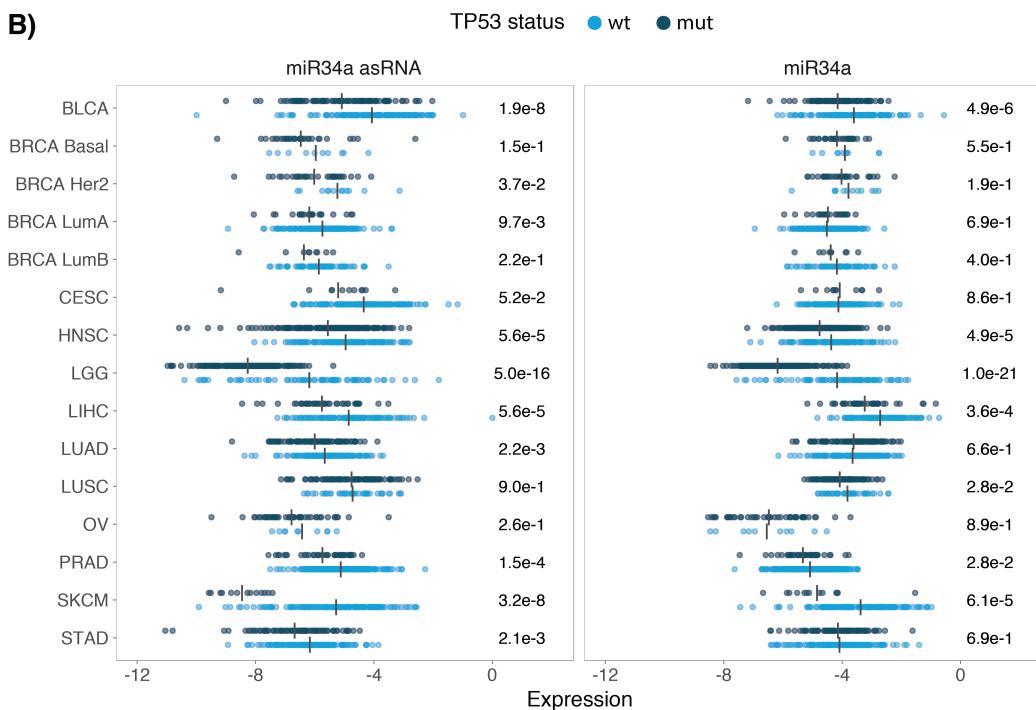
805 Supplementary Document 2: A table of primers used in this study.

806 **Supplementary Figures**

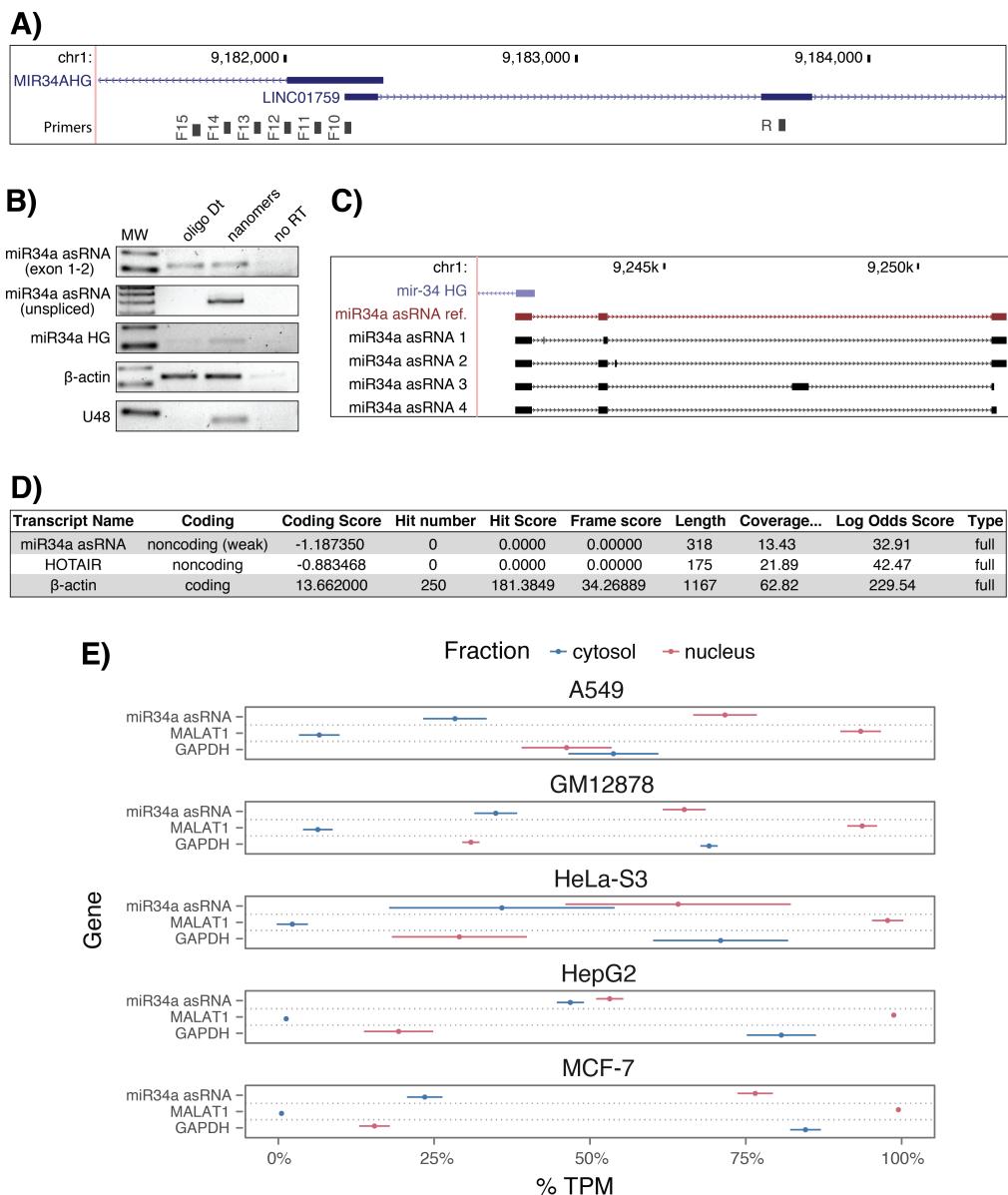
807

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)808
809

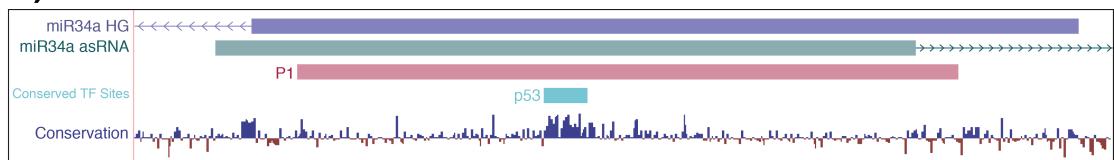
810 **Figure 1 Supplement 1: TCGA normalized expression levels and correlation analysis statistics.**
811 A) Spearman's rho and p-values (p) from the correlation analysis in Figure 1a between miR34a and
812 miR34a asRNA expression in TP53 wild type (wt) and mutated (mut) samples within TCGA cancer
813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



820
821
822
823
824
825
826
827
828
829
830
831
832

Figure 1 Supplement 2: Molecular characteristics of miR34a asRNA. A) A schematic representation of the primer placement in the primer walk assay. B) Polyadenylation status of spliced and unspliced miR34a asRNA in HEK293T cells. C) Sequencing results from the analysis of miR34a asRNA isoforms in U2OS cells. miR34a AS ref. refers to the full-length transcript as defined by the 3'-RACE and primer walk assay. D) Analysis of coding potential of the miR34a asRNA transcript using the Coding-potential Calculator. E) RNAseq data from five fractionated cell lines in the ENCODE project showing the percentage of transcripts per million (TPM) for miR34a asRNA. MALAT1 (nuclear localization) and GAPDH (cytoplasmic localization) are included as fractionation controls. Points represent the mean and horizontal lines represent the standard deviation from two biological replicates.

A)



B)



833
834
835
836
837
838

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.

839
840
841
842
843
844
845

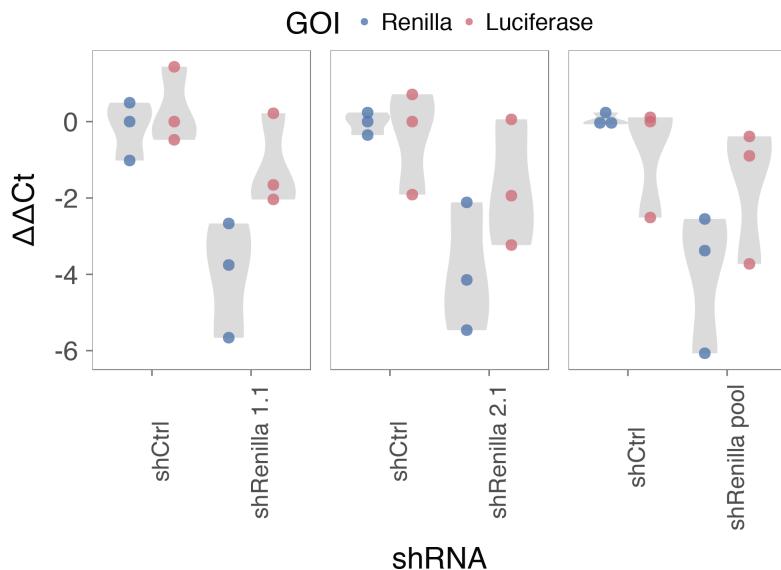
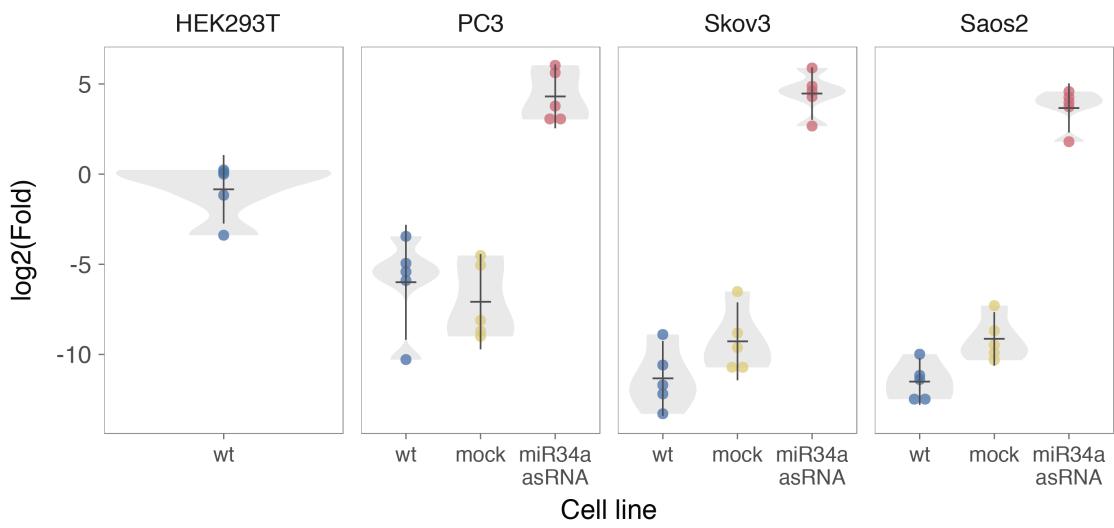
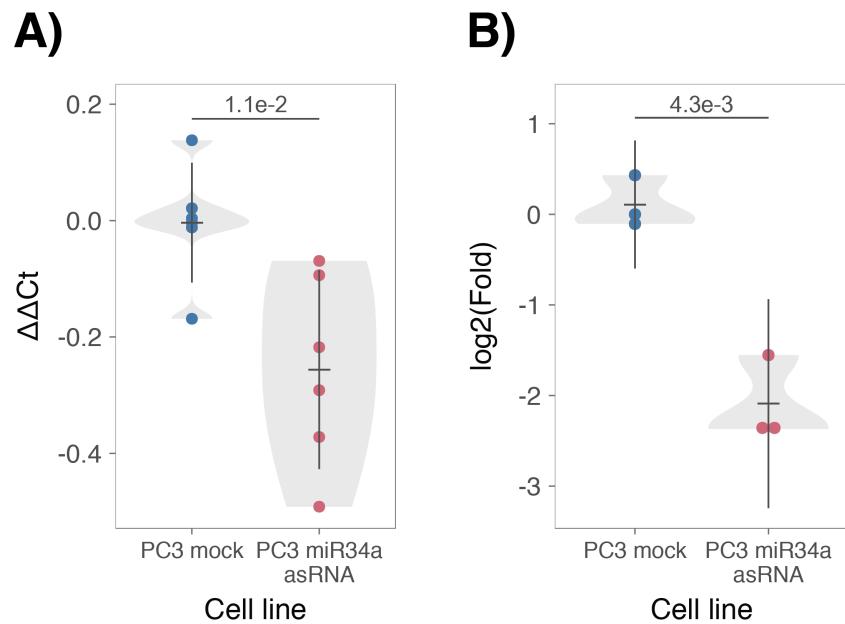


Figure 2 Supplement 2: Evaluating the effects of miR34a asRNA down-regulation. HEK293T cells were co-transfected with the p1 construct and either shRenilla or shControl. Renilla and luciferase levels were measured with Q-PCR 48 hours after transfection. Individual points represent independent experiments with the gray shadow indicating the density of the points. The experiment was performed in biological triplicate.



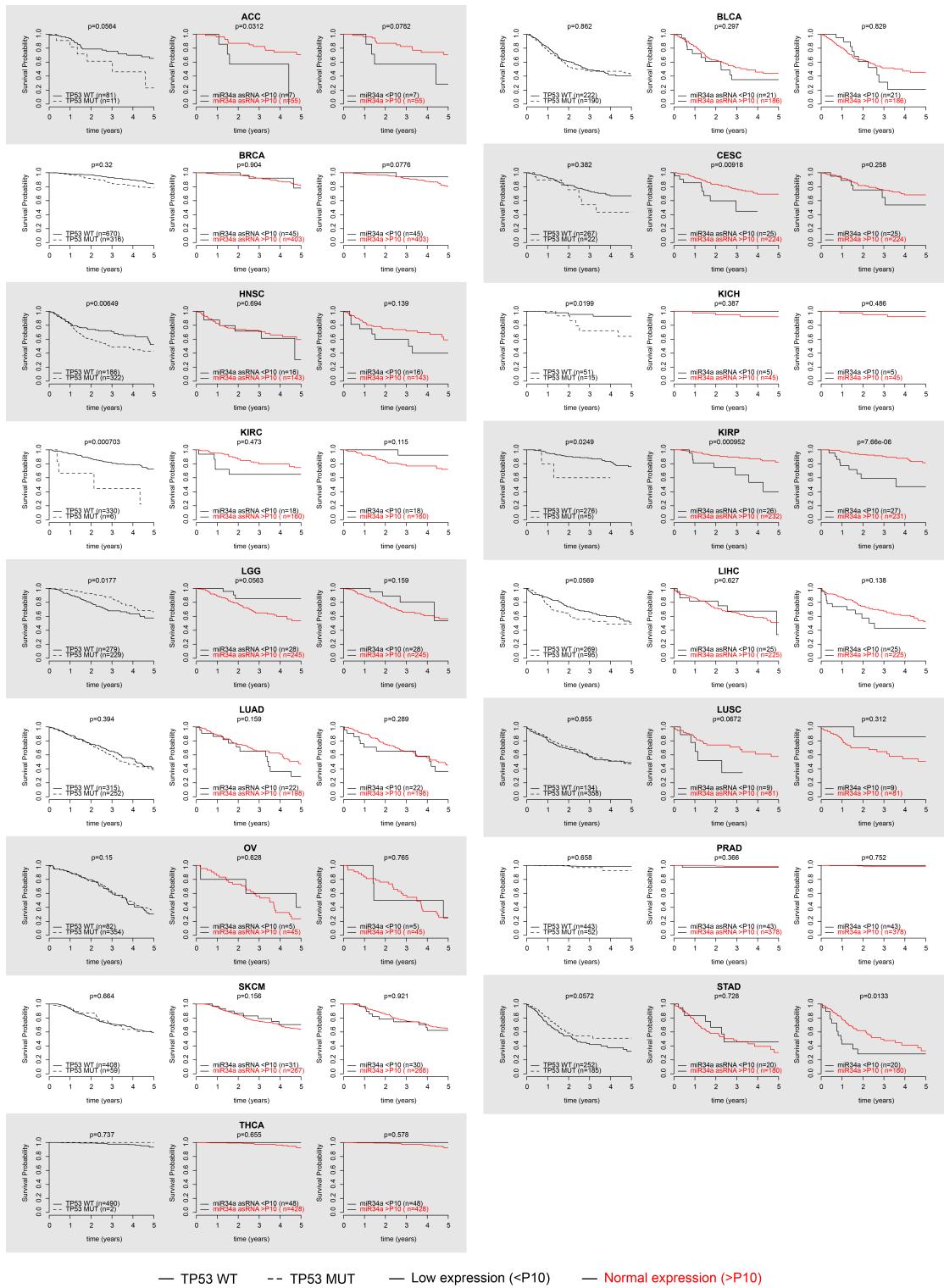
846
847

848 **Figure 3 Supplement 1: Physiological relevance of miR34a asRNA overexpression.** Comparison
849 of miR34a asRNA expression in HEK293T cells (high endogenous miR34a asRNA), and the wild-type
850 (wt), mock, and miR34a asRNA over-expressing stable cell lines.



851
852
853
854
855

Figure 3 Supplement 2: Effects of miR34a asRNA overexpression on cyclin D1. CCND1 expression (A) and western blot quantification of protein levels (B) in *miR34a* asRNA over-expressing PC3 stable cell lines. Experiments were performed in biological sextuplets (A) or triplicates (B).



856

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

861

862

863 References

864

865 Agostini, M., P. Tucci, R. Killick, E. Candi, B. S. Sayan, P. Rivetti di Val Cervo, P.
866 Nicotera, F. McKeon, R. A. Knight, T. W. Mak and G. Melino (2011). "Neuronal
867 differentiation by TAp73 is mediated by microRNA-34a regulation of synaptic
868 protein targets." *Proc Natl Acad Sci U S A* **108**(52): 21093-21098. DOI:
869 10.1073/pnas.1112061109

870

871 Ahn, Y. H., D. L. Gibbons, D. Chakravarti, C. J. Creighton, Z. H. Rizvi, H. P. Adams, A.
872 Pertsemlidis, P. A. Gregory, J. A. Wright, G. J. Goodall, E. R. Flores and J. M. Kurie
873 (2012). "ZEB1 drives prometastatic actin cytoskeletal remodeling by
874 downregulating miR-34a expression." *J Clin Invest* **122**(9): 3170-3183. DOI:
875 10.1172/JCI63608

876

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

880

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

884

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

887

888 Ashouri, A., V. I. Sayin, J. Van den Eynden, S. X. Singh, T. Papagiannakopoulos and
889 E. Larsson (2016). "Pan-cancer transcriptomic analysis associates long non-
890 coding RNAs with key mutational driver events." *Nature Communications* **7**:
891 13197. DOI: 10.1038/ncomms13197

892

893 Balbin, O. A., R. Malik, S. M. Dhanasekaran, J. R. Prensner, X. Cao, Y. M. Wu, D.
894 Robinson, R. Wang, G. Chen, D. G. Beer, A. I. Nesvizhskii and A. M. Chinnaiyan
895 (2015). "The landscape of antisense gene expression in human cancers." *Genome*
896 *Res* **25**(7): 1068-1079. DOI: 10.1101/gr.180596.114

897

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

902

903 Chang, T. C., D. Yu, Y. S. Lee, E. A. Wentzel, D. E. Arking, K. M. West, C. V. Dang, A.
904 Thomas-Tikhonenko and J. T. Mendell (2008). "Widespread microRNA
905 repression by Myc contributes to tumorigenesis." *Nat Genet* **40**(1): 43-50. DOI:
906 10.1038/ng.2007.30

907

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

910

- 911 Chen, J., M. Sun, W. J. Kent, X. Huang, H. Xie, W. Wang, G. Zhou, R. Z. Shi and J. D.
912 Rowley (2004). "Over 20% of human transcripts might form sense-antisense
913 pairs." *Nucleic Acids Res* **32**(16): 4812-4820. DOI: 10.1093/nar/gkh818
- 914
- 915 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.
916 S. Zheng (2010). "The impact of miR-34a on protein output in hepatocellular
917 carcinoma HepG2 cells." *Proteomics* **10**(8): 1557-1572. DOI:
918 10.1002/pmic.200900646
- 919
- 920 Chim, C. S., K. Y. Wong, Y. Qi, F. Loong, W. L. Lam, L. G. Wong, D. Y. Jin, J. F. Costello
921 and R. Liang (2010). "Epigenetic inactivation of the miR-34a in hematological
922 malignancies." *Carcinogenesis* **31**(4): 745-750. DOI: 10.1093/carcin/bgq033
- 923
- 924 Cole, K. A., E. F. Attiyeh, Y. P. Mosse, M. J. Laquaglia, S. J. Diskin, G. M. Brodeur and
925 J. M. Maris (2008). "A functional screen identifies miR-34a as a candidate
926 neuroblastoma tumor suppressor gene." *Mol Cancer Res* **6**(5): 735-742. DOI:
927 10.1158/1541-7786.MCR-07-2102
- 928
- 929 Conley, A. B. and I. K. Jordan (2012). "Epigenetic regulation of human cis-natural
930 antisense transcripts." *Nucleic Acids Res* **40**(4): 1438-1445. DOI:
931 10.1093/nar/gkr1010
- 932
- 933 Consortium, E. P. (2012). "An integrated encyclopedia of DNA elements in the
934 human genome." *Nature* **489**(7414): 57-74. DOI: 10.1038/nature11247
- 935
- 936 Ding, N., H. Wu, T. Tao and E. Peng (2017). "NEAT1 regulates cell proliferation
937 and apoptosis of ovarian cancer by miR-34a-5p/BCL2." *Onco Targets Ther* **10**:
938 4905-4915. DOI: 10.2147/OTT.S142446
- 939
- 940 Djebali, S., C. A. Davis, A. Merkel, A. Dobin, T. Lassmann, A. Mortazavi, A. Tanzer, J.
941 Lagarde, W. Lin, F. Schlesinger, C. Xue, G. K. Marinov, J. Khatun, B. A. Williams, C.
942 Zaleski, J. Rozowsky, M. Roder, F. Kokocinski, R. F. Abdelhamid, T. Alioto, I.
943 Antoshechkin, M. T. Baer, N. S. Bar, P. Batut, K. Bell, I. Bell, S. Chakrabortty, X.
944 Chen, J. Chrast, J. Curado, T. Derrien, J. Drenkow, E. Dumais, J. Dumais, R.
945 Duttagupta, E. Falconnet, M. Fastuca, K. Fejes-Toth, P. Ferreira, S. Foissac, M. J.
946 Fullwood, H. Gao, D. Gonzalez, A. Gordon, H. Gunawardena, C. Howald, S. Jha, R.
947 Johnson, P. Kapranov, B. King, C. Kingswood, O. J. Luo, E. Park, K. Persaud, J. B.
948 Preall, P. Ribeca, B. Risk, D. Robyr, M. Sammeth, L. Schaffer, L. H. See, A. Shahab, J.
949 Skancke, A. M. Suzuki, H. Takahashi, H. Tilgner, D. Trout, N. Walters, H. Wang, J.
950 Wrobel, Y. Yu, X. Ruan, Y. Hayashizaki, J. Harrow, M. Gerstein, T. Hubbard, A.
951 Reymond, S. E. Antonarakis, G. Hannon, M. C. Giddings, Y. Ruan, B. Wold, P.
952 Carninci, R. Guigo and T. R. Gingeras (2012). "Landscape of transcription in
953 human cells." *Nature* **489**(7414): 101-108. DOI: 10.1038/nature11233
- 954
- 955 Gallardo, E., A. Navarro, N. Vinolas, R. M. Marrades, T. Diaz, B. Gel, A. Quera, E.
956 Bandres, J. Garcia-Foncillas, J. Ramirez and M. Monzo (2009). "miR-34a as a
957 prognostic marker of relapse in surgically resected non-small-cell lung cancer."
958 *Carcinogenesis* **30**(11): 1903-1909. DOI: 10.1093/carcin/bgp219
- 959

- 960 Grossman, R. L., A. P. Heath, V. Ferretti, H. E. Varmus, D. R. Lowy, W. A. Kibbe and
961 L. M. Staudt (2016). "Toward a Shared Vision for Cancer Genomic Data." N Engl J
962 Med **375**(12): 1109-1112. DOI: 10.1056/NEJMmp1607591
- 963
- 964 Hunten, S., M. Kaller, F. Drepper, S. Oeljeklaus, T. Bonfert, F. Erhard, A. Dueck, N.
965 Eichner, C. C. Friedel, G. Meister, R. Zimmer, B. Warscheid and H. Hermeking
966 (2015). "p53-Regulated Networks of Protein, mRNA, miRNA, and lncRNA
967 Expression Revealed by Integrated Pulsed Stable Isotope Labeling With Amino
968 Acids in Cell Culture (pSILAC) and Next Generation Sequencing (NGS) Analyses."
969 Mol Cell Proteomics **14**(10): 2609-2629. DOI: 10.1074/mcp.M115.050237
- 970
- 971 International Human Genome Sequencing, C. (2004). "Finishing the euchromatic
972 sequence of the human genome." Nature **431**(7011): 931-945. DOI:
973 10.1038/nature03001
- 974
- 975 Johnsson, P., A. Ackley, L. Vidarsdottir, W. O. Lui, M. Corcoran, D. Grander and K.
976 V. Morris (2013). "A pseudogene long-noncoding-RNA network regulates PTEN
977 transcription and translation in human cells." Nat Struct Mol Biol **20**(4): 440-
978 446. DOI: 10.1038/nsmb.2516
- 979
- 980 Katayama, S., Y. Tomaru, T. Kasukawa, K. Waki, M. Nakanishi, M. Nakamura, H.
981 Nishida, C. C. Yap, M. Suzuki, J. Kawai, H. Suzuki, P. Carninci, Y. Hayashizaki, C.
982 Wells, M. Frith, T. Ravasi, K. C. Pang, J. Hallinan, J. Mattick, D. A. Hume, L. Lipovich,
983 S. Batalov, P. G. Engstrom, Y. Mizuno, M. A. Faghihi, A. Sandelin, A. M. Chalk, S.
984 Mottagui-Tabar, Z. Liang, B. Lenhard, C. Wahlestedt, R. G. E. R. Group, G. Genome
985 Science and F. Consortium (2005). "Antisense transcription in the mammalian
986 transcriptome." Science **309**(5740): 1564-1566. DOI: 10.1126/science.1112009
- 987
- 988 Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler and D.
989 Haussler (2002). "The human genome browser at UCSC." Genome Res **12**(6):
990 996-1006. DOI: 10.1101/gr.229102. Article published online before print in May
991 2002
- 992
- 993 Kim, K. H., H. J. Kim and T. R. Lee (2017). "Epidermal long non-coding RNAs are
994 regulated by ultraviolet irradiation." Gene **637**: 196-202. DOI:
995 10.1016/j.gene.2017.09.043
- 996
- 997 Kong, L., Y. Zhang, Z. Q. Ye, X. Q. Liu, S. Q. Zhao, L. Wei and G. Gao (2007). "CPC:
998 assess the protein-coding potential of transcripts using sequence features and
999 support vector machine." Nucleic Acids Res **35**(Web Server issue): W345-349.
1000 DOI: 10.1093/nar/gkm391
- 1001
- 1002 Lal, A., M. P. Thomas, G. Altschuler, F. Navarro, E. O'Day, X. L. Li, C. Concepcion, Y.
1003 C. Han, J. Thiery, D. K. Rajani, A. Deutsch, O. Hofmann, A. Ventura, W. Hide and J.
1004 Lieberman (2011). "Capture of microRNA-bound mRNAs identifies the tumor
1005 suppressor miR-34a as a regulator of growth factor signaling." PLoS Genet **7**(11):
1006 e1002363. DOI: 10.1371/journal.pgen.1002363
- 1007

- 1008 Leveille, N., C. A. Melo, K. Rooijers, A. Diaz-Lagares, S. A. Melo, G. Korkmaz, R.
1009 Lopes, F. Akbari Moqadam, A. R. Maia, P. J. Wijchers, G. Geeven, M. L. den Boer, R.
1010 Kalluri, W. de Laat, M. Esteller and R. Agami (2015). "Genome-wide profiling of
1011 p53-regulated enhancer RNAs uncovers a subset of enhancers controlled by a
1012 lncRNA." *Nature Communications* **6**: 6520. DOI: 10.1038/ncomms7520
- 1013
- 1014 Liu, C., K. Kelnar, B. Liu, X. Chen, T. Calhoun-Davis, H. Li, L. Patrawala, H. Yan, C.
1015 Jeter, S. Honorio, J. F. Wiggins, A. G. Bader, R. Fagin, D. Brown and D. G. Tang
1016 (2011). "The microRNA miR-34a inhibits prostate cancer stem cells and
1017 metastasis by directly repressing CD44." *Nat Med* **17**(2): 211-215. DOI:
1018 10.1038/nm.2284
- 1019
- 1020 Memczak, S., M. Jens, A. Elefsinioti, F. Torti, J. Krueger, A. Rybak, L. Maier, S. D.
1021 Mackowiak, L. H. Gregersen, M. Munschauer, A. Loewer, U. Ziebold, M.
1022 Landthaler, C. Kocks, F. le Noble and N. Rajewsky (2013). "Circular RNAs are a
1023 large class of animal RNAs with regulatory potency." *Nature* **495**(7441): 333-
1024 338. DOI: 10.1038/nature11928
- 1025
- 1026 O'Leary, N. A., M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad, R. McVeigh, B.
1027 Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y.
1028 Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C. M.
1029 Farrell, T. Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V. S. Joardar, V. K.
1030 Kodali, W. Li, D. Maglott, P. Masterson, K. M. McGarvey, M. R. Murphy, K. O'Neill,
1031 S. Pujar, S. H. Rangwala, D. Rausch, L. D. Riddick, C. Schoch, A. Shkeda, S. S. Storz,
1032 H. Sun, F. Thibaud-Nissen, I. Tolstoy, R. E. Tully, A. R. Vatsan, C. Wallin, D. Webb,
1033 W. Wu, M. J. Landrum, A. Kimchi, T. Tatusova, M. DiCuccio, P. Kitts, T. D. Murphy
1034 and K. D. Pruitt (2016). "Reference sequence (RefSeq) database at NCBI: current
1035 status, taxonomic expansion, and functional annotation." *Nucleic Acids Res*
1036 **44**(D1): D733-745. DOI: 10.1093/nar/gkv1189
- 1037
- 1038 Ozsolak, F., P. Kapranov, S. Foissac, S. W. Kim, E. Fishilevich, A. P. Monaghan, B.
1039 John and P. M. Milos (2010). "Comprehensive polyadenylation site maps in yeast
1040 and human reveal pervasive alternative polyadenylation." *Cell* **143**(6): 1018-
1041 1029. DOI: 10.1016/j.cell.2010.11.020
- 1042
- 1043 Polson, A., E. Durrett and D. Reisman (2011). "A bidirectional promoter reporter
1044 vector for the analysis of the p53/WDR79 dual regulatory element." *Plasmid*
1045 **66**(3): 169-179. DOI: 10.1016/j.plasmid.2011.08.004
- 1046
- 1047 Rashi-Elkeles, S., H. J. Warnatz, R. Elkon, A. Kupershtain, Y. Chobod, A. Paz, V.
1048 Amstislavskiy, M. Sultan, H. Safer, W. Nietfeld, H. Lehrach, R. Shamir, M. L. Yaspo
1049 and Y. Shiloh (2014). "Parallel profiling of the transcriptome, cistrome, and
1050 epigenome in the cellular response to ionizing radiation." *Sci Signal* **7**(325): rs3.
1051 DOI: 10.1126/scisignal.2005032
- 1052
- 1053 Raver-Shapira, N., E. Marciano, E. Meiri, Y. Spector, N. Rosenfeld, N. Moskovits, Z.
1054 Bentwich and M. Oren (2007). "Transcriptional activation of miR-34a
1055 contributes to p53-mediated apoptosis." *Mol Cell* **26**(5): 731-743. DOI:
1056 10.1016/j.molcel.2007.05.017

- 1057
1058 Rinn, J. L., M. Kertesz, J. K. Wang, S. L. Squazzo, X. Xu, S. A. Brugmann, L. H.
1059 Goodnough, J. A. Helms, P. J. Farnham, E. Segal and H. Y. Chang (2007).
1060 "Functional demarcation of active and silent chromatin domains in human HOX
1061 loci by noncoding RNAs." *Cell* **129**(7): 1311-1323. DOI:
1062 10.1016/j.cell.2007.05.022
1063
1064 Rokavec, M., M. G. Oner, H. Li, R. Jackstadt, L. Jiang, D. Lodygin, M. Kaller, D. Horst,
1065 P. K. Ziegler, S. Schwitalla, J. Slotta-Huspenina, F. G. Bader, F. R. Greten and H.
1066 Hermeking (2015). "Corrigendum. IL-6R/STAT3/miR-34a feedback loop
1067 promotes EMT-mediated colorectal cancer invasion and metastasis." *J Clin Invest*
1068 **125**(3): 1362. DOI: 10.1172/JCI81340
1069
1070 Schneider, C. A., W. S. Rasband and K. W. Eliceiri (2012). "NIH Image to ImageJ:
1071 25 years of image analysis." *Nat Methods* **9**(7): 671-675.
1072
1073 Serviss, J. T. (2017). miR34AasRNAproject.
1074 https://github.com/GranderLab/miR34a_asRNA_project
1075
1076 Serviss, J. T., P. Johnsson and D. Grander (2014). "An emerging role for long non-
1077 coding RNAs in cancer metastasis." *Front Genet* **5**: 234. DOI:
1078 10.3389/fgene.2014.00234
1079
1080 Slabakova, E., Z. Culig, J. Remsik and K. Soucek (2017). "Alternative mechanisms
1081 of miR-34a regulation in cancer." *Cell Death Dis* **8**(10): e3100. DOI:
1082 10.1038/cddis.2017.495
1083
1084 Stahlhut, C. and F. J. Slack (2015). "Combinatorial Action of MicroRNAs let-7 and
1085 miR-34 Effectively Synergizes with Erlotinib to Suppress Non-small Cell Lung
1086 Cancer Cell Proliferation." *Cell Cycle* **14**(13): 2171-2180. DOI:
1087 10.1080/15384101.2014.1003008
1088
1089 Su, X., D. Chakravarti, M. S. Cho, L. Liu, Y. J. Gi, Y. L. Lin, M. L. Leung, A. El-Naggar,
1090 C. J. Creighton, M. B. Suraokar, I. Wistuba and E. R. Flores (2010). "TAp63
1091 suppresses metastasis through coordinate regulation of Dicer and miRNAs."
1092 *Nature* **467**(7318): 986-990. DOI: 10.1038/nature09459
1093
1094 Sun, F., H. Fu, Q. Liu, Y. Tie, J. Zhu, R. Xing, Z. Sun and X. Zheng (2008).
1095 "Downregulation of CCND1 and CDK6 by miR-34a induces cell cycle arrest."
1096 *FEBS Lett* **582**(10): 1564-1568. DOI: 10.1016/j.febslet.2008.03.057
1097
1098 Tarasov, V., P. Jung, B. Verdoodt, D. Lodygin, A. Epanchintsev, A. Menssen, G.
1099 Meister and H. Hermeking (2007). "Differential regulation of microRNAs by p53
1100 revealed by massively parallel sequencing: miR-34a is a p53 target that induces
1101 apoptosis and G1-arrest." *Cell Cycle* **6**(13): 1586-1593. DOI:
1102 10.4161/cc.6.13.4436
1103
1104 Team, R. C. (2017). "R: A Language and Environment for Statistical Computing."
1105 from <https://www.R-project.org/>.

- 1106
1107 Therneau, T. M. (2015). A Package for Survival Analysis in S. version 2.38.
1108 <https://CRAN.R-project.org/package=survival>
1109
1110 Turner, A. M., A. M. Ackley, M. A. Matrone and K. V. Morris (2012).
1111 "Characterization of an HIV-targeted transcriptional gene-silencing RNA in
1112 primary cells." *Hum Gene Ther* **23**(5): 473-483. DOI: 10.1089/hum.2011.165
1113
1114 Vogt, M., J. Munding, M. Gruner, S. T. Liffers, B. Verdoodt, J. Hauk, L.
1115 Steinstraesser, A. Tannapfel and H. Hermeking (2011). "Frequent concomitant
1116 inactivation of miR-34a and miR-34b/c by CpG methylation in colorectal,
1117 pancreatic, mammary, ovarian, urothelial, and renal cell carcinomas and soft
1118 tissue sarcomas." *Virchows Arch* **458**(3): 313-322. DOI: 10.1007/s00428-010-
1119 1030-5
1120
1121 Wang, L., P. Bu, Y. Ai, T. Srinivasan, H. J. Chen, K. Xiang, S. M. Lipkin and X. Shen
1122 (2016). "A long non-coding RNA targets microRNA miR-34a to regulate colon
1123 cancer stem cell asymmetric division." *eLife* **5**. DOI: 10.7554/eLife.14620
1124
1125 Wang, L., H. J. Park, S. Dasari, S. Wang, J. P. Kocher and W. Li (2013). "CPAT:
1126 Coding-Potential Assessment Tool using an alignment-free logistic regression
1127 model." *Nucleic Acids Res* **41**(6): e74. DOI: 10.1093/nar/gkt006
1128
1129 Wang, X., J. Li, K. Dong, F. Lin, M. Long, Y. Ouyang, J. Wei, X. Chen, Y. Weng, T. He
1130 and H. Zhang (2015). "Tumor suppressor miR-34a targets PD-L1 and functions
1131 as a potential immunotherapeutic target in acute myeloid leukemia." *Cell Signal*
1132 **27**(3): 443-452. DOI: 10.1016/j.cellsig.2014.12.003
1133
1134 Wickham, H. (2016). gtable: Arrange 'Grobs' in Tables. R package version 0.2.0.
1135 <https://CRAN.R-project.org/package=gtable>
1136
1137 Wickham, H. (2017). scales: Scale Functions for Visualization. R package version
1138 0.5.0. <https://CRAN.R-project.org/package=scales>
1139
1140 Wickham, H. (2017). tidyverse: Easily Install and Load the 'Tidyverse'. R package
1141 version 1.2.1. <https://CRAN.R-project.org/package=tidyverse>
1142
1143 Wickham, L. H. a. H. (2017). rlang: Functions for Base Types and Core R and
1144 'Tidyverse' Features. R package version 0.1.4. [https://CRAN.R-
1145 project.org/package=rlang](https://CRAN.R-project.org/package=rlang)
1146
1147 Wickham, S. M. B. a. H. (2014). magrittr: A Forward-Pipe Operator for R. R
1148 package version 1.5. <https://CRAN.R-project.org/package=magrittr>
1149
1150 Wilkins, D. gggenes: Draw Gene Arrow Maps in 'ggplot2'. R package version
1151 0.2.0.9003. <https://github.com/wilkox/gggenes>
1152
1153 Xiao, N. (2017). liftr: Containerize R Markdown Documents. R package version
1154 0.7. <https://CRAN.R-project.org/package=liftr>

- 1155
1156 Xie, Y. (2017). knitr: A General-Purpose Package for Dynamic Report Generation
1157 in R. R package version 1.17. <https://yihui.name/knitr/>
1158
1159 Yang, P., Q. J. Li, Y. Feng, Y. Zhang, G. J. Markowitz, S. Ning, Y. Deng, J. Zhao, S.
1160 Jiang, Y. Yuan, H. Y. Wang, S. Q. Cheng, D. Xie and X. F. Wang (2012). "TGF-beta-
1161 miR-34a-CCL22 signaling-induced Treg cell recruitment promotes venous
1162 metastases of HBV-positive hepatocellular carcinoma." *Cancer Cell* **22**(3): 291-
1163 303. DOI: 10.1016/j.ccr.2012.07.023
1164
1165 Yap, K. L., S. Li, A. M. Munoz-Cabello, S. Raguz, L. Zeng, S. Mujtaba, J. Gil, M. J.
1166 Walsh and M. M. Zhou (2010). "Molecular interplay of the noncoding RNA ANRIL
1167 and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional
1168 silencing of INK4a." *Mol Cell* **38**(5): 662-674. DOI: 10.1016/j.molcel.2010.03.021
1169
1170 Yu, W., D. Gius, P. Onyango, K. Muldoon-Jacobs, J. Karp, A. P. Feinberg and H. Cui
1171 (2008). "Epigenetic silencing of tumour suppressor gene p15 by its antisense
1172 RNA." *Nature* **451**(7175): 202-206. DOI: 10.1038/nature06468
1173
1174 Zenz, T., J. Mohr, E. Eldering, A. P. Kater, A. Buhler, D. Kienle, D. Winkler, J. Durig,
1175 M. H. van Oers, D. Mertens, H. Dohner and S. Stilgenbauer (2009). "miR-34a as
1176 part of the resistance network in chronic lymphocytic leukemia." *Blood* **113**(16):
1177 3801-3808. DOI: 10.1182/blood-2008-08-172254
1178
1179
1180
1181