Cell

Widespread Shortening of 3′UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells

Christine Mayr[1,2,3,4,*] and David P. Bartel[1,2,3,*]
[1]Howard Hughes Medical Institute
[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA
[4]Present address: Cancer Biology and Genetics Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA
*Correspondence: dbartel@wi.mit.edu (D.P.B.), mayrc@mskcc.org (C.M.)
DOI 10.1016/j.cell.2009.06.016

SUMMARY

In cancer cells, genetic alterations can activate proto-oncogenes, thereby contributing to tumorigenesis. However, the protein products of oncogenes are sometimes overexpressed without alteration of the proto-oncogene. Helping to explain this phenomenon, we found that when compared to similarly proliferating nontransformed cell lines, cancer cell lines often expressed substantial amounts of mRNA isoforms with shorter 3′ untranslated regions (UTRs). These shorter isoforms usually resulted from alternative cleavage and polyadenylation (APA). The APA had functional consequences, with the shorter mRNA isoforms exhibiting increased stability and typically producing ten-fold more protein, in part through the loss of microRNA-mediated repression. Moreover, expression of the shorter mRNA isoform of the proto-oncogene *IGF2BP1/IMP-1* led to far more oncogenic transformation than did expression of the full-length, annotated mRNA. The high incidence of APA in cancer cells, with consequent loss of 3′UTR repressive elements, suggests a pervasive role for APA in oncogene activation without genetic alteration.

INTRODUCTION

Oncogene activation plays a central role in the pathogenesis of cancer. Mammalian genomes contain proto-oncogenes that function to regulate normal cell proliferation and differentiation. In cancer cells, these genes frequently have been activated—sometimes by a mutation that changes the encoded proteins and other times by a mutation that increases the expression of the gene. The increased expression can occur through several mechanisms, including gene amplifications that increase copy number and chromosomal translocations that replace the original promoter with a much more active one (Bishop, 1991; Rowley, 2001).

A recently recognized mechanism of oncogene activation is the loss of microRNA (miRNA) complementary sites (Mayr et al., 2007; Lee and Dutta, 2007). miRNAs specify posttranscriptional repression by pairing to short sites in mRNA targets, usually in 3′UTRs (Bartel, 2009). In the case of *HMGA2* activation, chromosomal translocations swap the *HMGA2* 3′UTR for that of another gene, which results in loss of complementary sites for the *let-7* miRNA and escape from repression (Mayr et al., 2007; Lee and Dutta, 2007). In principle, point substitutions or other lesions that disrupt individual target sites or that shorten the 3′UTR through premature cleavage and polyadenylation could have the same effect (Majoros and Ohler, 2007). Hinting at this possibility are events associated with a more aggressive subgroup of mantle cell lymphomas. Mantle cell lymphomas are characterized by the t(11;14)(q13;q32) translocation, which brings *Cyclin D1* near the highly active immunoglobulin heavy chain gene promoter, leading to Cyclin D1 overexpression. In a subset of mantle cell lymphomas a second alteration of *Cyclin D1*, the shortening of its 3′UTR, which might relieve miRNA-directed repression, leads to an additional 1.6-fold increase in Cyclin D1 expression and correlates with both increased proliferation of the lymphoma cells and decreased overall survival of patients (Rosenwald et al., 2003). The reasons for 3′UTR shortening are unknown in a third of cases, but in the remaining cases either genomic deletions encompass much of the 3′UTR or point substitutions create premature cleavage and polyadenylation signals (Wiestner et al., 2007). For some cases in which DNA lesions are not found, the *Cyclin D1* 3′UTR might be shortened through alternative cleavage and polyadenylation (APA).

Cleavage and polyadenylation is required for the maturation of most mRNA transcripts (Proudfoot, 1991; Colgan and Manley, 1997). The pre-mRNA is cleaved 10–30 nt after the polyadenylation signal (AAUAAA and variants thereof) and an untemplated poly(A) tract is added. Although a strong polyadenylation signal usually is located at the 3′ end of the 3′UTR, nearly all genes have additional polyadenylation signals in their 3′UTRs, with about half of human genes possessing conserved APA signals with usage supported by expressed sequence tags (ESTs) (Tian et al., 2005). Use of APA signals often eliminates large parts of the 3′UTR, enabling escape from the stronger regulatory potential of longer 3′UTRs. Besides miRNA regulation, the lost

Table 1. Candidate Genes

Gene	Isoforms[a]	miRNA	Sites[b]	Function
HMGA2	>1	*let-7*	7	Architectural transcription factor; cancer, stem cell biology
C12orf28	0	*let-7*	4	Unknown
LIN28B	1	*let-7*	5	Homolog of *Lin28*, reprogramming of human fibroblasts
TRIM71	0	*let-7*	2	Development
IMP-1	3	*let-7*	5	RNA-binding protein; mRNA stability, translational control, mRNA transport
ARID3B	1	*let-7*	5	DNA-binding protein; development, oncogene
FIGN	0	*let-7*	3	Development
PUNC	0	*let-7*	3	Development
SMARCAD1	0	*let-7*	2	Helicase; chromatin remodeling
YOD1	1	*let-7*	4	Deubiquitinating enzyme
BACH1	1	*let-7*	2	Transcription factor
CPEB2	1	*let-7*	3	RNA-binding protein; development, cell division, senescence, synaptic plasticity
IMP-3	2	*let-7*	1	RNA binding protein
IMP-2	2	*let-7*	2	RNA binding protein
FGF2	3	miR-15/16	5	Mitogen; development, wound healing, tumor growth
PLAG1	1	miR-15/16	2	Zink finger protein; rearranged in pleomorphic adenoma; rearranged in pleomorphic adenoma
MYB	1	miR-15/16	2	Transcription factor; hematopoiesis, oncogene
TACC1	0	miR-15/16	2	Oncogene
CCND2	2	miR-15/16	3	Cyclin; oncogene
CCND1	2	miR-15/16	2	Cyclin; oncogene
DICER1	2-3	miR-103/107	6	miRNA biogenesis
RAB10	2	miR-103/107	3	Ras-related GTP-binding protein
RUNX1	0	miR-27	2	Subunit of transcription factor; hematopoiesis

[a] 0, not expressed in the cell lines and tissues investigated.
[b] Listed are the number of sites for the indicated miRNAs.

regulatory sequences in the 3′UTR can influence mRNA nuclear export and cytoplasmic localization, as well as non-miRNA-mediated changes in mRNA stability and translational efficiency (Moore, 2005). Alternative mRNAs that differ in their 3′UTRs can exist in different tissues or developmental stages, and studies have shown that these mRNA isoforms can have different stability or translational activity (Miyamoto et al., 1996; Takagaki et al., 1996; Edwalds-Gilbert et al., 1997; Lutz, 2008). Over 30 years ago, stimulation of lymphocytes was shown to increase both RNA polyadenylation and the rate of protein synthesis, without a change in the rate of transcription (Coleman et al., 1974; Hauser et al., 1978). Recently, a genome-wide study identified shorter 3′UTRs in activated T cells compared to resting T cells, and found that in general shorter 3′UTRs were associated with cell proliferation (Sandberg et al., 2008).

We set out to explore the possibility that APA might be a mechanism by which genes can escape miRNA-mediated repression in cancer. We found that shorter 3′UTRs were indeed associated with tumor cells, over and above the association expected from their proliferative state. Reporter assays and immunoblots revealed the functional consequences of APA, showing that the shorter transcripts produced substantially more protein than did their full-length counterparts, in part through escape of miRNA-mediated targeting. Moreover, expressing the shorter but not the full-length isoform of the proto-oncogene *Insulin-like growth factor 2 mRNA binding protein 1* (*IGF2BP1/IMP-1/CRD-BP/ZBP-1)* led to oncogenic transformation, illustrating that loss of repressive 3′UTR elements through APA can promote the oncogenic phenotype.

RESULTS

Shorter 3′UTRs Are Associated with Transformation

To find altered mRNA length and abundance, we used northern blots probing RNA from 27 cancer cell lines from different tissues. The cell lines were derived from sarcomas and breast, lung and colon cancers, and compared to immortalized nontransformed epithelial cell lines and normal corresponding tissues. We chose to analyze 23 genes. Gene function was not considered when selecting these genes. One criterion for selecting a gene was that there were potential APA signals in the 3′UTR. AAUAAA is the canonical polyadenylation signal, but variants of this signal, including AUUAAA, AAGAAA, UAUAAA, AGUAAA are also used. These variants are used less often as the polyadenylation signal for the longest isoform but more often as proximal polyadenylation signals (Tian et al., 2005). When including these

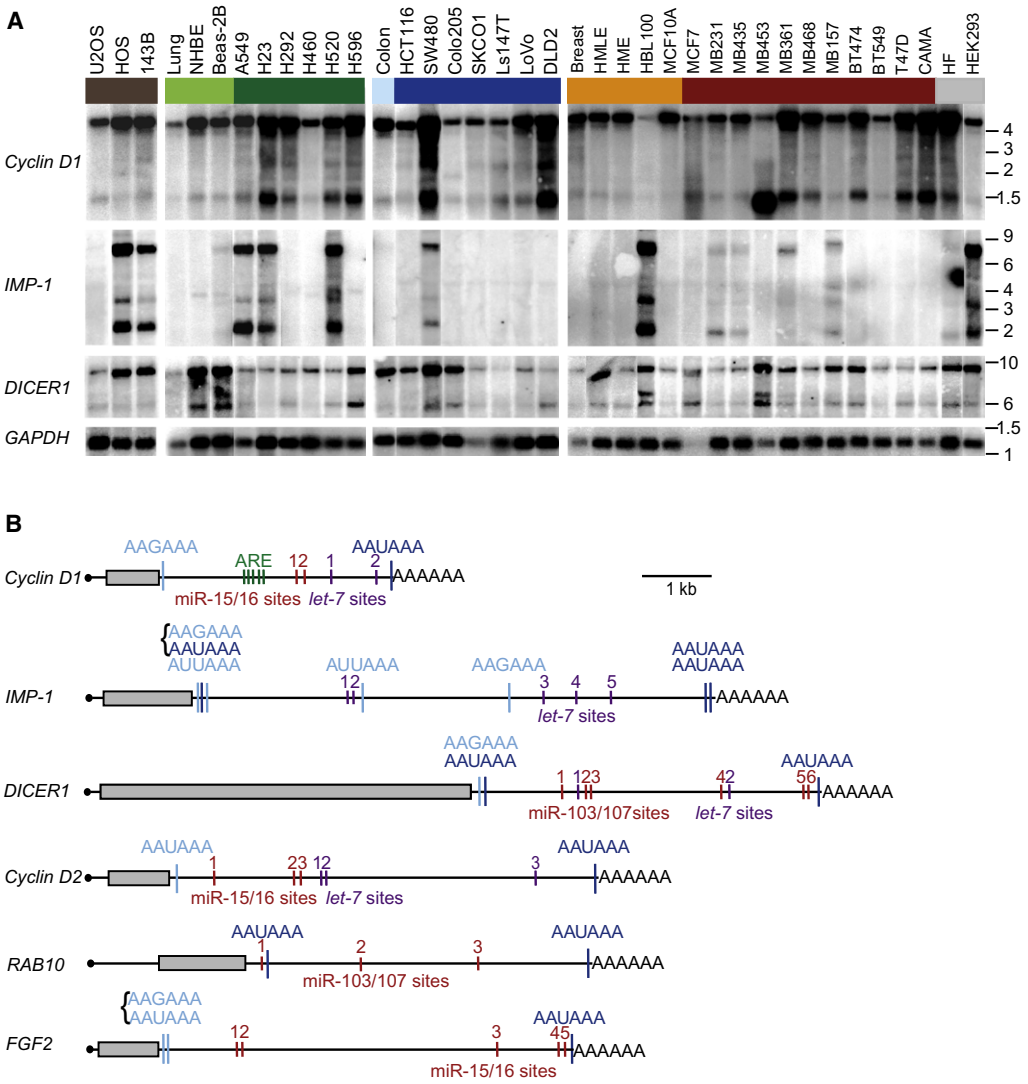

Figure 1. APA Leads to Shorter 3′UTRs

(A) Northern blots of human cell lines and tissues, successively stripped and reprobed for the indicated mRNAs. Groupings include sarcoma cell lines (brown), normal lung tissue, cultured human bronchial epithelial cells (NHBE) and immortalized lung epithelial cell line (Beas-2B, light green), lung cancer cell lines (dark green), normal colon tissue (light blue), colon cancer cell lines (dark blue), normal breast tissue and immortalized breast epithelial cell lines (orange), breast cancer cell lines (red), and other cell lines (gray), which were immortalized fibroblasts (HF) and an embryonic kidney cell line (HEK293). *GAPDH* expression served as loading control. Migration of markers with lengths indicated (kb) is shown at the right.

(B) Schematic illustration of mRNAs with alternative isoforms due to APA. Grey boxes show protein coding region; black line represents untranslated regions, and AAAAAA represents the poly(A) tail. Shown are functional APA signals identified by 3′ RACE, with signals conserved at the orthologous position in four genomes (human, mouse, rat, dog) in dark blue, and more poorly conserved signals in light blue. Brackets indicate closely spaced signals. Also shown are negative regulatory elements identified in the 3′UTRs, including AU-rich elements (ARE, green), *let-7* sites (purple), and sites to the other widely expressed miRNA with the most sites in that 3′UTR (red).

variants, nearly every gene had potential APA signals, and thus this criterion imposed only a minor constraint. The other selection criterion was that the mRNA have strong potential for miRNA regulation, as predicted by TargetScan (version 4.2, (Lewis et al., 2005; Grimson et al., 2007). With one exception (IMP-3), mRNAs with at least two sites to a single miRNA were chosen (Table 1). Reasoning that the miRNAs would be relevant only if expressed in the cancer cell lines, we considered only widely expressed miRNAs (e.g., *let-7*, miR-15/16 and miR-103/107) and confirmed their expression in the cell lines by northern blots (Figure S1 available with this article online). For example, we investigated the top ten predicted targets of *let-7*.

Sixteen of the 23 candidate genes were expressed in the cell lines. In nine of these 16, more than one mRNA isoform was detected (Figures 1A and S2). To investigate if the difference in mRNA length was due to APA, 3′ RACE (rapid amplification of cDNA ends) was performed. For three of the nine genes with multiple isoforms, the differences in mRNA lengths were due to

alternative splicing (data not shown), but for the other six, the differences were due to APA (Figure 1B). For the shorter mRNA isoforms of these six genes, the canonical AAUAAA signal was used in three cases, the variant AAGAAA in three cases, the variant AUUAAA in two cases, and in the other two cases the signals were clustered (AAGAAAAAUAAA), making it difficult to assign which one was used (Figure 1B). Thirty percent of these signals (all AAUAAA) were conserved at the orthologous position in all four mammals examined (human, mouse, rat and dog), and in many of the other cases signals were present at nonorthologous positions. The most 3′ located polyadenylation signal was always AAUAAA and always conserved at the orthologous position.

GU- or U-rich sequences are often found downstream of functional cleavage and polyadenylation signals and are important for signal recognition (Proudfoot, 1991; Colgan and Manley, 1997). Indeed, examining composition of the 200 nucleotides surrounding the functional proximal polyadenylation signals and comparing it with that of the nucleotides surrounding the apparently nonfunctional polyadenylation signals in the genes for which only the full-length mRNA isoform was detected revealed an overrepresentation of U-rich sequences within 50 nt downstream of the functional poly(A) signals ($p = 0.02$; Figure S3).

The genes with alternative isoforms included *Cyclin D1*, *DICER1* and *RAB10*, which were expressed in all samples, and *IMP-1*, *Cyclin D2* and *FGF2*, which were expressed in at least a third of the samples. The shorter mRNAs usually were detectable in normal tissues but more prominent in cancer lines (Figures 1A and S2). The bands of the longest and the shortest mRNA isoform were quantified to determine the expression ratio. Comparing the ratios between the cell lines and their normal corresponding tissues revealed that a majority of the cell lines expressed a higher fraction of the shorter mRNA isoform ($p < 10^{-3}$) (Figure 2A, purple). These results concurred with those of Sandberg et al. (2008), who associated shorter 3′UTRs with a proliferative state. However, the previous study did not distinguish between cancer cell lines and nontransformed proliferating cells. When making this distinction for the lines used in our study, higher amounts of shorter mRNAs were detected significantly more often in cancer cell lines than in nontransformed cell lines ($p < 10^{-4}$), indicating that shorter mRNAs are associated with transformation even more than they are associated with proliferation (Figure 2B). Indeed, no significant difference was found between the nontransformed cell lines compared with normal tissues ($p = 0.35$; Figure 2B).

One possibility raised by our findings was that APA-mediated shortening of 3′UTRs is a cancer-associated phenomenon, without any association with proliferation after accounting for the association with cellular transformation. Alternatively, a more global analysis might still reveal an association with proliferation. To address this issue, we revisited the dataset of Sandberg et al. (2008), who analyzed array data across a broad panel of mammalian cell lines and tissues for differences in both proliferation and 3′UTR length. They created a tandem UTR length index (TLI) to assess aggregate expression of extended 3′UTR regions relative to overall gene-expression levels, and correlated this UTR-length measurement with a gene-signature-based measure of cellular proliferation, called the proliferation index, across a panel of 135 samples. These samples included 22 cancer cell lines, one cancer sample, five nontransformed cell lines, ten samples derived from blood cells and 97 tissues. After subgrouping their data based on these sample types (Figure 2C), we confirmed that even after excluding the cancer lines, proliferation influenced 3′UTR length: Stimulation of blood cells, including T cells, B cells and monocytes, significantly reduced 3′UTR length when compared to their unstimulated counterparts ($p = 0.04$) or normal tissues ($p = 0.005$), and nontransformed cell lines also had shorter 3′UTRs than normal tissues ($p = 0.02$; Figure 2D). But most strikingly, the cancer cell lines had the shortest 3′UTRs (Figures 2C and 2D, red), which were significantly shorter than the 3′UTRs of nontransformed cell lines ($p = 0.007$) despite comparable proliferation indices ($p = 0.6$). This analysis, which extended our northern blot results from a limited number of genes to a genome-wide scale, confirmed that shorter 3′UTRs are associated with transformation over and above the association expected from their proliferative state.

Shorter mRNAs Have Greater Stability and Produce More Protein

What might be the functional consequence of shorter 3′UTRs? In all cases investigated, predicted miRNA target sites were lost in the shorter isoform (Figure 1B). Because miRNAs often destabilize their target mRNAs (Bartel, 2009), the stabilities of the longer and the shorter mRNAs were examined in nine cell lines from diverse tissues for three genes with the most robust mRNA northern signals (*IMP-1*, *Cyclin D1* and *Cyclin D2*). Cells were treated with Actinomycin D to inhibit transcription, total RNA was collected, and the decay of the mRNA isoforms was investigated by northern blot analysis. Across different tissues and different genes, the shorter mRNA was on average 2.6 times more stable than was the longer mRNA (Table 2).

In addition to mRNA destabilization, miRNAs can also direct translational repression of their target mRNAs, which adds to their impact on protein output (Lee et al., 1993; Wightman et al., 1993; Filipowicz et al., 2008). To investigate if different amounts of protein were produced from the shorter and the longer isoforms, we used reporter assays. Both the short and the long 3′UTR isoforms of *IMP-1*, *Cyclin D2, and DICER1* were each cloned downstream of a luciferase reporter gene. To ensure that each long isoform was expressed, for these reporters all the proximal polyadenylation signals that were functional (as detected by 3′ RACE) were mutated (Figure 1B). A strong poly(A) signal from SV40 virus was added downstream of the endogenous polyadenylation signal of the short isoform, as well as that of the long isoform, to ensure both isoforms were expressed as intended. Luciferase reporter gene expression was measured for all three genes in 16 cell lines from different tissues. For all genes in all the cell lines tested, the shorter mRNA isoform produced more protein (Figure 3). The protein expression derived from the shorter mRNA was 1.6- to 42-fold higher than that from the longer mRNA (*IMP-1* median, 8.0-fold; *DICER1* median, 5.5-fold; *Cyclin D2* median, 14.8-fold; average difference overall, 10-fold).

Contribution of miRNA Regulation on Differential Protein Expression Levels

Although the greater mRNA destabilization and translational repression observed for the longer isoforms was consistent

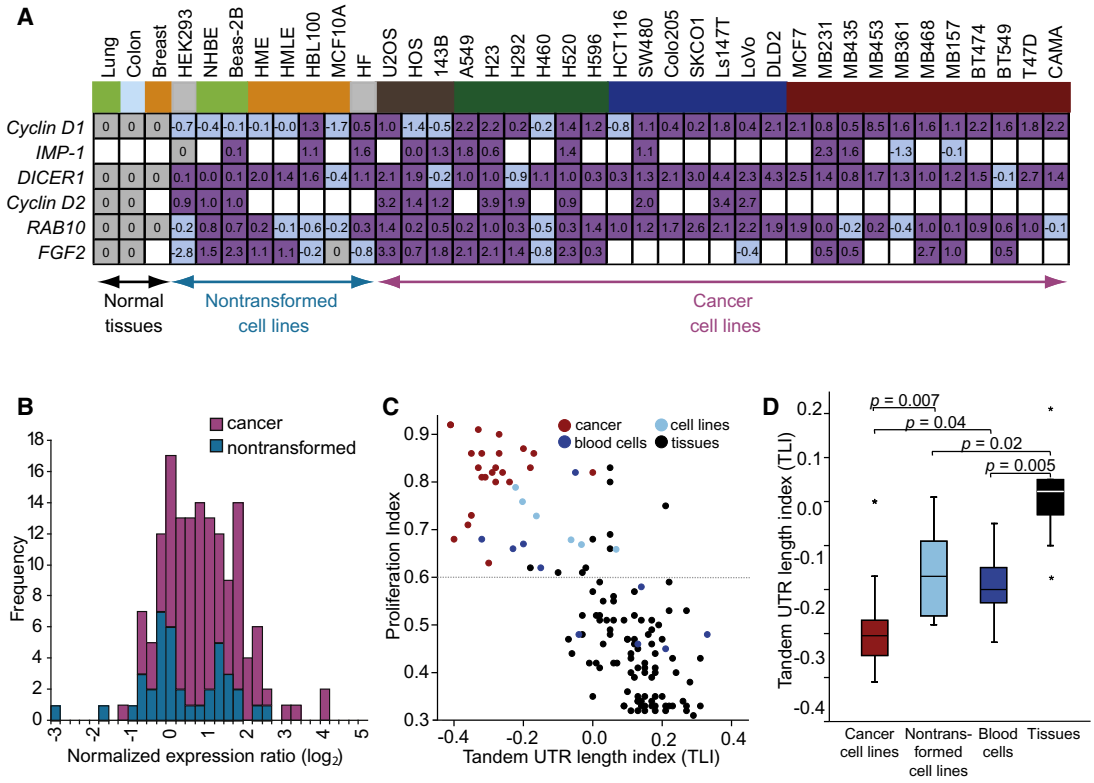

Figure 2. APA Leads to Shorter 3′UTRs Preferentially in Cancer Cells

(A) Quantification of alternative mRNA isoforms. Expression ratio of shortest to longest mRNA isoform from Figures 1A and S2 was normalized to that of the corresponding normal tissue and shown as $\log_2$ values. For the sarcoma cell lines no normal tissue was available, therefore the values were normalized to the median of expression of all normal tissues. *IMP-1* was not expressed in any of the normal tissues, and thus *IMP-1* ratios were normalized to that of HEK293, because for the other genes examined the expression ratios of the mRNA isoforms in HEK293 cells were similar to those of the normal tissues. *FGF2* was not expressed in normal breast tissue; therefore *FGF2* ratios of the breast cell lines were normalized to that MCF10A, a nontransformed breast epithelial line, whose expression ratio of short and long mRNA isoforms resembled that of normal breast tissue for the other genes analyzed. Ratios used for normalization are shaded gray; those with a higher fraction of shorter isoform are purple; those with a lower fraction of shorter isoform are shaded light blue, and cases without detectable mRNA are blank. Cells and tissues are color-coded as in Figure 1A.

(B) Distributions of alternative mRNAs for cancer cells and nontransformed cells. Plotted are the values from Figure 2A, which are the expression ratios of shorter to longer mRNA isoforms, normalized to the corresponding normal tissue. The value for *Cyclin D1* in the breast cancer cell line MB453, which was 8.5 ($\log_2$), is not included because a deletion encompassing the 3′UTR is responsible for the shorter mRNA (Lebwohl et al., 1994). For nontransformed cell lines, the distribution is not significantly displaced from zero ($p = 0.35$), indicating similar ratios of shorter to longer mRNA isoforms as the normal tissues. For cancer cell lines, the ratio is significantly displaced from both zero ($p < 10^{-4}$) and the distribution for nontransformed cell lines ($p < 10^{-4}$), indicating increased ratios of the shorter isoforms.

(C) Genome-wide analyses of 3′UTR length and cellular proliferation. The figure is redrawn from Sandberg et al. (2008), except data points are colored to indicate different types of samples, with cancer samples red (including 22 cancer cell lines and 1 cancer sample), nontransformed cell lines light blue, blood cells (including resting and stimulated T cells, B cells and monocytes) dark blue and normal tissues black. The tandem UTR length index (TLI) is a relative measure of 3′UTR length derived from array data (Sandberg et al., 2008).

(D) The 3′UTR length as measured by TLI for different sample types, focusing only on proliferating cells. TLI values from samples of panel (C) above a proliferation threshold of proliferation index = 0.6 (dotted line in panel [C]) were plotted (horizontal line, median; box, 25th through 75th percentile; error bars, range; asterisk, outliers).

with regulation by the miRNAs expressed endogenously within the cell lines, other types of posttranscriptional repression lost in the shorter isoforms also might have explained some or all of the observed differences. To begin to determine the impact of miRNA regulation on protein expression, we mutated the miRNA complementary sites for the miRNAs with the highest number of sites. The five *let-7* sites in the *IMP-1* 3′UTR, the six sites for miR-103/107 in the *DICER1* 3′UTR and the three sites for miR-15/16 in the *Cyclin D2* 3′UTR were mutated in the context of the long isoform (Figure 1B). Luciferase activity was examined, comparing activity from reporters with intact sites with that from those with mutant sites using the endogenous expression levels of the miRNAs. Loss of these miRNA sites led to expression increases of 1.0- to 1.8-fold across 16 cell lines from various tissues (*IMP-1* median, 1.4, range, 1.0–1.6; *Cyclin D2* median, 1.3, range, 1.1–1.8; *DICER1* median, 1.2, range, 1.0–1.5)—rather modest effects when considering the number of sites mutated (Figures 4A–4C). When the endogenous miRNAs were supplemented by transfecting more of the cognate miRNAs, the reporter gene expression differences increased up to 3.3-fold (data not shown), implying regulatory potential beyond that mediated by the miRNAs at their endogenous levels and

Table 2. mRNA Half-Life

Cancer Tissue	Cell Line	Long mRNA $t_{1/2}$ (hr)[a]	Short mRNA $t_{1/2}$ (hr)[a]	Short/Long
IMP-1				
Lung	A549	4.2 ± 1.4	9.0 ± 2.5	2.1
Colon	SW480	6.2 ± 0.4	12.3 ± 4.3	2.0
Breast	HBL100	2.3 ± 0.9	6.4 ± 1.0	2.8
Breast	MDA-MB231	4.1 ± 0.5	6.7 ± 4.0	1.6
Breast	MDA-MB361	1.8 ± 0.3	4.5 ± 1.6	2.5
Cyclin D1				
Breast	MCF-7	1.9 ± 0.2	3.8 ± 0.7	2.0
Breast	MDA-MB361	8.0 ± 0.1	22.4 ± 6.3	2.8
Breast	MDA-MB453	1.5 ± 0.5	6.3 ± 2.7	4.2
Cyclin D2				
Sarcoma	U2OS	2.3 ± 0.2	3.5 ± 1.1	1.5
Colon	SW480	3.1 ± 0.9	20.6 ± 12.4	6.6
Colon	LoVo	13.8 ± 1.4	9.3 ± 1.8	0.7

[a] Reported is the mean ± standard deviation.

suggesting that the variability in response to the endogenous miRNAs might depend on expression levels of the cognate miRNAs. Quantitative northern blots of miRNAs across the cell lines showed that, overall, *let-7* levels were the highest, miR-15/16 levels were intermediate, and miR-103/107 levels were about 5 times lower than those of *let-7* (Figure S1). In addition to its six miR-103/107 sites, the *DICER1* 3′UTR has two *let-7* sites. Mutating these sites led to a median expression increase of 1.3 (range 1.0- to 1.8-fold), which was higher than that of mutating the six miR-103/107 sites, consistent with the idea that *let-7*, based on its higher expression, might have been more effective in these cells (Figure 4D).

To find additional miRNAs endogenously expressed in these cell lines and to measure the relative expression of the miRNAs examined, we profiled the miRNAs by high-throughput sequencing. The *let-7* miRNA family was either the highest or the second highest sequenced family in the investigated cell lines (with the exception of HEK293 cells where it ranked eighth highest) and corresponded to 3.3%–34% (median, 16.8%) of all miRNA sequencing reads in the cells (Table S1). Consistent with the northern blots, miR-103/107 was sequenced less frequently, ranking 3–16 of all miRNA families expressed in these cells and corresponding to 0.74%–5.8% of the reads (median, 2.4%). Considering the expression levels of the different miRNAs in the different cell types, variability in miRNA expression explained a significant fraction of the variability in reporter expression (r^2 = 0.20, p = 0.001, Figure S4). By taking this correlation into account and extrapolating to the other miRNA sites (conserved and non-conserved) predicted by TargetScan to match coexpressed miRNAs, we estimated the maximum repression attributable to miRNAs to range from 1.9- to 4.1-fold for the three 3′UTRs in the 12 cell lines with profiled miRNAs (*IMP-1* median, 2.8, range, 2.5–3.5; *DICER1* median, 2.7, range, 1.9–3.2; *Cyclin D2* median, 2.4, range, 1.9–4.1; Figure S4 and Table S2). These rough estimates were typically lower than the average increases in luciferase reporter gene expression for the short 3′UTR in these 12 lines (*IMP-1*, 7.7-fold, *DICER1*: 6.6-fold, *Cyclin D2*: 15.3-fold). Thus, loss of miRNA regulation appears to account for a quarter to two-thirds of the increase in protein-expression levels observed with the short 3′UTRs. In 20%–40% of the cell

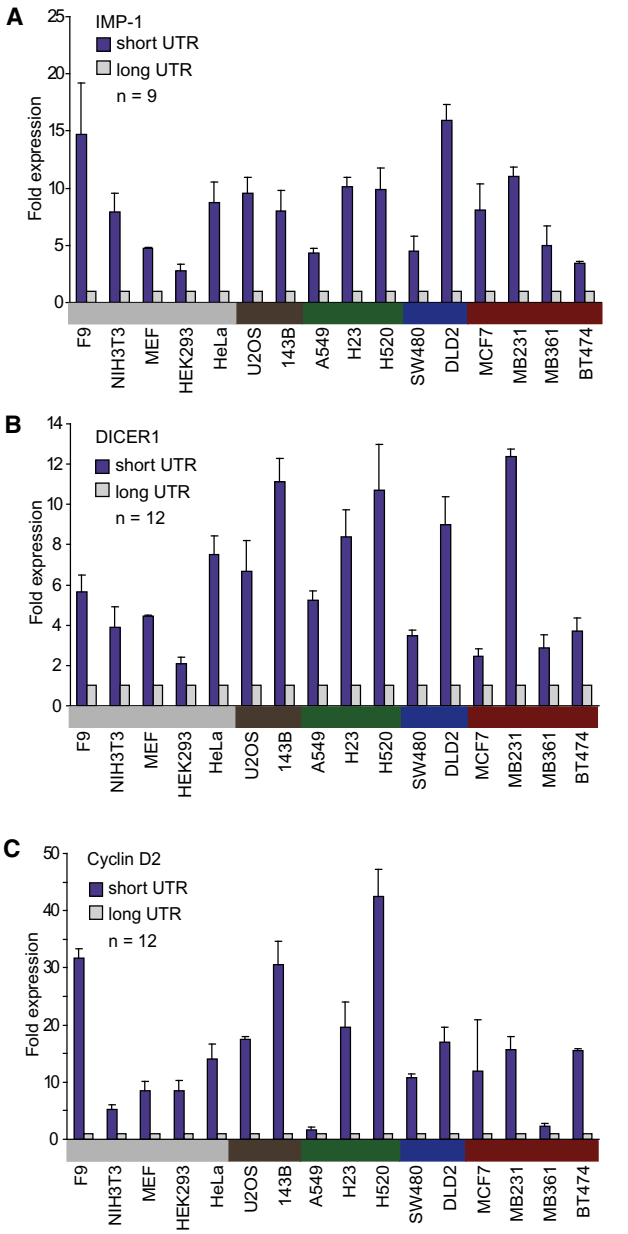

Figure 3. The Shortest mRNA Isoform Leads to Higher Protein Expression Than Does the Full-Length Isoform

(A) Luciferase expression from a reporter containing the 3′UTR of the shortest *IMP-1* isoform, compared to that from the reporter containing the 3′UTR of the full-length *IMP-1* isoform. Assays were performed in 16 cell lines from various tissues, color-coded as in Figure 1A. The number of transfections is shown (n = 9, performed in three independent experiments, each with three transfections of each reporter), with error bars indicating the standard deviation for these transfections.

(B) As in (A), but using reporters with the *DICER1* 3′UTRs.

(C) As in (A), but using reporters with the *Cyclin D2* 3′UTRs.

lines (dependent on the gene investigated) loss of miRNA regulation might even explain all the difference in protein expression. However, for the majority of cell lines miRNA regulation explained only a part of the difference in protein expression between the short and the long mRNAs, indicating a net repressive effect of other regulatory phenomena lost in the shorter UTRs as a consequence of APA.

Impact of 3′UTR Sequence versus 3′UTR Length

One explanation for the repression not attributable to miRNA targeting might be the length of the mRNA itself. The difference in length between the shorter and the longer mRNA isoforms was substantial (*IMP-1*, 6.3 kb, or 71% of the mRNA; *DICER1*, 4.3 kb, or 42% of the mRNA; *Cyclin D2*, 3.5 kb, or 54% of the mRNA). Shorter mRNAs might more readily form a closed-loop structure, which enhances translation, as shown recently in yeast (Amrani et al., 2008). To find out if the length or specific sequences within the longer mRNA were responsible for the difference in luciferase activity, the reverse complement (rc) of the UTR segment present in the long but not the short *IMP-1* mRNA isoform was cloned downstream the short 3′UTR, while disrupting the proximal poly(A) signal. In four of the 15 cell lines tested, the short and rc reporters had comparable protein production, which averaged four times higher than that from a reporter with the long 3′UTR, suggesting that repressive sequences recognized by *trans*-acting factors present preferentially in these cell lines specify translational repression or mRNA destabilization (Figure 4E). In the other cell lines, however, expression of the rc construct led to intermediate gene expression. Although 3′UTR length might have some influence on gene expression in these cell lines, our results for the other four lines speak against this as a general phenomenon in human cells.

The Shorter *Cyclin D2* mRNA Leads to More Cells in S Phase

Having found that the shorter mRNA isoforms were more stable and led to higher protein expression, we wanted to test if there might be a phenotypic consequence of expressing the shorter rather than the longer isoform. For *Cyclin D2*, we expressed either the full-length mRNA or its shorter isoform in breast cancer cell lines. *Cyclin D2* is not expressed endogenously in these cell lines because of promoter hypermethylation (Evron et al., 2001). In MCF7 cells, we observed an effect on the cell-cycle profile, with a lower fraction of cells in G1 accompanied by a higher fraction of cells that had entered S phase (Figure 5A), which was concordant with the function of *Cyclin D2* in overcoming the G1/S checkpoint (Lukas et al., 1995). This effect was significantly greater in cells expressing the short mRNA of *Cyclin D2* compared with those expressing the long mRNA, and correlated with *Cyclin D2* mRNA and protein expression in these cells (Figure S5).

The Shorter *IMP-1* mRNA Is Oncogenic

IMP-1 is an RNA-binding protein that is overexpressed in a variety of human cancers, including those of the breast, lung and colon, where it potentially plays a role in stabilizing several target mRNAs, including c-*myc,* β-*catenin* and β-*TrCP1* (Ross et al., 1997; Doyle et al., 1998; Nielsen et al., 1999; Noubissi et al., 2006). To learn the relative effects of the *IMP-1* isoforms, we assayed colony formation in soft-agar after transducing cells with retroviral vectors expressing either the short or long mRNA isoform. Expressing the full-length *IMP-1* isoform (with mutations designed to prevent APA) produced a number of colonies comparable to that of expressing the empty vector ($p = 0.09$), whereas expressing the short isoform greatly promoted cell transformation ($p = 0.002$, Figure 5B). Much of this transformation was attributable to loss of miRNA targeting, in that mutating the *let-7* sites significantly enhanced transformation from the long isoform ($p = 0.002$) and correlated with IMP-1 mRNA and protein expression in these cells (Figure S6). The shorter isoform was not only able to transform fibroblasts but also human breast epithelial cell lines (Figure S7). Taken together, our results indicate that APA associated with cancer cells can activate an oncogene, thereby potentially contributing to the pathogenesis of cancer.

DISCUSSION

Recent reports propose that a change in 3′UTR length by means of APA is a coordinated mechanism for altering expression of many genes during T cell activation (Sandberg et al., 2008), neuronal activation (Flavell et al., 2008) or embryonic development (Ji et al., 2009). For a better understanding of gene regulation, it will be important to identify additional cellular, developmental and disease states that lead to increased use of proximal APA signals. We observed shorter 3′UTRs in cancer cell lines compared with nontransformed cell lines, despite similar proliferation rates of the transformed and nontransformed lines, thereby linking 3′UTR shortening with oncogenic transformation even more than with cellular proliferation.

Many proto-oncogenes play roles in proliferation and differentiation of normal cells and must be highly regulated to prevent malignant transformation. In normal cells, the full-length, annotated transcripts were expressed and presumably responsible for normal proto-oncogene function, but in cancer cells, substantially more shorter isoforms were also expressed, which typically differed from the full-length mRNAs only in the length of their 3′UTRs. Our reporter assays and immunoblots revealed that the APA-mediated 3′UTR shortening can have striking functional consequences in the cancer cell lines, with the shorter mRNA isoforms typically producing ten-times more protein. These results suggested that the cancer-associated shortening of 3′UTRs could activate oncogenes, thereby reinforcing the transformed state. In support of this hypothesis, we found that the shorter *IMP-1* isoform promoted oncogenic transformation far more efficiently than did the longer one.

The oncogenic potential of *IMP-1* was previously demonstrated using a transgenic mouse model with *IMP-1* under the promoter of whey acidic protein, which is induced in mammary epithelial cells in pregnant and lactating female mice (Tessier et al., 2004). The construct overexpressing *IMP-1* in these mice contained only the first 288 nt of the 6.3 kb annotated 3′UTR. Our results, which showed both that cancer cells endogenously express a shorter isoform with a 369-nt 3′UTR (with the orthologous mouse isoform estimated to have a 336-nt 3′UTR) and that this shorter isoform is the one that was oncogenic, support the in vivo relevance of the previous transgenic model.

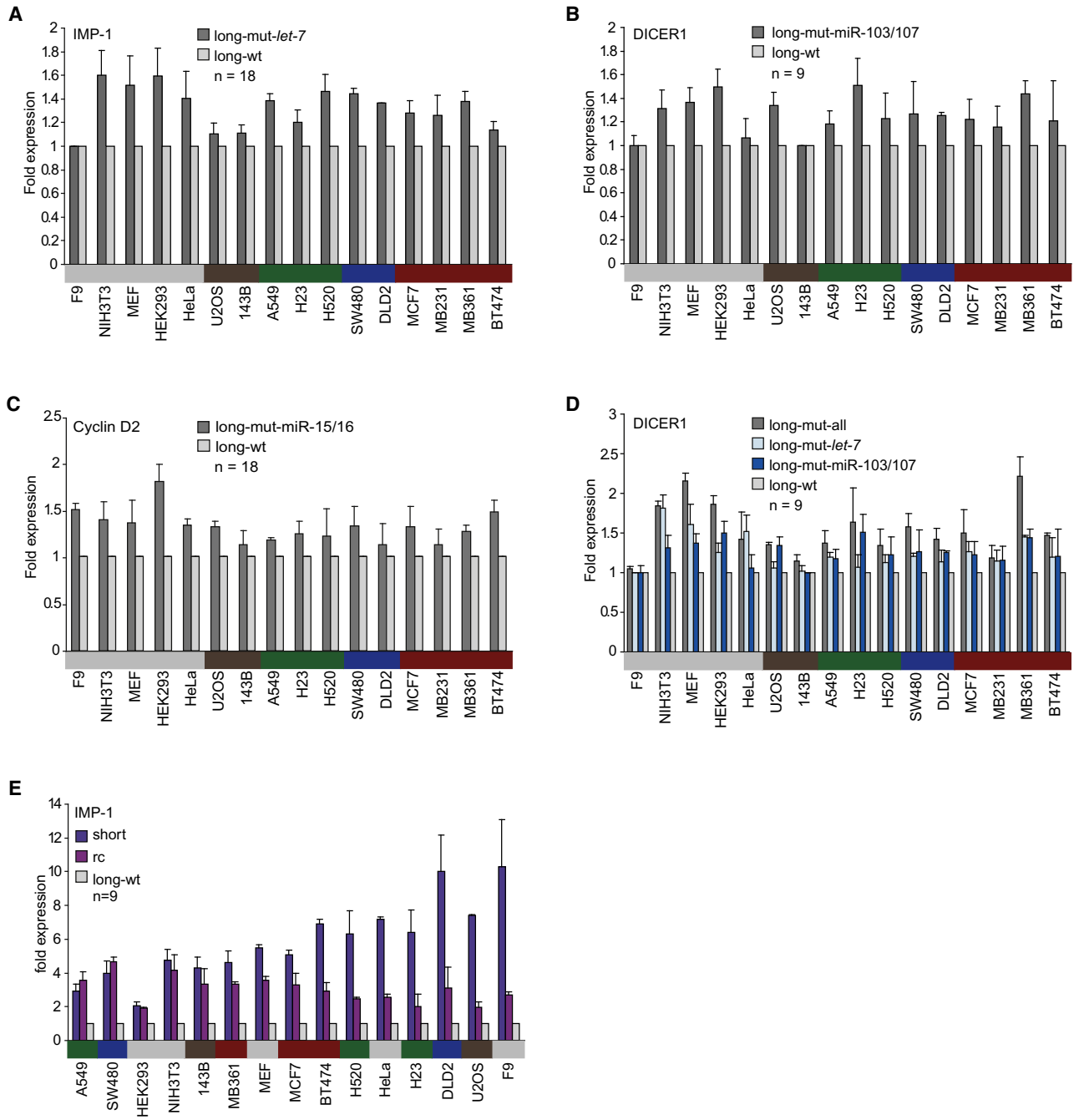

Figure 4. Contributions of miRNA Regulation and 3′UTR Length to the Difference in Protein Expression Observed between the Long and the Short mRNA Isoforms

(A) Repression of the *IMP-1* 3′UTR by the *let-7* miRNA endogenously expressed within the cells. Luciferase expression of a reporter possessing the full-length *IMP-1* 3′UTR with five mutant *let-7* sites (long-mut-*let-7*) is compared with that of a reporter with five intact *let-7* sites (long-wt, Figure 1B). The number of transfections is shown (n = 18, performed in six independent experiments, each with three transfections of each reporter), with error bars indicating standard deviation. Cells are color-coded as in Figure 1A.

(B) Repression of the *DICER1* 3′UTR by the miR-103/107 endogenously expressed within the cells. As in (A) but examining the contribution of the six miR-103/107 sites in the full-length *DICER1* 3′UTR.

(C) Repression of the *Cyclin D2* 3′UTR by the miR-15/16 endogenously expressed within the cells. As in (A) but examining the contribution of the three miR-15/16 sites in the full-length *Cyclin D2* 3′UTR.

(D) Repression of the *DICER1* 3′UTR by the *let-7* miRNA endogenously expressed within the cells. As in (B) but examining the contribution of the two *let-7* sites, disrupting either only the two *let-7* sites (long-mut-*let-7*) or both the two *let-7* sites and the six miR-103/107 sites (long-mut-all).

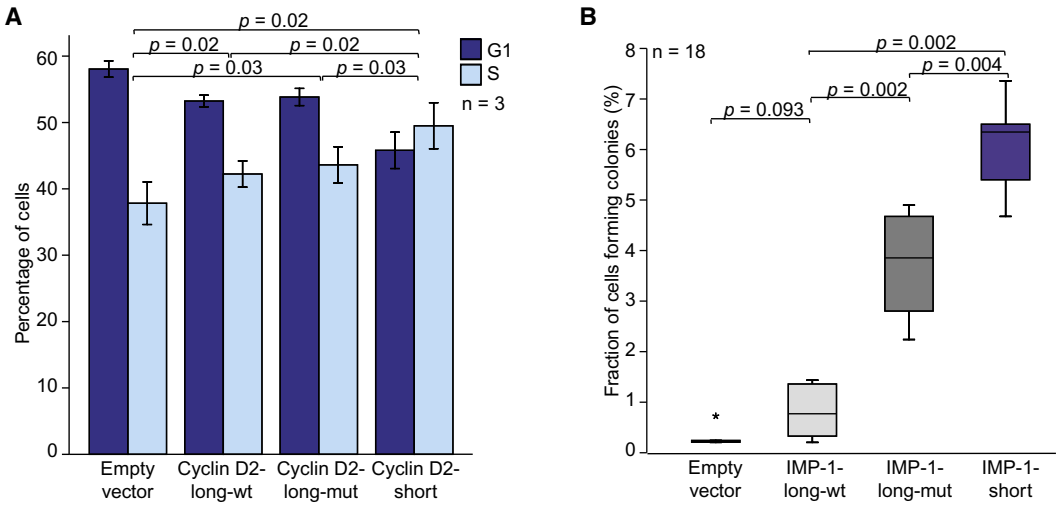

Figure 5. Functional Consequences of Expressing Short or Long mRNA Isoforms

(A) The short mRNA isoform of *Cyclin D2* increases the fraction of cells in S phase. For MCF7 cells stably transduced with retroviral vectors, the percentage of cells that were in G1 and S phase were measured using fluorescence-activated cell sorting (FACS) analysis. Plotted is the mean (±SD) of three independent experiments.

(B) The short mRNA isoform of *IMP-1* promotes oncogenic transformation.

For NIH 3T3 cells stably transduced with retroviral vectors, the percentage that yielded colonies after 28 days was plotted (horizontal line, median; box, 25th through 75th percentile; error bars, range; asterisk, outliers; n = 18 from six independent experiments, each in triplicate). The vectors expressed either the long mRNA isoform of *IMP-1* with mutations designed to prevent APA but retaining intact *let-7* sites (IMP-1-long-wt) or the analogous long mRNA isoform with mutant *let-7* sites (IMP-1-long-mut) or the shortest mRNA isoform of *IMP-1* (IMP-1-short).

Part of the *IMP-1* oncogene activation observed in our soft-agar assays was explained by escape of the shorter isoform from miRNA-mediated repression. Since the discovery of miRNAs in mammalian cells, much has been learned about their roles in tumorigenesis. Some miRNAs, including the miR-17∼92 cluster are amplified in human cancers and act as oncogenes when overexpressed in mice (Ota et al., 2004; He et al., 2005). Others, including miR-15/16, miR-34 and the *let-7* miRNAs, are deleted or downregulated in cancers and are reported to act as tumor-suppressor genes (Calin et al., 2002; Johnson et al., 2005; Mayr et al., 2007; Yu et al., 2007; He et al., 2007). These miRNAs playing tumor-suppressor roles can do so only when their targets retain the cognate regulatory sites. Thus, another important mechanism for oncogenic transformation is the loss of miRNA sites in the mRNAs of oncogenes. This loss can occur through genetic aberrations, such as the translocations that abrogate *let-7* repression of the *HMGA2* oncogene (Mayr et al., 2007; Lee and Dutta, 2007). Our results uncovered an epigenetic mechanism (using the broader sense of the word "epigenetic") that achieved the same effect: oncogenes can escape miRNA-mediated repression through 3′UTR shortening due to the APA prevalent in cancer cells. Our data suggests that this epigenetic mechanism is more prevalent than is the genetic one because cancer cells had the shortest UTRs as indicated by their low TLI (a transcriptome-wide index), and APA generated the different transcripts in over a third of the expressed genes examined in detail.

Escape from miRNA-mediated repression explained only part of the upregulation conferred by APA. In addition to miRNAs, AU-rich or GU-rich elements can specify repression (Chen and Shyu, 1995; Vlasova et al., 2008), but a search of the tested 3′UTRs did not reveal these elements. Thus, our results implicated additional, as-yet-undefined regulatory elements in these 3′UTRs. Interestingly, the net effect of these elements was never activating—whenever an effect was observed, it was repressive. Although activating regulatory elements undoubtedly occur in some 3′UTRs and the net effect of all RNA elements might be activating in some 3′UTRs or cell types, our results support the idea that regulatory phenomena acting on 3′UTRs are generally repressive. This conclusion holds despite the bias in selecting for study 3′UTRs with multiple sites to coexpressed miRNAs, because even after all those sites were mutated, we still observed only repression. The overall repressive character of 3′UTRs could help explain why nearly all successful transgenic mouse models of tumorigenesis overexpress oncogenes missing large segments of their annotated 3′UTRs (Ruther et al., 1987; Fedele et al., 2002; Primrose and Twyman, 2006). As with *IMP-1*, our results indicate that these models expressing shorter versions of the mRNA are more biologically relevant than might have been anticipated when they were first generated.

One of the most interesting open questions regarding APA in cancer cells is, what mechanism underlies the recognition and increased utilization of proximal polyadenylation signals? In the

(E) Influence of the length of the *IMP-1* 3′UTR on protein expression. The reverse complement (rc) of the difference between the shortest and the full-length mRNA isoforms of *IMP-1* was cloned downstream the shortest 3′UTR, while mutating the proximal poly(A) signal. Expression of the reporter possessing the shortest 3′UTR and of that also possessing the rc, each normalized to expression of the reporter containing the full-length isoform. The number of transfections is shown (n = 9, performed in three independent experiments, each with three transfections of each reporter), with error bars indicating standard deviation.

sequence surrounding the proximal polyadenylation sites we never found point mutations that would have changed the strength of the polyadenylation signal, with the caveat that by 3′ RACE we investigated only the sequences upstream of the cleavage sites. Although mutations downstream of the cleavage sites cannot be excluded, we hypothesize that differential expression of *trans*-acting factors explains the use of proximal polyadenylation sites in cancer cells. Factors that might influence the choice of polyadenylation signal include limiting components of the polyadenylation machinery, RNA-binding proteins that bind in the vicinity of the proximal signal and influence recognition by the polyadenylation machinery (Takagaki et al., 1996; Martincic et al., 1998; Veraldi et al., 2001; Lutz, 2008; Wang et al., 2008), and perhaps factors that influence transcriptional elongation rate (Kornblihtt, 2005). To begin to identify such factors, we examined published array data comparing breast cancer cells with a breast epithelial cell line, MCF10A (Hoeflich et al., 2009). A survey of the constitutive components of the polyadenylation machinery and other candidates from the literature revealed several that were significantly upregulated in the cancer lines (Figure S8). These included the mRNAs of cleavage and polyadenylation specificity factor 1 (CPSF1) and cleavage stimulation factor 2 (CSTF2), which recognize the poly(A) signal and accessory sequences including the downstream G/U-rich sequence, respectively, raising the intriguing possibility that an increase of these factors might help increase utilization of suboptimal proximal poly(A) signals in cancer cells.

We imagine a complex scenario in which some *trans*-acting factors act globally, some act tissue specifically, and some act gene specifically, with the combinatorial expression of all the different *trans*-acting factors determining the probability of using each proximal polyadenylation signal. The observation of higher amounts of shorter mRNAs in cancer cells compared with normal cells, with no examples of the reverse for any of the genes and cell types studied, suggested a role for globally acting factors. That some cell lines showed high amounts of shorter transcripts for all genes investigated further implicated the role of globally acting factors. Nonetheless, the differences observed for different genes in different cell types suggested a role for additional factors acting more specifically. Such complexity could explain the differential impact of different oncogenes in different tissues. Indeed, some of the genes for which we did not observe alternative mRNAs are known oncogenes in tissues not included in our panel of cell lines (*ARID3B* in neuroblastomas, *MYB* and *PLAG1* in hematological malignancies). Perhaps shorter mRNA transcripts might be found in the tissues where these genes have oncogenic effects. Moreover, the prospect of some factors acting more specifically opens the possibility for exceptions to the trend of shorter isoforms expressed preferentially in cancer cells. Combinatorial use of tissue-specific and gene-specific *trans*-acting factors could for some genes (most intriguingly, for tumor-suppressor genes) lead to higher amounts of shorter mRNAs in normal cells rather than in cancer cells.

The prevalence of APA in cancer lines brought to the fore the question of what influence it might have on oncogenes, and our results for *IMP-1* supported the idea that APA was activating. However, APA creates shorter, less repressed isoforms of more than just known oncogenes, as illustrated for *DICER1*. Because some of these genes likely act in opposition to oncogenes, the net functional significance of APA in cancer cells is unknown, and in principle could even be tumor suppressive. During both normal development and cancer development there is often a dichotomy of cell proliferation and differentiation (Derynck and Wagner, 1995). The association of APA with cell proliferation suggests that it might represent a coordinated gene-expression program that antagonizes differentiation during normal development. Accordingly, we propose that cancer cells coopt and exaggerate this proliferation/de-differentiation program with the net effect of enhanced tumorigenesis.

Our observations that the shorter mRNAs were found in transformed cells and that expression of the shorter mRNA of IMP-1 (and presumably other oncogenes) can lead to transformation suggests a APA-mediated feed-forward loop in cancer that might lead to a more aggressive phenotype. This proposal is in agreement with the report on mantle cell lymphomas, in which the patients that have shorter *Cyclin D1* 3′UTRs have the worst prognosis (Rosenwald et al., 2003; Wiestner et al., 2007). However, when this process of generating shorter mRNAs would start and whether APA might play a role in early tumorigenesis is unclear. In general, shorter mRNAs were not observed in non-transformed cell lines, which usually already have one hit toward cancer but are not yet fully transformed. Perhaps if an early lesion activates a signaling pathway that is able to change the expression of a key *trans*-acting factor, then APA would contribute to early steps in oncogenesis.

Oncogenes are reported to be overexpressed in human tumors much more frequently than genetic lesions are detected at these loci. Our results could explain some of this discrepancy, because overexpression of oncogenes due to APA-mediated shortening of mRNA transcripts was a widespread phenomenon in the cancer cells investigated. The epigenetic nature of this mechanism for oncogene activation suggests that it can be reversed, perhaps providing a new strategy for cancer treatment.

EXPERIMENTAL PROCEDURES

Northern Blots

Total RNA was isolated using Tri-reagent (Ambion). miRNAs were detected as described, loading 20 μg per lane (Mayr et al., 2007). mRNA was detected using a protocol adapted from Cold Spring Harbor Protocols (Sambrook and Russell, 2006), loading in each lane 1.5–2 μg polyadenylated RNA purified from total RNA with Oligotex (QIAGEN).

mRNA Stability

Cell lines were treated with Actinomycin D (10 μg/ml), and after 0, 2, 4, 6, and 8 hr total RNA was isolated and analyzed on northern blots.

3′ RACE

This protocol was adapted from Cold Spring Harbor Protocols (Sambrook and Russell, 2006). 1 μg total RNA was used to generate cDNA with Superscript II reverse transcriptase (Invitrogen) according to the manufacturer's instructions using TAP as a primer. The first PCR was done with a gene-specific forward primer and AP as reverse primer. Nested PCR was done with a nested gene-specific forward primer and MAP as reverse primer. The PCR product from the nested PCR was separated on an agarose gel, cloned and sequenced. If only full-length mRNAs were found by 3′ RACE, northern blots were reprobed with different probes in the ORF to find alternatively spliced isoforms.

Luciferase Assays

Luciferase assays were performed as described (Mayr et al., 2007). To account for differences in plasmid preparations, values for each wild-type reporter were normalized to those for its cognate mutant-site reporter by using values from cells that lacked or expressed the relevant miRNA at low levels (F9 cells for the *let-7* miRNA, DLD2 cells for miR-15/16, and 143B for miR-103/107).

Small RNA Cloning

Small RNAs were cloned and analyzed as described previously (Grimson et al., 2008). The miRNA cloning data have been deposited at the NCBI Gene Expression Omnibus (GEO) repository (GSE16579).

Cell Cycle Analysis by FACS

Cells were plated at comparable densities, harvested after 48h, fixed with ethanol and stained with propidium iodide (50 μg/ml) and RNase A (40 U/ml) and DNA content was measured on a FACS Calibur HTS (Becton Dickinson). The percentage of diploid cells in G1, S and G2 were analyzed by ModFitLT V3.1.

Soft-Agar Assay

Soft-agar assays were performed as described (Mayr et al., 2007).

Statistics

The Kruskal-Wallis test was used to analyze the difference between several independent subgroups. Mann Whitney test was applied to analyze the difference between two independent supgroups. The Wilcoxon test was used to make pairwise comparisons using SPSS 14.0.

SUPPLEMENTAL DATA

Supplemental Data include eight figures, two tables, Supplemental Experimental Procedures, and Supplemental References and can be found with this article online at http://www.cell.com/supplemental/S0092-8674(09)00716-8.

ACKNOWLEDGMENTS

We thank Calvin Jan (C.J.) and Andrew Grimson (A.G.) for very helpful advice throughout the project; Wendy Johnston for preparing the small RNA libraries for the cell lines; Gina Lafkas for help with the soft-agar assay; A.G. and Rosaria Chang for analyzing the Solexa reads; and Olivia Rissland, A.G., C.J., and Alena Shkumatava for helpful comments on the manuscript. Supported by grants from the Deutsche Forschungsgemeinschaft (C.M.) and the National Institutes of Health (D.B.). D.B. is a Howard Hughes Medical Institute Investigator.

Received: January 12, 2009
Revised: April 21, 2009
Accepted: June 9, 2009
Published: August 20, 2009

REFERENCES

Amrani, N., Ghosh, S., Mangus, D.A., and Jacobson, A. (2008). Translation factors promote the formation of two states of the closed-loop mRNP. Nature *453*, 1276–1280.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215–233.

Bishop, J.M. (1991). Molecular themes in oncogenesis. Cell *64*, 235–248.

Calin, G.A., Dumitru, C.D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., et al. (2002). Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc. Natl. Acad. Sci. USA *99*, 15524–15529.

Chen, C.Y., and Shyu, A.B. (1995). AU-rich elements: characterization and importance in mRNA degradation. Trends Biochem. Sci. *20*, 465–470.

Coleman, M.S., Hutton, J.J., and Bollum, F.J. (1974). Terminal riboadenylate transferase in human lymphocytes. Nature *248*, 407–409.

Colgan, D.F., and Manley, J.L. (1997). Mechanism and regulation of mRNA polyadenylation. Genes Dev. *11*, 2755–2766.

Derynck, R., and Wagner, E.F. (1995). Cell differentiation. Curr. Opin. Cell Biol. *7*, 843–844.

Doyle, G.A., Betz, N.A., Leeds, P.F., Fleisig, A.J., Prokipcak, R.D., and Ross, J. (1998). The c-*myc* coding region determinant-binding protein: a member of a family of KH domain RNA-binding proteins. Nucleic Acids Res. *26*, 5036–5044.

Edwalds-Gilbert, G., Veraldi, K.L., and Milcarek, C. (1997). Alternative poly(A) site selection in complex transcription units: means to an end? Nucleic Acids Res. *25*, 2547–2561.

Evron, E., Umbricht, C.B., Korz, D., Raman, V., Loeb, D.M., Niranjan, B., Buluwela, L., Weitzman, S.A., Marks, J., and Sukumar, S. (2001). Loss of cyclin D2 expression in the majority of breast cancers is associated with promoter hypermethylation. Cancer Res. *61*, 2782–2787.

Fedele, M., Battista, S., Kenyon, L., Baldassarre, G., Fidanza, V., Klein-Szanto, A.J., Parlow, A.F., Visone, R., Pierantoni, G.M., Outwater, E., et al. (2002). Overexpression of the *HMGA2* gene in transgenic mice leads to the onset of pituitary adenomas. Oncogene *21*, 3190–3198.

Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat. Rev. Genet. *9*, 102–114.

Flavell, S.W., Kim, T.K., Gray, J.M., Harmin, D.A., Hemberg, M., Hong, E.J., Markenscoff-Papadimitriou, E., Bear, D.M., and Greenberg, M.E. (2008). Genome-wide analysis of MEF2 transcriptional program reveals synaptic target genes and neuronal activity-dependent polyadenylation site selection. Neuron *60*, 1022–1038.

Grimson, A., Farh, K.K., Johnston, W.K., Garrett-Engele, P., Lim, L.P., and Bartel, D.P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol. Cell *27*, 91–105.

Grimson, A., Srivastava, M., Fahey, B., Woodcroft, B.J., Chiang, H.R., King, N., Degnan, B.M., Rokhsar, D.S., and Bartel, D.P. (2008). Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. Nature *455*, 1193–1197.

Hauser, H., Knippers, R., and Schafer, K.P. (1978). Increased rate of RNA-polyadenylation. An early response in Concanavalin A activated lymphocytes. Exp. Cell Res. *111*, 175–184.

He, L., Thomson, J.M., Hemann, M.T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S.W., Hannon, G.J., and Hammond, S.M. (2005). A microRNA polycistron as a potential human oncogene. Nature *435*, 828–833.

He, X., He, L., and Hannon, G.J. (2007). The guardian's little helper: microRNAs in the p53 tumor suppressor network. Cancer Res. *67*, 11099–11101.

Hoeflich, K.P., O'Brien, C., Boyd, Z., Cavet, G., et al. (2009). In vivo antitumor activity of MEK and phosphatidylinositol 3-kinase inhibitors in basal-like breast cancer models. Clin. Cancer Res. *15*, 4649–4664.

Ji, Z., Lee, J.Y., Pan, Z., Jiang, B., and Tian, B. (2009). Progressive lengthening of 3′ untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. Proc. Natl. Acad. Sci. USA *106*, 7028–7033.

Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D., and Slack, F.J. (2005). RAS is regulated by the *let-7* microRNA family. Cell *120*, 635–647.

Kornblihtt, A.R. (2005). Promoter usage and alternative splicing. Curr. Opin. Cell Biol. *17*, 262–268.

Lebwohl, D.E., Muise-Helmericks, R., Sepp-Lorenzino, L., Serve, S., Timaul, M., Bol, R., Borgen, P., and Rosen, N. (1994). A truncated cyclin D1 gene encodes a stable mRNA in a human breast cancer cell line. Oncogene *9*, 1925–1929.

Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. Cell *75*, 843–854.

Lee, Y.S., and Dutta, A. (2007). The tumor suppressor microRNA *let-7* represses the *HMGA2* oncogene. Genes Dev. *21*, 1025–1030.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15–20.

Lukas, J., Bartkova, J., Welcker, M., Petersen, O.W., Peters, G., Strauss, M., and Bartek, J. (1995). Cyclin D2 is a moderately oscillating nucleoprotein required for G1 phase progression in specific cell types. Oncogene *10*, 2125–2134.

Lutz, C.S. (2008). Alternative polyadenylation: a twist on mRNA 3′ end formation. ACS Chem. Biol. *3*, 609–617.

Majoros, W.H., and Ohler, U. (2007). Spatial preferences of microRNA targets in 3′ untranslated regions. BMC Genomics *8*, 152.

Martincic, K., Campbell, R., Edwalds-Gilbert, G., Souan, L., Lotze, M.T., and Milcarek, C. (1998). Increase in the 64-kDa subunit of the polyadenylation/cleavage stimulatory factor during the G0 to S phase transition. Proc. Natl. Acad. Sci. USA *95*, 11095–11100.

Mayr, C., Hemann, M.T., and Bartel, D.P. (2007). Disrupting the pairing between *let-7* and *Hmga2* enhances oncogenic transformation. Science *315*, 1576–1579.

Miyamoto, S., Chiorini, J.A., Urcelay, E., and Safer, B. (1996). Regulation of gene expression for translation initiation factor eIF-2 alpha: importance of the 3′ untranslated region. Biochem. J. *315*, 791–798.

Moore, M.J. (2005). From birth to death: the complex lives of eukaryotic mRNAs. Science *309*, 1514–1518.

Nielsen, J., Christiansen, J., Lykke-Andersen, J., Johnsen, A.H., Wewer, U.M., and Nielsen, F.C. (1999). A family of insulin-like growth factor II mRNA-binding proteins represses translation in late development. Mol. Cell. Biol. *19*, 1262–1270.

Noubissi, F.K., Elcheva, I., Bhatia, N., Shakoori, A., Ougolkov, A., Liu, J., Minamoto, T., Ross, J., Fuchs, S.Y., and Spiegelman, V.S. (2006). CRD-BP mediates stabilization of betaTrCP1 and c-*myc* mRNA in response to beta-catenin signalling. Nature *441*, 898–901.

Ota, A., Tagawa, H., Karnan, S., Tsuzuki, S., Karpas, A., Kira, S., Yoshida, Y., and Seto, M. (2004). Identification and characterization of a novel gene, C13orf25, as a target for 13q31-q32 amplification in malignant lymphoma. Cancer Res. *64*, 3087–3095.

Primrose, S.B., and Twyman, R.M. (2006). Principles of Gene Manipulation and Genomics (Blackwell Publishing).

Proudfoot, N. (1991). Poly(A) signals. Cell *64*, 671–674.

Rosenwald, A., Wright, G., Wiestner, A., Chan, W.C., Connors, J.M., Campo, E., Gascoyne, R.D., Grogan, T.M., Muller-Hermelink, H.K., Smeland, E.B., et al. (2003). The proliferation gene expression signature is a quantitative integrator of oncogenic events that predicts survival in mantle cell lymphoma. Cancer Cell *3*, 185–197.

Ross, A.F., Oleynikov, Y., Kislauskis, E.H., Taneja, K.L., and Singer, R.H. (1997). Characterization of a beta-actin mRNA zipcode-binding protein. Mol. Cell. Biol. *17*, 2158–2165.

Rowley, J.D. (2001). Chromosome translocations: dangerous liaisons revisited. Nat. Rev. Cancer *1*, 245–250.

Ruther, U., Garber, C., Komitowski, D., Muller, R., and Wagner, E.F. (1987). Deregulated c-*fos* expression interferes with normal bone development in transgenic mice. Nature *325*, 412–416.

Sambrook, J., and Russell, D.W. (2006). Separation of RNA according to size: electrophoresis of glyoxylated RNA through agarose gels. Cold Spring Harbor Protocols 2006:pdb.prot4057.

Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A., and Burge, C.B. (2008). Proliferating cells express mRNAs with shortened 3′ untranslated regions and fewer microRNA target sites. Science *320*, 1643–1647.

Takagaki, Y., Seipelt, R.L., Peterson, M.L., and Manley, J.L. (1996). The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. Cell *87*, 941–952.

Tessier, C.R., Doyle, G.A., Clark, B.A., Pitot, H.C., and Ross, J. (2004). Mammary tumor induction in transgenic mice expressing an RNA-binding protein. Cancer Res. *64*, 209–214.

Tian, B., Hu, J., Zhang, H., and Lutz, C.S. (2005). A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res. *33*, 201–212.

Veraldi, K.L., Arhin, G.K., Martincic, K., Chung-Ganster, L.H., Wilusz, J., and Milcarek, C. (2001). hnRNP F influences binding of a 64-kilodalton subunit of cleavage stimulation factor to mRNA precursors in mouse B cells. Mol. Cell. Biol. *21*, 1228–1238.

Vlasova, I.A., Tahoe, N.M., Fan, D., Larsson, O., Rattenbacher, B., Sternjohn, J.R., Vasdewani, J., Karypis, G., Reilly, C.S., Bitterman, P.B., and Bohjanen, P.R. (2008). Conserved GU-rich elements mediate mRNA decay by binding to CUG-binding protein 1. Mol. Cell *29*, 263–270.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470–476.

Wiestner, A., Tehrani, M., Chiorazzi, M., Wright, G., Gibellini, F., Nakayama, K., Liu, H., Rosenwald, A., Muller-Hermelink, H.K., Ott, G., et al. (2007). Point mutations and genomic deletions in *CCND1* create stable truncated cyclin D1 mRNAs that are associated with increased proliferation rate and shorter survival. Blood *109*, 4599–4606.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. Cell *75*, 855–862.

Yu, F., Yao, H., Zhu, P., Zhang, X., Pan, Q., Gong, C., Huang, Y., Hu, X., Su, F., Lieberman, J., and Song, E. (2007). *let-7* regulates self renewal and tumorigenicity of breast cancer cells. Cell *131*, 1109–1123.

Cell, Volume *138*

Supplemental Data

Widespread Shortening of 3′UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells

Christine Mayr and David P. Bartel

Figure S1

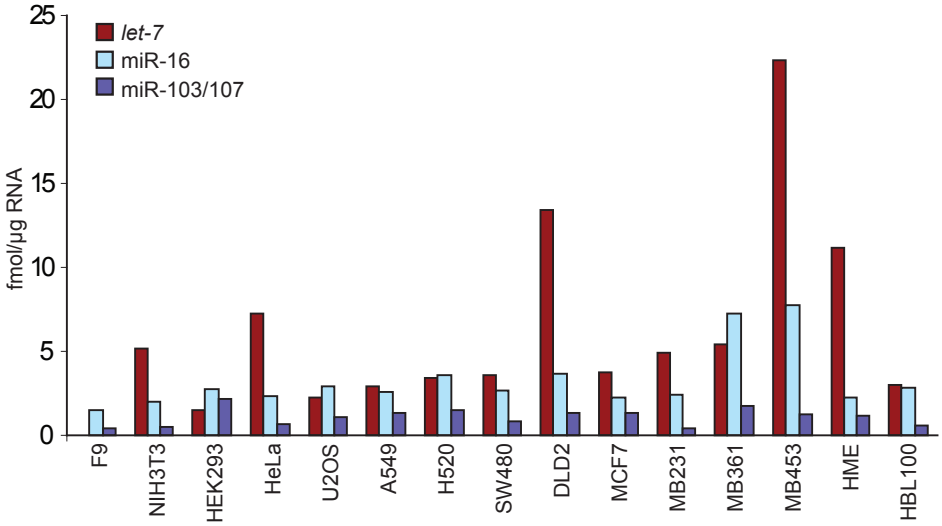

Figure S1. miRNA expression levels of *let-7*a–d, miR-16 and miR-103/107 in cell lines were determined by quantitative Northern blots.

Figure S2

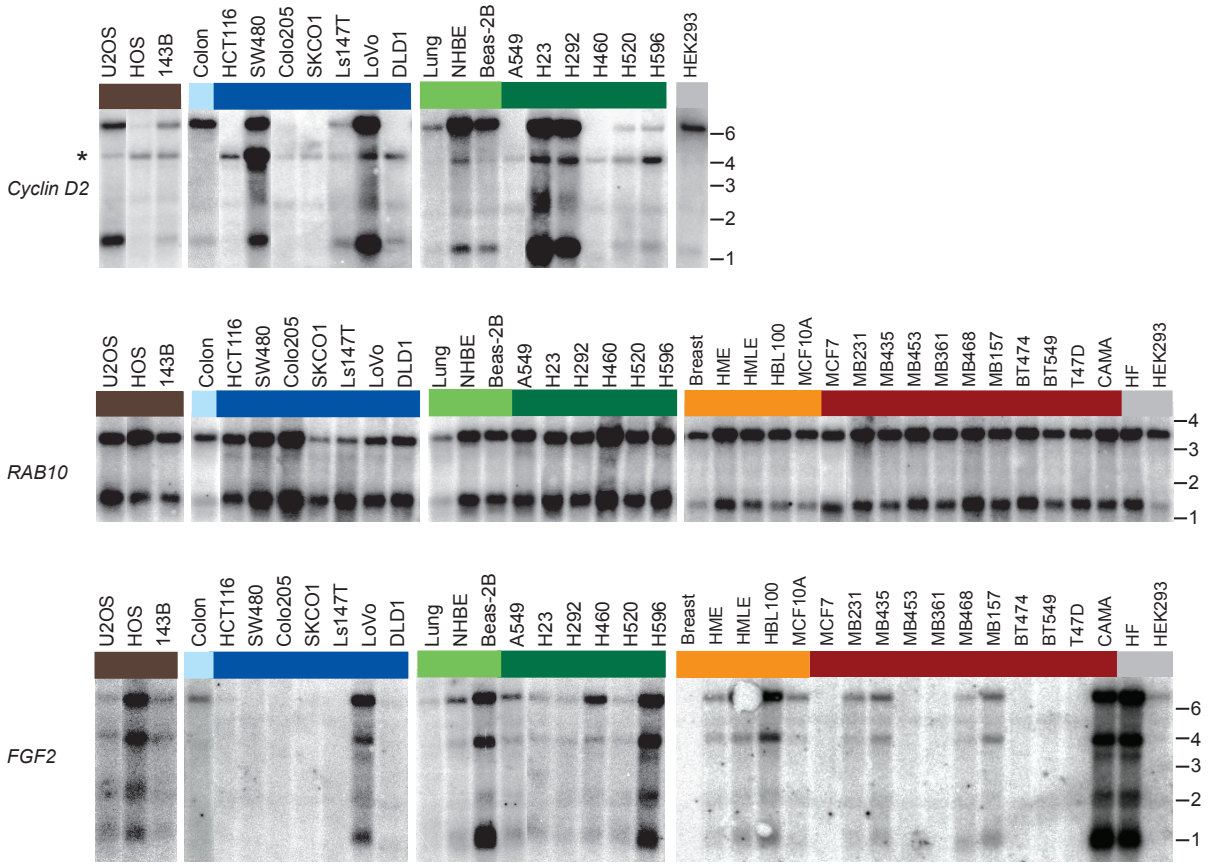

Figure S2. APA leads to shorter 3'UTRs in transformed cells.
(A) Northern blots of human cell lines and tissues. Sarcoma cell lines (brown), normal lung tissue, cultured human bronchial epithelial cells (NHBE) and immortalized lung epithelial cell line (Beas-2B, light green), lung cancer cell lines (dark green), normal colon tissue (light blue), colon cancer cell lines (dark blue), normal breast tissue and immortalized breast epithelial cell lines (orange), breast cancer cell lines (red), other cell lines (grey), immortalized fibroblasts (HF) and HEK293 (embryonic kidney cell line). In addition to the full-length annotated mRNA of *Cyclin D2* (6.5 kb), *RAB10* (3.5 kb) and *FGF2* (6.8 kb) shorter mRNAs are observed predominantly in cancer cells.
*, is cross hybridization to *Cyclin D1*.

Figure S3

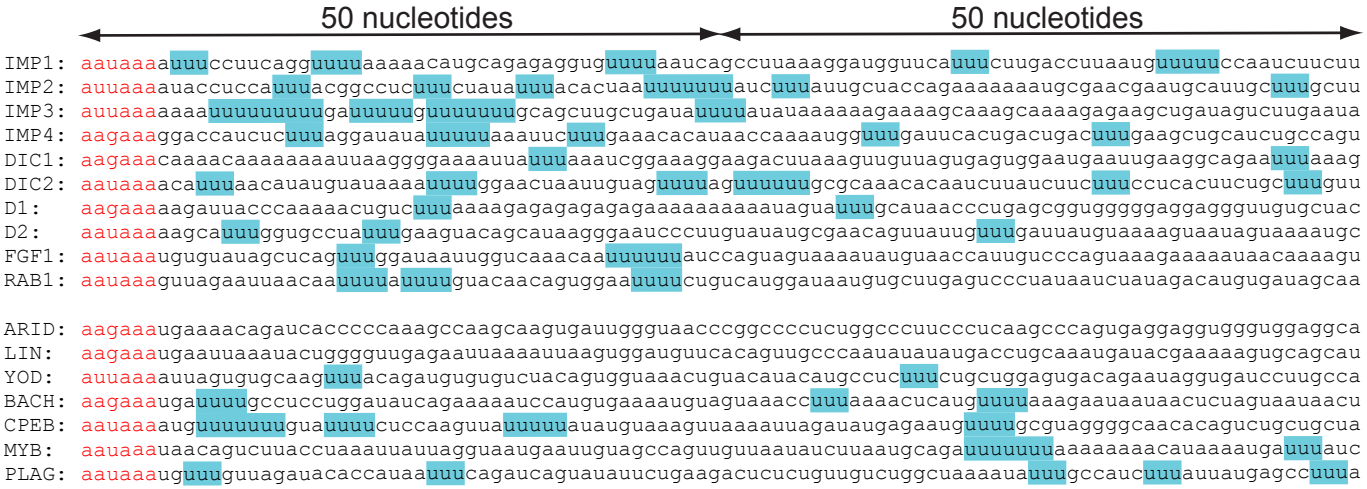

Figure S3. Sequence downstream of functional poly(A) signal.
Overrepresentation of U-rich sequences within 50 nucleotides of the APA signal (p = 0.02). Upper panel: genes with alternative mRNAs. Lower panel: genes without alternative mRNAs in cell lines and tissues investigated.

Figure S4

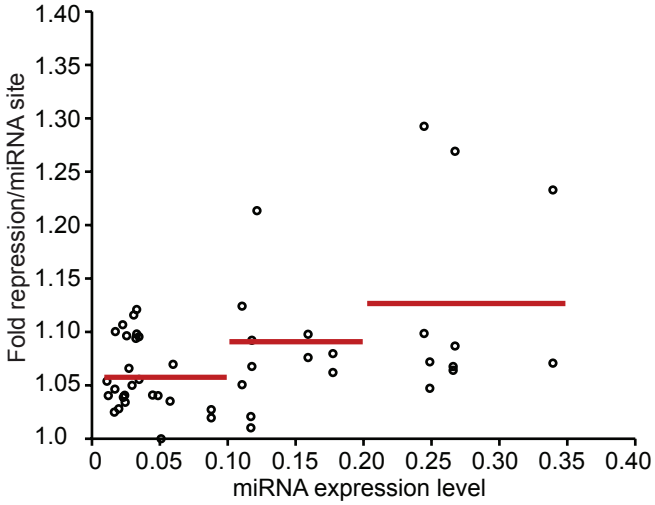

Figure S4. Correlation of endogenous miRNA expression in cell lines with the corresponding repression seen in luciferase reporter assays for three different 3'UTRs.

Relative miRNA expression was determined by small RNA sequencing. Luciferase reporter values were obtained by comparing reporter gene expression from constructs with intact miRNA sites and mutated miRNA sites. The repression per miRNA site was determined by taking the n^{th}-root of the total repression according to the number of miRNA sites (n) measured per 3'UTR. The average increase in luciferase expression per miRNA site (e) was determined by binning the relative miRNA expression in the tested cell lines (1–10%, 10–20% and >20%, red lines; Pearson's correlation coefficient, R = 0.43, p = 0.001).

The maximal increase in expression upon loss of all miRNA regulatory sites of miRNAs expressed higher than 1% in a given cell line was estimated by counting all conserved and non-conserved sites for a given 3'UTR as predicted by TargetScan for the miRNAs expressed higher than 1% in the cell line. According to the expression level of the miRNA, the corresponding luciferase reporter repression value (e) was applied. To determine the maximum repression, all possible repression values were multiplied.

Figure S5

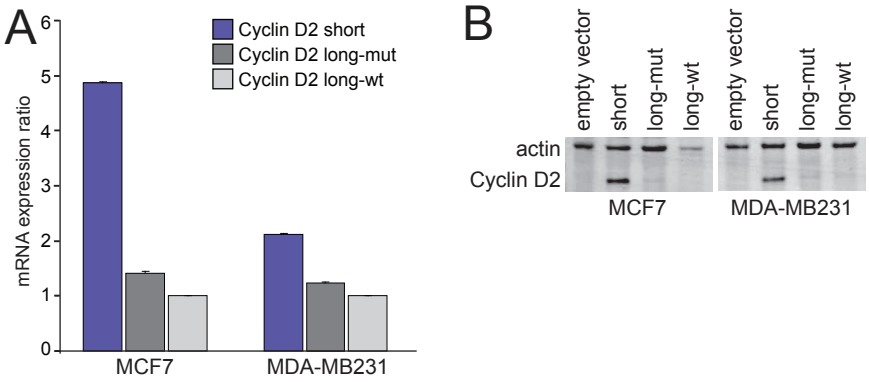

Figure S5. mRNA and protein expression of Cyclin D2 in breast cancer cell lines.
Figure S5A. mRNA expression of ectopically expressed mRNA isoforms of *Cyclin D2* measured by qRT-PCR. Shown is the average fold expression (± s. d.) of six experiments normalized to *GAPDH* and *Puromycin*, and the amount of Cyclin D2 expressed from the long isoform with wt miRNA sites (long-wt). Long-mut: the miRNA sites for miR-15/16 and *let-7* in the 3'UTR of Cyclin D2 were mutated.
Figure S5B. Protein expression of Cyclin D2 (samples as in Figure S5A).

Figure S6

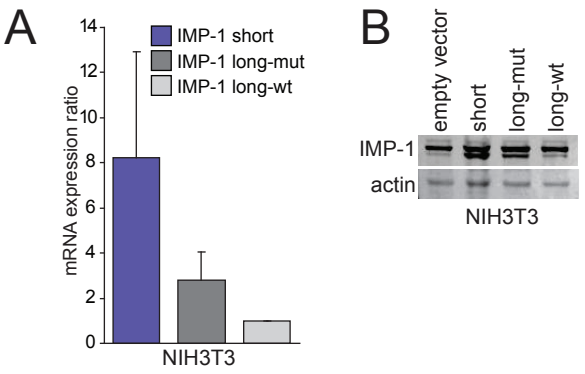

Figure S6. mRNA and protein expression of IMP-1 in NIH3T3 fibroblasts.
Figure S6A. mRNA expression of ectopically expressed mRNA isoforms of *IMP-1* measured by qRT-PCR. Shown is the average fold expression (± s. d.) of six experiments normalized to *GAPDH* and the amount of IMP-1 expressed from the long isoform with wild-type *let-7* sites (long-wt). Long-mut: mutated *let-7* sites.
Figure S6B. Protein expression of IMP-1 (samples as in Figure S6A). Endogenous IMP-1 is expressed as a doublet. Ectopic expression predominantly increases the lower migrating band.

Figure S7

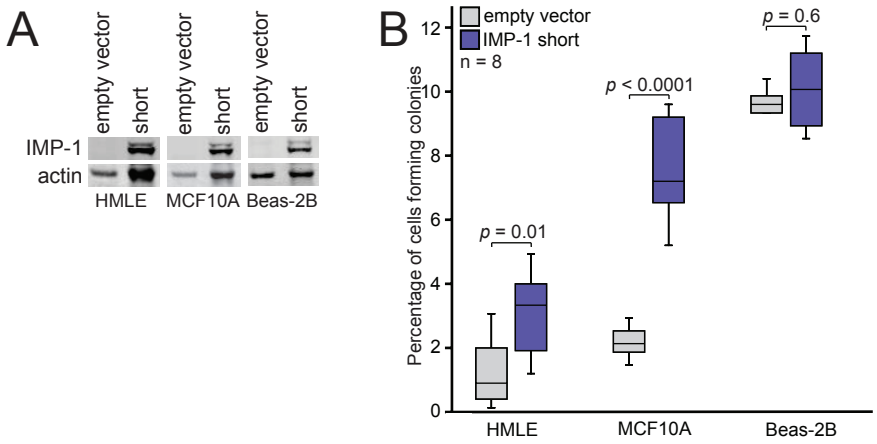

Figure S7. Protein expression of IMP-1 and oncogenic transformation of human epithelial cells by expression of the short mRNA of *IMP-1*.

Figure S7A. Protein expression of IMP-1 expressed from the short mRNA in human epithelial cells from breast (HMLE and MCF10A) and lung tissue (Beas-2B).

Figure S7B. Soft-agar assay of colony formation. The short mRNA isoform of *IMP-1* promotes oncogenic transformation. For breast epithelial cell lines stably transduced with retroviral vectors, the percentage that yielded colonies was plotted (median, horizontal line; 25th through 75th percentile, box; range, error bars; n = 8 from four independent experiments). The vectors expressed either the empty vector or the short mRNA isoform of *IMP-1*. In Beas-2B cells a high number of background colonies was observed that did not increase upon ectopic expression of IMP-1.

Figure S8

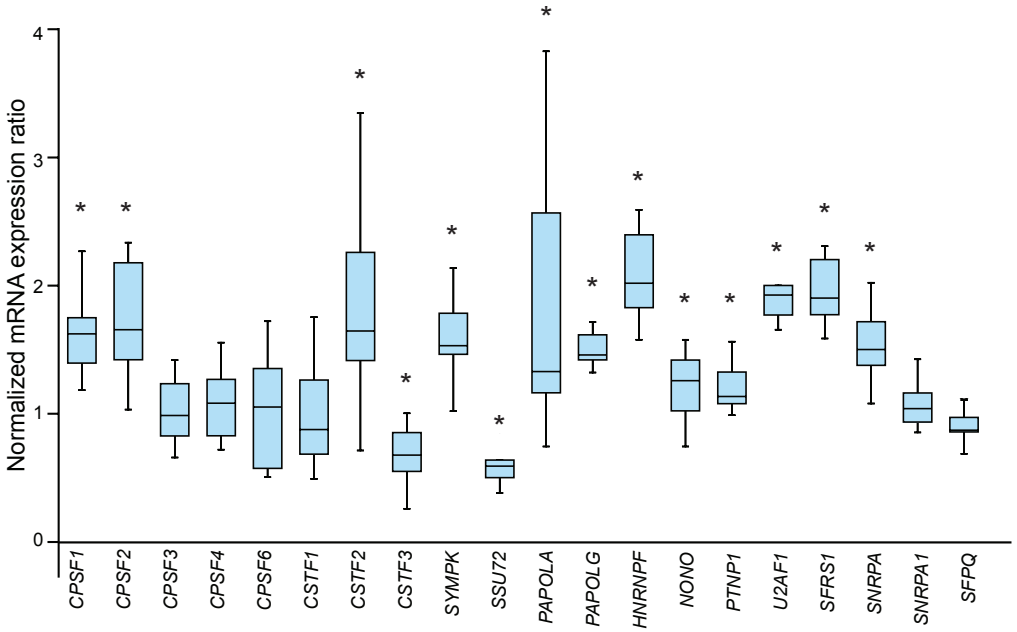

Figure S8. Normalized mRNA expression of proteins involved in polyadenylation. Shown is the mRNA expression of 20 potential *trans*-acting factors (Lou et al., 1999; Vagner et al., 2000; Millevoi et al., 2006; Hall-Pogar et al., 2007; Shi et al., 2009) from nine breast cancer cell lines normalized to the mRNA expression from MCF10A (n=3), a breast epithelial cell line (median, horizontal line; 25th through 75th percentile, box; range, error bars; *, $p<0.013$).

Table S1. miRNA expression in cell lines

Shown is the relative abundance of the 20 highest expressed miRNA families per cell line.

NIH3T3		MEF	
>mmu-let-7	0.249	>mmu-let-7	0.267
>mmu-mir-21	0.098	>mmu-mir-29	0.160
>mmu-mir-26	0.069	>mmu-mir-21	0.123
>mmu-mir-125	0.062	>mmu-mir-24	0.049
>mmu-mir-29	0.055	>mmu-mir-22	0.032
>mmu-mir-17/20/106	0.051	>mmu-mir-125	0.031
>mmu-mir-24	0.040	>mmu-mir-26	0.030
>mmu-mir-130	0.031	>mmu-mir-140	0.025
>mmu-mir-199	0.031	>mmu-mir-15/16	0.023
>mmu-mir-15/16	0.031	>mmu-mir-130	0.022
>mmu-mir-22	0.029	>mmu-mir-199	0.020
>mmu-mir-27	0.021	>mmu-mir-27	0.018
>mmu-mir-322	0.017	>mmu-mir-17/20/106	0.016
>mmu-mir-103/107	0.017	>mmu-mir-31	0.015
>mmu-mir-30	0.015	>mmu-mir-191	0.011
>mmu-mir-19	0.014	>mmu-mir-214	0.011
>mmu-mir-143	0.008	>mmu-mir-103/107	0.011
>mmu-mir-34c	0.008	>mmu-mir-23	0.009
>mmu-mir-23	0.006	>mmu-mir-143	0.008
>mmu-mir-503	0.006	>mmu-mir-221/222	0.007

	HEK293		HeLa
>hsa-mir-17/20/106	0.189	>hsa-let-7	0.339
>hsa-mir-15/16	0.144	>hsa-mir-21	0.184
>hsa-mir-103/107	0.060	>hsa-mir-27	0.052
>hsa-mir-221/222	0.049	>hsa-mir-17/20/106	0.048
>hsa-mir-19	0.047	>hsa-mir-26	0.040
>hsa-mir-26	0.039	>hsa-mir-24	0.034
>hsa-mir-25/32/92/363/367	0.039	>hsa-mir-30	0.027
>hsa-mir-191	0.035	>hsa-mir-92	0.024
>hsa-let-7	0.033	>hsa-mir-19	0.022
>hsa-mir-7	0.031	>hsa-mir-15/16	0.017
>hsa-mir-18	0.024	>hsa-mir-22	0.017
>hsa-mir-148	0.022	>hsa-mir-29	0.017
>hsa-mir-10	0.019	>hsa-mir-125	0.012
>hsa-mir-93	0.016	>hsa-mir-93	0.008
>hsa-mir-130	0.014	>hsa-mir-191	0.008
>hsa-mir-186	0.014	>hsa-mir-103/107	0.007
>hsa-mir-30	0.013	>hsa-mir-143	0.007
>hsa-mir-29	0.013	>hsa-mir-100	0.007
>hsa-mir-378	0.012	>hsa-mir-23	0.006
>hsa-mir-196	0.012	>hsa-mir-186	0.006

	U2OS		143B
>hsa-mir-21	0.445	>hsa-mir-21	0.280
>hsa-let-7	0.088	>hsa-let-7	0.117
>hsa-mir-26	0.041	>hsa-mir-29	0.101
>hsa-mir-130	0.037	>hsa-mir-103/107	0.051
>hsa-mir-15/16	0.035	>hsa-mir-15/16	0.045
>hsa-mir-103/107	0.030	>hsa-mir-27	0.041
>hsa-mir-29	0.028	>hsa-mir-24	0.039
>hsa-mir-27	0.027	>hsa-mir-26	0.035
>hsa-mir-7	0.026	>hsa-mir-17/20/106	0.034
>hsa-mir-24	0.024	>hsa-mir-125	0.027
>hsa-mir-17/20/106	0.024	>hsa-mir-221/222	0.026
>hsa-mir-22	0.023	>hsa-mir-22	0.018
>hsa-mir-221/222	0.015	>hsa-mir-25/32/92/363/367	0.013
>hsa-mir-148	0.011	>hsa-mir-19	0.012
>hsa-mir-93	0.011	>hsa-mir-148	0.012
>hsa-mir-135	0.009	>hsa-mir-130	0.010
>hsa-mir-25/32/92/363/367	0.007	>hsa-mir-146	0.010
>hsa-mir-186	0.007	>hsa-mir-100	0.008
>hsa-mir-19	0.007	>hsa-mir-320	0.007
>hsa-mir-151	0.005	>hsa-mir-101	0.006

	A549		H520
>hsa-mir-21	0.423	>hsa-let-7	0.177
>hsa-let-7	0.087	>hsa-mir-29	0.106
>hsa-mir-24	0.056	>hsa-mir-21	0.102
>hsa-mir-27	0.045	>hsa-mir-103/107	0.058
>hsa-mir-29	0.038	>hsa-mir-200b,c	0.049
>hsa-mir-103/107	0.033	>hsa-mir-24	0.047
>hsa-mir-15/16	0.029	>hsa-mir-26	0.042
>hsa-mir-29	0.026	>hsa-mir-17/20/106	0.040
>hsa-mir-26	0.025	>hsa-mir-221/222	0.032
>hsa-mir-22	0.020	>hsa-mir-27	0.028
>hsa-mir-194	0.018	>hsa-mir-15/16	0.027
>hsa-mir-30	0.018	>hsa-mir-148	0.024
>hsa-mir-221/222	0.015	>hsa-mir-25/32/92/363/367	0.013
>hsa-mir-23	0.011	>hsa-mir-221/222	0.012
>hsa-mir-25/32/92/363/367	0.009	>hsa-mir-22	0.012
>hsa-mir-192	0.009	>hsa-mir-125	0.010
>hsa-mir-191	0.009	>hsa-mir-19	0.009
>hsa-mir-125	0.008	>hsa-mir-130	0.009
>hsa-mir-194	0.007	>hsa-mir-93	0.008
>hsa-mir-130	0.007	>hsa-mir-378	0.008

	SW480		DLD2
>hsa-mir-21	0.201	>hsa-let-7	0.266
>hsa-let-7	0.159	>hsa-mir-29	0.163
>hsa-mir-29	0.089	>hsa-mir-21	0.137
>hsa-mir-499	0.054	>hsa-mir-10	0.059
>hsa-mir-27	0.052	>hsa-mir-200c	0.054
>hsa-mir-200b,c	0.051	>hsa-mir-26	0.053
>hsa-mir-26	0.048	>hsa-mir-27	0.030
>hsa-mir-15/16	0.043	>hsa-mir-17/20/106	0.027
>hsa-mir-17/20/106	0.034	>hsa-mir-103/107	0.023
>hsa-mir-103/107	0.024	>hsa-mir-221/222	0.022
>hsa-mir-10	0.021	>hsa-mir-24	0.021
>hsa-mir-24	0.015	>hsa-mir-15/16	0.012
>hsa-mir-7	0.014	>hsa-mir-148	0.012
>hsa-mir-25/32/92/363/367	0.013	>hsa-mir-19	0.011
>hsa-mir-148	0.013	>hsa-mir-25/32/92/363/367	0.007
>hsa-mir-135	0.011	>hsa-mir-22	0.007
>hsa-mir-93	0.010	>hsa-mir-191	0.006
>hsa-mir-19b	0.009	>hsa-mir-210	0.005
>hsa-mir-221/222	0.008	>hsa-mir-93	0.005
>hsa-mir-96	0.008	>hsa-mir-320	0.005

	MCF7		MDA-MB231
>hsa-mir-21	0.570	>hsa-let-7	0.249
>hsa-let-7	0.121	>hsa-mir-29	0.153
>hsa-mir-26	0.033	>hsa-mir-21	0.126
>hsa-mir-103/107	0.030	>hsa-mir-15/16	0.049
>hsa-mir-200b,c	0.024	>hsa-mir-146	0.037
>hsa-mir-27	0.024	>hsa-mir-27	0.030
>hsa-mir-15/16	0.021	>hsa-mir-17/20/106	0.028
>hsa-mir-17/20/106	0.017	>hsa-mir-30	0.019
>hsa-mir-29	0.016	>hsa-mir-26	0.019
>hsa-mir-141	0.014	>hsa-mir-125	0.019
>hsa-mir-24	0.013	>hsa-mir-103/107	0.016
>hsa-mir-191	0.013	>hsa-mir-221/222	0.015
>hsa-mir-23	0.012	>hsa-mir-25/32/92/363/367	0.013
>hsa-mir-148	0.007	>hsa-mir-23	0.012
>hsa-mir-342	0.006	>hsa-mir-24	0.011
>hsa-mir-96	0.005	>hsa-mir-19	0.011
>hsa-mir-93	0.005	>hsa-mir-22	0.011
>hsa-mir-125	0.005	>hsa-mir-363	0.007
>hsa-mir-135	0.005	>hsa-mir-378	0.007
>hsa-mir-30	0.004	>hsa-mir-96	0.006

Table S2. Estimation of maximum repression due to miRNAs

	miRNA expression	repression per site	Number of miRNA sites per UTR			estimated repression per miRNA family		
	NIH3T3		**IMP-1**	**DICER1**	**Cyclin D2**	**IMP-1**	**DICER1**	**Cyclin D2**
>mmu-let-7	0.249	1.130	5	2	3	1.842	1.277	1.443
>mmu-mir-21	0.098	1.059	2	1	1	1.121	1.059	1.059
>mmu-mir-26	0.069	1.059	0	0	1	1.000	1.000	1.059
>mmu-mir-125	0.062	1.059	0	1	0	1.000	1.059	1.000
>mmu-mir-29	0.055	1.059	0	1	0	1.000	1.059	1.000
>mmu-mir-17/20/106	0.051	1.059	3	1	1	1.188	1.059	1.059
>mmu-mir-24	0.040	1.059	1	0	0	1.059	1.000	1.000
>mmu-mir-130	0.031	1.059	1	0	1	1.059	1.000	1.059
>mmu-mir-199	0.031	1.059	0	1	0	1.000	1.059	1.000
>mmu-mir-15/16	0.031	1.059	0	2	3	1.000	1.121	1.188
>mmu-mir-22	0.029	1.059	2	0	0	1.121	1.000	1.000
>mmu-mir-27	0.021	1.059	0	0	0	1.000	1.000	1.000
>mmu-mir-322	0.017	1.059	0	0	0	1.000	1.000	1.000
>mmu-mir-103/107	0.017	1.059	0	6	0	1.000	1.411	1.000
>mmu-mir-30	0.015	1.059	0	0	1	1.000	1.000	1.059
>mmu-mir-19	0.014	1.059	0	1	2	1.000	1.059	1.121

	miRNA expression	repression per site	Number of miRNA sites per UTR			estimated repression per miRNA family		
	MEF		**IMP-1**	**DICER1**	**Cyclin D2**	**IMP-1**	**DICER1**	**Cyclin D2**
mmu-let-7	0.267	1.130	5	2	3	1.842	1.277	1.443
mmu-mir-29	0.160	1.081	0	1	0	1.000	1.081	1.000
mmu-mir-21	0.123	1.081	2	1	1	1.169	1.081	1.081
mmu-mir-24	0.049	1.059	1	0	0	1.059	1.000	1.000
mmu-mir-22	0.032	1.059	2	0	0	1.121	1.000	1.000
mmu-mir-125	0.031	1.059	0	1	0	1.000	1.059	1.000
mmu-mir-26	0.030	1.059	0	0	1	1.000	1.000	1.059
mmu-mir-140	0.025	1.059	2	1	0	1.121	1.059	1.000
mmu-mir-15/16	0.023	1.059	0	2	3	1.000	1.121	1.188
mmu-mir-130	0.022	1.059	1	0	1	1.059	1.000	1.059
mmu-mir-199	0.020	1.059	0	1	0	1.000	1.059	1.000
mmu-mir-27	0.018	1.059	0	0	0	1.000	1.000	1.000
mmu-mir-17/20/106	0.016	1.059	3	1	1	1.188	1.059	1.059
mmu-mir-31	0.015	1.059	0	0	0	1.000	1.000	1.000
mmu-mir-191	0.011	1.059	0	0	0	1.000	1.000	1.000
mmu-mir-214	0.011	1.059	0	1	2	1.000	1.059	1.121
mmu-mir-103/107	0.011	1.059	0	6	0	1.000	1.411	1.000

	miRNA expression	repression per site	Number of miRNA sites per UTR			estimated repression per miRNA family		
	HEK293		**IMP-1**	**DICER1**	**Cyclin D2**	**IMP-1**	**DICER1**	**Cyclin D2**
hsa-mir-17/20/106	0.189	1.081	3	1	1	1.263	1.081	1.081
hsa-mir-15/16	0.144	1.081	0	2	3	1.000	1.169	1.263
hsa-mir-103/107	0.060	1.059	0	6	0	1.000	1.411	1.000
hsa-mir-1221/222	0.049	1.059	0	1	1	1.000	1.059	1.059
hsa-mir-19	0.047	1.059	0	1	2	1.000	1.059	1.121
hsa-mir-26	0.039	1.059	0	0	1	1.000	1.000	1.059
hsa-mir-25/32/92/363/367	0.039	1.059	0	0	0	1.000	1.000	1.000
hsa-mir-191	0.035	1.059	2	0	1	1.121	1.000	1.059
hsa-let-7	0.033	1.059	5	2	3	1.332	1.121	1.188
hsa-mir-7	0.031	1.059	0	0	2	1.000	1.000	1.121
hsa-mir-18	0.024	1.059	0	2	2	1.000	1.121	1.121
hsa-mir-148	0.022	1.059	0	0	0	1.000	1.000	1.000
hsa-mir-10	0.019	1.059	2	0	0	1.121	1.000	1.000
hsa-mir-93	0.016	1.059	1	0	2	1.059	1.000	1.121
hsa-mir-130	0.014	1.059	1	0	1	1.059	1.000	1.059
hsa-mir-186	0.014	1.059	0	0	1	1.000	1.000	1.059
hsa-mir-30	0.013	1.059	0	0	1	1.000	1.000	1.059
hsa-mir-29	0.013	1.059	0	1	0	1.000	1.059	1.000
hsa-mir-378	0.012	1.059	1	1	3	1.059	1.059	1.188
hsa-mir-196	0.012	1.059	0	1	0	1.000	1.059	1.000

	miRNA expression	repression per site	Number of miRNA sites per UTR			estimated repression per miRNA family		
	HeLa		**IMP-1**	**DICER1**	**Cyclin D2**	**IMP-1**	**DICER1**	**Cyclin D2**
>hsa-let-7	0.339	1.130	5	2	3	1.842	1.277	1.443
>hsa-mir-21	0.184	1.081	2	1	1	1.169	1.081	1.081
>hsa-mir-27	0.052	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-17/20/106	0.048	1.059	3	1	1	1.188	1.059	1.059
>hsa-mir-26	0.040	1.059	0	0	1	1.000	1.000	1.059
>hsa-mir-24	0.034	1.059	1	0	0	1.059	1.000	1.000
>hsa-mir-30	0.027	1.059	0	0	1	1.000	1.000	1.059
>hsa-mir-25/32/92/363/367	0.024	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-19	0.022	1.059	0	1	2	1.000	1.059	1.121
>hsa-mir-15/16	0.017	1.059	0	2	3	1.000	1.121	1.188
>hsa-mir-22	0.017	1.059	2	0	0	1.121	1.000	1.000
>hsa-mir-29	0.017	1.059	0	1	0	1.000	1.059	1.000
>hsa-mir-125	0.012	1.059	0	1	0	1.000	1.059	1.000

	miRNA expression	repression per site	Number of miRNA sites per UTR			estimated repression per miRNA family		
	U2OS		**IMP-1**	**DICER1**	**Cyclin D2**	**IMP-1**	**DICER1**	**Cyclin D2**
>hsa-mir-21	0.445	1.130	2	1	1	1.277	1.130	1.130
>hsa-let-7	0.088	1.059	5	2	3	1.332	1.121	1.188
>hsa-mir-26	0.041	1.059	0	0	1	1.000	1.000	1.059
>hsa-mir-130	0.037	1.059	1	0	1	1.059	1.000	1.059
>hsa-mir-15/16	0.035	1.059	0	2	3	1.000	1.121	1.188
>hsa-mir-103/107	0.030	1.059	0	6	0	1.000	1.411	1.000
>hsa-mir-29	0.028	1.059	0	1	0	1.000	1.059	1.000
>hsa-mir-27	0.027	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-7	0.026	1.059	0	0	2	1.000	1.000	1.121
>hsa-mir-24	0.024	1.059	1	0	0	1.059	1.000	1.000
>hsa-mir-17/20/106	0.024	1.059	3	1	1	1.188	1.059	1.059
>hsa-mir-22	0.023	1.059	2	0	0	1.121	1.000	1.000
>hsa-mir-221/222	0.015	1.059	0	1	1	1.000	1.059	1.059
>hsa-mir-148	0.011	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-93	0.011	1.059	1	0	2	1.059	1.000	1.121

	miRNA expression	repression per site	Number of miRNA sites per UTR			estimated repression per miRNA family		
	143B		**IMP-1**	**DICER1**	**Cyclin D2**	**IMP-1**	**DICER1**	**Cyclin D2**
>hsa-mir-21	0.280	1.130	2	1	1	1.277	1.130	1.130
>hsa-let-7	0.117	1.081	5	2	3	1.476	1.169	1.263
>hsa-mir-29	0.101	1.081	0	1	0	1.000	1.081	1.000
>hsa-mir-103/107	0.051	1.059	0	6	0	1.000	1.411	1.000
>hsa-mir-15/16	0.045	1.059	0	2	3	1.000	1.121	1.188
>hsa-mir-27	0.041	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-24	0.039	1.059	1	0	0	1.059	1.000	1.000
>hsa-mir-26	0.035	1.059	0	0	1	1.000	1.000	1.059
>hsa-mir-17/20/106	0.034	1.059	3	1	1	1.188	1.059	1.059
>hsa-mir-125	0.027	1.059	0	1	0	1.000	1.059	1.000
>hsa-mir-221/222	0.026	1.059	0	1	1	1.000	1.059	1.059
>hsa-mir-22	0.018	1.059	2	0	0	1.121	1.000	1.000
>hsa-mir-25/32/92/363/367	0.013	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-19	0.012	1.059	0	1	2	1.000	1.059	1.121
>hsa-mir-148	0.012	1.059	0	0	0	1.000	1.000	1.000

	miRNA expression	repression per site	Number of miRNA sites per UTR			estimated repression per miRNA family		
	A549		**IMP-1**	**DICER1**	**Cyclin D2**	**IMP-1**	**DICER1**	**Cyclin D2**
>hsa-mir-21	0.423	1.130	2	1	1	1.277	1.130	1.130
>hsa-let-7	0.087	1.081	5	2	3	1.476	1.169	1.263
>hsa-mir-24	0.056	1.059	1	0	0	1.059	1.000	1.000
>hsa-mir-27	0.045	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-29	0.038	1.059	0	1	0	1.000	1.059	1.000
>hsa-mir-103/107	0.033	1.059	1	6	0	1.000	1.411	1.000
>hsa-mir-15/16	0.029	1.059	1	2	3	1.000	1.121	1.188
>hsa-mir-26	0.025	1.059	0	0	1	1.000	1.000	1.059
>hsa-mir-22	0.020	1.059	2	0	0	1.121	1.000	1.000
>hsa-mir-194	0.018	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-30	0.018	1.059	0	0	1	1.000	1.000	1.059
>hsa-mir-221/222	0.015	1.059	0	1	1	1.000	1.059	1.059
>hsa-mir-23	0.011	1.059	1	0	0	1.059	1.000	1.000

	miRNA expression	repression per site	Number of miRNA sites per UTR			estimated repression per miRNA family		
	H520		**IMP-1**	**DICER1**	**Cyclin D2**	**IMP-1**	**DICER1**	**Cyclin D2**
>hsa-let-7	0.177	1.081	5	2	3	1.476	1.169	1.263
>hsa-mir-29	0.106	1.081	0	1	0	1.000	1.081	1.000
>hsa-mir-21	0.102	1.081	2	1	1	1.169	1.081	1.081
>hsa-mir-103/107	0.058	1.059	0	6	0	1.000	1.411	1.000
>hsa-mir-200b,c	0.049	1.059	1	0	0	1.059	1.000	1.000
>hsa-mir-24	0.047	1.059	1	0	0	1.059	1.000	1.000
>hsa-mir-26	0.042	1.059	0	0	1	1.000	1.000	1.059
>hsa-mir-1720/106	0.040	1.059	3	1	1	1.188	1.059	1.059
>hsa-mir-221/222	0.032	1.059	0	1	1	1.000	1.059	1.059
>hsa-mir-27	0.028	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-15/16	0.027	1.059	0	2	3	1.000	1.121	1.188
>hsa-mir-148	0.024	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-25/32/92/363/367	0.013	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-221/222	0.012	1.059	0	1	1	1.000	1.059	1.059
>hsa-mir-22	0.012	1.059	2	0	0	1.121	1.000	1.000
>hsa-mir-125	0.010	1.059	0	1	0	1.000	1.059	1.000

	miRNA expression	repression per site	Number of miRNA sites per UTR			estimated repression per miRNA family		
	SW480		**IMP-1**	**DICER1**	**Cyclin D2**	**IMP-1**	**DICER1**	**Cyclin D2**
>hsa-mir-21	0.201	1.130	2	1	1	1.277	1.130	1.130
>hsa-let-7	0.159	1.081	5	2	3	1.476	1.169	1.263
>hsa-mir-29	0.089	1.059	0	1	0	1.000	1.059	1.000
>hsa-mir-499	0.054	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-27	0.052	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-200c	0.051	1.059	1	0	0	1.059	1.000	1.000
>hsa-mir-26	0.048	1.059	0	0	1	1.000	1.000	1.059
>hsa-mir-/1516	0.043	1.059	0	2	3	1.000	1.121	1.188
>hsa-mir-17/20/106	0.034	1.059	3	1	1	1.188	1.059	1.059
>hsa-mir-103/107	0.024	1.059	0	6	0	1.000	1.411	1.000
>hsa-mir-10	0.021	1.059	2	0	0	1.121	1.000	1.000
>hsa-mir-24	0.015	1.059	1	0	0	1.059	1.000	1.000
>hsa-mir-7	0.014	1.059	0	0	2	1.000	1.000	1.121
>hsa-mir-25/32/92/363/367	0.013	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-148	0.013	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-135	0.011	1.059	2	0	0	1.121	1.000	1.000

	miRNA expression	repression per site	Number of miRNA sites per UTR			estimated repression per miRNA family		
	DLD2		**IMP-1**	**DICER1**	**Cyclin D2**	**IMP-1**	**DICER1**	**Cyclin D2**
>hsa-let-7	0.266	1.130	5	2	3	1.842	1.277	1.443
>hsa-mir-29	0.163	1.081	0	1	0	1.000	1.081	1.000
>hsa-mir-21	0.137	1.081	2	1	1	1.169	1.081	1.081
>hsa-mir-10	0.059	1.059	2	0	0	1.121	1.000	1.000
>hsa-mir-200c	0.054	1.059	1	0	0	1.059	1.000	1.000
>hsa-mir-26	0.053	1.059	0	0	1	1.000	1.000	1.059
>hsa-mir-27	0.030	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-17/20/106	0.027	1.059	3	1	1	1.188	1.059	1.059
>hsa-mir-103/107	0.023	1.059	0	6	0	1.000	1.411	1.000
>hsa-mir-221222	0.022	1.059	0	1	1	1.000	1.059	1.059
>hsa-mir-24	0.021	1.059	1	0	0	1.059	1.000	1.000
>hsa-mir-15/16	0.012	1.059	0	2	3	1.000	1.121	1.188
>hsa-mir-148	0.012	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-19	0.011	1.059	0	1	2	1.000	1.059	1.121

	miRNA expression	repression per site	Number of miRNA sites per UTR			estimated repression per miRNA family		
	MCF7		**IMP-1**	**DICER1**	**Cyclin D2**	**IMP-1**	**DICER1**	**Cyclin D2**
>hsa-mir-21	0.570	1.130	2	1	1	1.277	1.130	1.130
>hsa-let-7	0.121	1.081	5	2	3	1.476	1.169	1.263
>hsa-mir-26	0.033	1.059	0	0	1	1.000	1.000	1.059
>hsa-mir-103/107	0.030	1.059	0	6	0	1.000	1.411	1.000
>hsa-mir-200b,c	0.024	1.059	1	0	0	1.059	1.000	1.000
>hsa-mir-27	0.024	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-15/16	0.021	1.059	0	2	3	1.000	1.121	1.181
>hsa-mir-29	0.016	1.059	0	1	0	1.000	1.059	1.000
>hsa-mir-24	0.013	1.059	1	0	0	1.059	1.000	1.000
>hsa-mir-191	0.013	1.059	2	0	1	1.121	1.000	1.059
>hsa-mir-23	0.012	1.059	1	0	0	1.057	1.000	1.000
>hsa-mir-17/20/106	0.010	1.059	3	1	1	1.188	1.059	1.059

	miRNA expression	repression per site	Number of miRNA sites per UTR			estimated repression per miRNA family		
	MDA-MB231		**IMP-1**	**DICER1**	**Cyclin D2**	**IMP-1**	**DICER1**	**Cyclin D2**
>hsa-let-7	0.249	1.130	5	2	3	1.842	1.277	1.443
>hsa-mir-29	0.153	1.081	0	1	0	1.000	1.081	1.000
>hsa-mir-21	0.126	1.081	2	1	1	1.169	1.081	1.081
>hsa-mir-15/16	0.049	1.059	0	2	3	1.000	1.121	1.188
>hsa-mir-146	0.037	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-27	0.030	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-17/20/106	0.028	1.059	3	1	1	1.188	1.059	1.059
>hsa-mir-30	0.019	1.059	0	0	1	1.000	1.000	1.059
>hsa-mir-26	0.019	1.059	0	0	1	1.000	1.000	1.059
>hsa-mir-125	0.019	1.059	0	1	0	1.000	1.059	1.000
>hsa-mir-103/107	0.016	1.059	0	6	0	1.000	1.411	1.000
>hsa-mir-222	0.015	1.059	0	1	1	1.000	1.059	1.059
>hsa-mir-92	0.013	1.059	0	0	0	1.000	1.000	1.000
>hsa-mir-23	0.012	1.059	1	2	0	1.059	1.121	1.000
>hsa-mir-24	0.011	1.059	1	0	0	1.059	1.000	1.000
>hsa-mir-19	0.011	1.059	0	1	2	1.000	1.059	1.121
>hsa-mir-22	0.011	1.059	2	0	0	1.121	1.000	1.000

Supplemental Experimental procedures

Cell lines

The following cell lines were purchased from ATCC and cultured as indicated: HEK293, HEK293T, HeLa, NIH3T3, MEF, F9, human sarcoma cell lines (U2OS, Saos-2, HOS, 143B), human breast cancer cell lines (A549, NCI-H23, NCI-H292, NCI-H460, NCI-H520, NCI-H596), the human immortalized bronchial epithelial cell line Beas-2B, human immortalized fibroblasts (HF) and human colon cancer cell lines (HCT116, SW480, HT-29, Colo-205, Lovo and DLD2). NHBE (normal human bronchial epithelial cells) were purchased from Lonza and cultured as indicated. The colon cancer cell lines SKCO1 and Ls147T were a gift from Thijn Brummelkamp (Whitehead Institute, Cambridge, USA). Immortalized breast epithelial cell lines (HME, HMLE, HBL-100, MCF10A) and breast cancer cell lines MCF7, MDA-MB231, MDA-MB435, MDA-MB453, MDA-MB361, MDA-MB468, MDA-MB157, BT474, BT549, T47D, CAMA) were a gift from Robert Weinberg's lab (Whitehead Institute, Cambridge, USA). Phoenix cells were a gift from Michael Hemann (MIT, Cambridge, USA).

Prediction of miRNA sites

Targetscan (version 4.2) was used for prediction of miRNA targets (Lewis et al., 2005).

Northern blots to detect miRNAs

Total RNA was isolated from the above mentioned cell lines using Tri-reagent (Ambion) and 20 µg was loaded per lane. RNA blotting was performed as described (Mayr et al., 2007) (http://web.wi.mit.edu/bartel/pub/protocols/), with the following DNA oligo probes:

let-7a,b,c,d 5'-CAACCTACTACCTCA

let-7e 5'-ACTATACAACCTCCTACCTCA

let-7f 5'-ACTATACAATCTACTACCTCA

let-7g,i 5'-ACAAACTACTACCTCA

miR-15a/b 5'-AAACCATGATGTGCTGCTA
miR-16 5'-CGCCAATATTTACGTGCTGCTA
miR-103/107 5'-ATAGCCCTGTACAATGCTGCT
U6 snRNA 5'-TTGCGTGTCATCCTTGCGCAGG

Northern blots to detect mRNAs

Total RNA was isolated from the above mentioned cell lines using Tri-reagent (Ambion). Polyadenylated RNA was purified with Oligotex (Qiagen) and 1.5-2 µg was used per sample. The protocol for the Northern blot was adapted from Cold Spring Harbor Protocols (Sambrook & Russell, 2006). 2 µl of poly(A)$^+$ RNA was incubated with 10 µl glyoxal reaction mixture [6 ml DMSO, 2 ml of deionized glyoxal, 1.2 ml of 10X BPTE electrophoresis buffer (100 mM PIPES, 300 mM Bis-Tris, 10 mM EDTA), 0.6 ml of 80% glycerol in H_2O, 0.2 ml ethidium bromide (10 mg/ml in H_2O)] at 55 °C for 1 hour. Glyoxal-treated RNA was separated on an agarose gel and transferred onto Nytran SuPerCharge Turboblotter Membrane (Whatman) overnight. The glyoxal reaction was reversed by incubation of the membrane in 20 mM Tris-HCL (pH = 8), then crosslinked and baked. Blots were prehybridized in UltraHyb solution (Ambion), hybridized and washed according to the manufacturer's instructions. PCR probes were gel-purified and labeled with ^{32}P according to the manufacturer's instructions using the Megaprime labeling kit (Amersham). Blots were scanned on a Phosphoimager, and bands were quantified using MultiGauge V2.2.

mRNA stability

Cell lines were treated with Actinomycin D (10 µg/ml), total RNA was isolated at 0h, 2h, 4h, 6h and 8h and Northern blots were performed as described above. Half-life was calculated: $t_{1/2} = 0.693/[(\ln c1-\ln c2)/t]$, with t = time interval between c1 and c2, c1 = amount of mRNA at t = 0 and c2 = mRNA amount at t = 2h, 4h or 6h.

3' RACE

This protocol was adapted from Cold Spring Harbor Protocols (Sambrook & Russell, 2006). 1 µg total RNA was used to generate cDNA with Superscript II reverse transcriptase (Invitrogen) according to the manufacturer's instructions using TAP as a primer. The first PCR was done with a gene-specific forward primer and AP as reverse primer. Nested PCR was done with a nested gene-specific forward primer and MAP as reverse primer. The PCR product from the nested PCR was separated on an agarose gel, cloned and sequenced.

TAP: 5'-GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT;

AP: 5'-GACTCGAGTCGACATCG;

MAP: 5'-CGACATCGATTTTTTTTTT.

If only full-length mRNAs were found by 3' RACE, Northern blots were reprobed with different probes in the ORF to find alternatively spliced isoforms.

Constructs

The long and the short 3'UTRs of the human *IGF2BP1/IMP-1, DICER1* and *Cyclin D2* were PCR amplified with Pfu UltraII (Stratagene) and subcloned into the *Renilla* luciferase vector pIS1 (Mayr et al., 2007). The following primers were used:

IMP-1 fw: 5'-CAAACCGGTCCAGCCCCTCCCTGTCCCTTCG

IMP-1 long rev: 5'-CGTGCGGCCGCTTGTCTTTCAATATAAGTTTCCT

IMP-1 short rev: 5'-CGTGCGGCCGCAAATGAACCATCCTTTAAGGC

DICER1 fw: 5'-GCCTCTAGAAACCGCTTTTTAAAATTC

DICER1 long rev: 5'-GATGCGGCCGCGAACAGACGATAAC

DICER1 short rev: 5'- GATGCGGCCGCTATACATATGTTAAATGTTTTATTCTG

Cyclin D2 fw: 5'-CAATCTAGAGGATGCCAGTTGGGCCGAAAG

Cyclin D2 long rev: 5'-CGTGCGGCCGCAGGTCAAGGTGAGTTTATTGTCC

Cyclin D2 short rev: 5'-CGTGCGGCCGCTTCAAATAGGCACCAAATGC

To generate only the long mRNAs poly(A) signals were mutated to a C at positions 2 and 4. MiRNA complementary sites were mutated at nucleotides

complementary to positions 3 and 5 of the seed of the miRNA using Quikchange Multi-site kit (Stratagene).

For functional analyses, the ORF of human IMP-1 together with either the short or the long 3'UTR (with mutation of all four poly(A) signals) with either wild-type *let-7* sites or mutant *let-7* sites was cloned into pPIG (which was a gift of Michael Hemann, MIT, USA), which was derived from pMSCVpuro (Clontech) and modified by cloning the IRES of encephalomyocarditis virus (ECMV) and GFP downstream of puromycin into the vector. The ORF of the human *Cyclin D2* gene with 120 bp of the 5'UTR (fw primer: 5'-TATTTAGCCAAAGGAAGGAGGTC) together with either the short or the long 3' UTR (with mutation of the first AAUAAA) with either wild-type miR-15/16 and *let-7* sites or mutant miR-15/16 and *let-7* sites was cloned into pMSCVpuro (Clontech).

Luciferase assays

All cell lines were transfected using Lipofectamine 2000 (Invitrogen) in 24-well plates with 100 ng firefly luciferase control reporter plasmid pISO (Mayr et al., 2007) and 400 ng *Renilla* luciferase reporter plasmid per well for the full-length mRNA. For the shorter mRNA same molar amounts were transfected. Firefly and *Renilla* luciferase activities were measured 24 hours after transfection with the Dual-luciferase assay (Promega). *Renilla* activity was normalized to firefly activity to control for transfection efficiency. To account for differences in plasmid preparations, values were then normalized to those of the reporter in F9 cells for the *let-7* miRNA, in DLD2 cells for miR-15/16 and in 143B for miR-103/107.

Cell culture

To produce retroviral particles for ectopic expression of Cyclin D2 HEK293T cells were transfected with Lipofectamine 2000 (Invitrogen) together with plasmids for VSV-G and MCV to transduce human cell lines. Supernatant containing viral particles was harvested 48 hours later. The breast cell lines were transduced

(with comparable MOIs) and 24 hours later puromycin (1 µg/ml) was added. Marker-selected populations were obtained after 7-10 days.

To produce retroviral particles for ectopic expression of IMP-1 Phoenix cells were transfected with Lipofectamine 2000 (Invitrogen) and supernatant containing viral particles was harvested 48 hours later. NIH3T3 cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% FCS (Invitrogen). NIH3T3 were transduced (with comparable MOIs) and 24 hours later puromycin (2 µg/ml) was added. Marker-selected populations were obtained after 7-10 days. To transduce human cell lines HEK293T cells were used as packaging cells and transfected together with plasmids for VSV-G and MCV. The transcript for expression of the short isoform of IMP-1 is 6,124 bp (5' PCMV LTR, ψ+ packaging signal, IMP-1, PGK, Puromycin, IRES, GFP) and the long isoform of IMP-1 is 12,412 bp. For transduction of mouse cells (NIH3T3) we used Phoenix cells as the packaging cell line and obtained NIH3T3 cells expressing human IMP-1. For transduction of human epithelial cell lines we used HEK293T cells for packaging. For the short isoform we obtained comparable viral titers as for transduction of NIH3T3 cells, but for the longer isoforms the titers obtained were low. Visualization of the mRNA transcripts on a Northern blot showed IMP-1 transcripts including the ORF for the NIH3T3 cells as well as for the shorter IMP-1 isoforms in the human cells, but splice variants without the IMP-1 ORF for the long isoforms. These spliced transcripts (including the antibiotic selection marker) were about 5 kb. We suspect that the transcript including the full-length IMP-1 is rather large for packaging in HEK293T cells and we were unable to express it in human cells.

Western blot

Cells were lysed in Laemmli buffer (Biorad). Samples were run on 4-12% Tris-HCL gels, transferred to PVDF membrane (Biorad), blocked with Odyssey blocking buffer (LI-COR) and probed with anti-IMP-1 antibody (1:200; Santa Cruz), anti-Cyclin D2 antibody (1:200; Santa Cruz) and anti-actin antibody

(1:10,000; Sigma). Blots were scanned and bands were quantified using the Odyssey Imager (LI-COR).

RT–PCR

RNA was extracted with Tri-reagent (Ambion), treated with DNase using the DNA-free kit (Ambion). DNA-free RNA was reverse-transcribed with random hexamers and Superscript II (Invitrogen) according to manufacturer's instructions. Quantitative RT–PCR was performed on an ABI 7900HT Real-Time PCR system with ABI SYBR Green reagents. The following primer pairs were used to amplify the specified mRNAs:
IMP-1 fw: 5'-GAAGAAGGTAGAGCAAGATACCG,
IMP-1 rev: 5'-CCCGAACTTTCTTCATTATTTCC,
Cyclin D2 fw: 5'- GAGCTGCTGGCTAAGATCACC,
Cyclin D2 rev: 5'-ATATCCCGCACGTCTGTAGG,
Puromycin fw: 5'-CACCAGGGCAAGGGTCTG,
Puromycin rev: 5'-GCTCGTAGAAGGGGAGGTTG,
human GAPDH fw: 5'- GGTCTCCTCTGACTTCAACAGC,
human GAPDH rev: 5'- GCTGTAGCCAAATTCGTTGTCATACC,
mouse GAPDH fw: 5'- CTCACTCAAGATTGTCAGCAATG,
mouse GAPDH rev: 5'-CACATTGGGGGTAGGAACAC.
Threshold cycle (Ct) and baseline were detected by ABI 7900HT SDS2.3 software.

Cell cycle analysis by FACS

Cells were plated at comparable densities, harvested after 48h, fixed with ethanol and stained with propidium iodide (50μg/ml) and RNase A (40 U/ml) and DNA content was measured on a FACS Calibur HTS (Becton Dickinson). The percentage of diploid cells in G1, S and G2 were analyzed by ModFitLT V3.1.

Soft-agar Assay

Soft-agar assays were performed as described (Mayr et al., 2007). 2×10^5 transduced NIH3T3 cells were suspended in 0.5% Noble Agar (Sigma) in Ham's F12 medium (Cellgro), supplemented with 12% FCS and puromycin (1 µg/ml), and plated over a first layer of 0.5% Noble Agar in Ham's F12 medium. The cells were grown at 37° C and 5% $C0_2$, and colonies were counted at day 28. Colonies of eight cells or more were counted. Six independent transductions were done for each of the constructs. Each transduction was plated in triplicate. For each plate, 30 fields of 30 cells were counted. HMLE and MCF10A cells were suspended in 0.5% Noble Agar (Sigma) in MEGM (Lonza) and Beas-2B in BEGM (Lonza) and supplemented with puromycin (1 µg/ml). Colonies (five cells or more) from at least three independent experiments were counted.

mRNA expression of *trans*-acting factors

Published Affymetrix array data (GSE12790) were analyzed. Three replicates of MCF10A were used to normalize the expression of 20 *trans*-acting factors from nine breast cancer cell lines (MCF7, MDA-MB231, MDA-MB361, MDA-MB468, MDA-MB453, BT474, BT549, T47D and CAMA).

Statistics

The Kruskal-Wallis test was used to analyze the difference between several independent subgroups. Mann Whitney test was applied to analyze the difference between two independent supgroups. The Wilcoxon test was used to make pairwise comparisons using SPSS 14.0.

Supplemental References

Hall-Pogar, T., Liang, S., Hague, L. K., and Lutz, C. S. (2007). Specific trans-acting proteins interact with auxiliary RNA polyadenylation elements in the COX-2 3'-UTR. Rna *13*, 1103-1115.

Lou, H., Helfman, D. M., Gagel, R. F., and Berget, S. M. (1999). Polypyrimidine tract-binding protein positively regulates inclusion of an alternative 3'-terminal exon. Mol Cell Biol *19*, 78-85.

Millevoi, S., Loulergue, C., Dettwiler, S., Karaa, S. Z., Keller, W., Antoniou, M., and Vagner, S. (2006). An interaction between U2AF 65 and CF I(m) links the splicing and 3' end processing machineries. Embo J *25*, 4854-4864.
Shi, Y., Di Giammartino, D. C., Taylor, D., Sarkeshik, A., Rice, W. J., Yates, J. R., 3rd, Frank, J., and Manley, J. L. (2009). Molecular architecture of the human pre-mRNA 3' processing complex. Mol Cell *33*, 365-376.
Vagner, S., Vagner, C., and Mattaj, I. W. (2000). The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing. Genes Dev *14*, 403-413.