Let-7 represses Nr6a1 and a mid-gestation developmental program in adult fibroblasts

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Let-7 represses *Nr6a1* and a mid-gestation developmental program in adult fibroblasts

Allan M. Gurtan, Arvind Ravi, Peter B. Rahl, et al.

*Genes Dev.* 2013 27: 941-954

Access the most recent version at doi:10.1101/gad.215376.113

Supplemental Material

http://genesdev.cshlp.org/content/suppl/2013/04/29/27.8.941.DC1.html

References

This article cites 68 articles, 23 of which can be accessed free at:
http://genesdev.cshlp.org/content/27/8/941.full.html#ref-list-1

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

To subscribe to *Genes & Development* go to:
http://genesdev.cshlp.org/subscriptions

Copyright © 2013 by Cold Spring Harbor Laboratory Press
Let-7 represses \textit{Nr6a1} and a mid-gestation developmental program in adult fibroblasts

Allan M. Gurtan,1 Arvind Ravi,2 Peter B. Rahl,3 Andrew D. Bosson,1,4 Courtney K. JnBaptiste,1,4 Arjun Bhutkar,1 Charles A. Whittaker,1 Richard A. Young,3 and Phillip A. Sharp1,4,5

1David H. Koch Institute for Integrative Cancer Research, Cambridge, Massachusetts 02139, USA; 2Harvard-Massachusetts Institute of Technology Health Sciences and Technology Program, Cambridge, Massachusetts 02139, USA; 3Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA; 4Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

MicroRNAs (miRNAs) are critical to proliferation, differentiation, and development. Here, we characterize gene expression in murine Dicer-null adult mesenchymal stem cell lines, a fibroblast cell type. Loss of Dicer leads to derepression of let-7 targets at levels that exceed 10-fold to 100-fold with increases in transcription. Direct and indirect targets of this miRNA belong to a mid-gestation embryonic program that encompasses known oncofetal genes as well as oncogenes not previously associated with an embryonic state. Surprisingly, this mid-gestation program represents a distinct period that occurs between the pluripotent state of the inner cell mass at embryonic day 3.5 (E3.5) and the induction of let-7 upon differentiation at E10.5. Within this mid-gestation program, we characterize the let-7 target \textit{Nr6a1}, an embryonic transcriptional repressor that regulates gene expression in adult fibroblasts following miRNA loss. In total, let-7 is required for the continual suppression of embryonic gene expression in adult cells, a mechanism that may underlie its tumor-suppressive function.

[Keywords: Dicer; miRNAs; let-7; embryo; cancer; oncofetal]

Supplemental material is available for this article.

Received February 2, 2013; revised version accepted March 27, 2013.
existing somatic in vitro studies have yielded significant insight into the activity of miRNAs in differentiated cell types (Dugas et al. 2010). Many somatic cells predomi-
nantly express let-7, a highly conserved miRNA known for its role in opposing self-renewal programs regulated by miR-290 (Melton et al. 2010). In multiple species, ranging from nematodes to mammals, mature let-7 becomes expressed in the developing embryo (Nimmo and Slack 2009, Ambros 2011). In nematodes, loss of the let-7 family results in the reiteration of larval stages. Let-7 is similarly thought to regulate developmental timing in mammals (Schulman et al. 2005). Exogenous expression of mature let-7 in ESCs in vitro represses pluripotency programs and antagonizes the activity of the miR-290 family (Melton et al. 2010). In vivo, however, the dispersion of let-7 genes across multiple genomic loci has confounded its genetic analysis in mammalian development.

In the adult, mature let-7 persists at high levels and functions as a tumor suppressor (Boyerialas et al. 2010; Trang et al. 2010). Many targets of let-7 are oncogenes, and reduction of let-7 levels is frequently observed in cancer. A small subset of let-7 targets falls into a conceptual class of genes termed “oncofetal” because of their expression profile (Boyerialas et al. 2008; Gurtan and Sharp 2013): highly expressed in the embryo, inactive in most adult tissues, and re-expressed in tumors. The only oncofetal targets of let-7 characterized thus far have been the RNA-binding families Igf2bp1–3 (Lee and Dutta 2007, Mayr et al. 2007, Boyerialas et al. 2008; Mayr Bartel 2009). These oncofetal targets regulate proliferation and growth. Transgenic mice that express elevated levels of Lin-28 are larger than control littermates [Zhu et al. 2010, 2011], while knockouts of Hmga2 [Zhou et al. 1995] or Igf2bp1 [Hansen et al. 2004] exhibit dwarf phenotypes. Transgenic overexpression of Hmga2 [Fedele et al. 2006; Zaidi et al. 2006] or Igf2bp1 [Tessier et al. 2004] results in tumorigenesis. Given the large number of predicted but uncharacterized targets of let-7, additional genes may populate networks typified by known oncofetal genes.

Here, we exploit a murine Dicer-null somatic cell line to identify gene expression programs regulated by let-7 specifically and somatic miRNAs generally in the context of the adult. Deletion of Dicer in adult MSCs, a fibroblast cell type that abundantly expresses let-7, results in specific, transcriptionally reinforced changes in gene expression, including global derepression of miRNA targets. Dicer-null MSCs largely retain their cellular identity but, within this landscape of mesenchymal gene expression, activate a discrete let-7-targeted mid-gestation developmental program that includes known oncofetal genes as well as oncogenes not previously associated with an embryonic state. Within this mid-gestation program, we characterize the let-7 target Ntr6a1, an embryonic transcriptional repressor that contributes to gene expression changes in adult fibroblasts following miRNA loss. In total, we report that somatic miRNAs such as let-7 are required for the continual suppression of embryonic gene expression in adult cells, a mechanism that may underlie their tumor-suppressive functions.

Results

Dicer knockout adult fibroblasts exhibit specific changes in gene expression

To understand the role of miRNAs in regulating gene expression in somatic cells, we performed expression profiling in a recently described model of Dicer loss. Specifically, we used immortalized clonal lines of Dicer1f/Dicer1f [Dicer wild-type] and Dicer1f/Dicer1f [Dicer knockout] MSCs, a fibroblast cell type present in adult bone marrow [Ravi et al. 2012]. To establish a baseline of miRNA expression in these fibroblast cell lines, we performed small RNA sequencing [small RNA-seq] [Supplemental Table S1]. In Dicer wild-type MSCs, the five most abundant miRNAs were miR-22 (17%), let-7c (16%), let-7b (9%), miR-16 (7%), and miR-145 (6%). Collapsing miRNAs by seeds, the let-7 family comprised 37% of all miRNA reads, consistent with published observations of let-7 predominance in somatic tissues [Marson et al. 2008]. This expression profile is comparable with previous studies in these cells [Gurtan et al. 2012] and closely resembles the profile observed in KrasG12D;Trp53−/−;Dicer1−/− sarcoma cells [Ravi et al. 2012]. Mature miRNAs constituted 55% of total small RNA reads in Dicer wild-type MSCs but only 0.5% of total reads in Dicer knockout MSCs, similar to the degree of loss observed in sarcoma cells. Per cell, most miRNAs were present at fewer than one copy in Dicer knockout MSCs [Supplemental Fig. S1A]. Members of the let-7 family were reduced from thousands of copies per cell in Dicer wild-type MSCs to <10 copies per cell in Dicer knockout MSCs [Supplemental Fig. S1A].

Next, we characterized the consequences of miRNA loss on gene expression. By exon microarray, we observed specific gene expression differences between Dicer wild-type and knockout MSCs [Fig. 1A; Supplemental Table S2], with unsupervised hierarchical clustering distinguishing the cells by genotype [Supplemental Fig. S1B]. Two-hundred-seventeen genes were differentially expressed [adjusted P-value ≤0.1] [Fig. 1A]. The median change for both up-regulated and down-regulated genes was approximately threefold [Fig. 1A, gray dashed lines] with a maximum change of ~50-fold.

Conserved targets of conserved miRNAs, as predicted by TargetScan, were largely up-regulated in Dicer knockout MSCs [Fig. 1B], particularly for let-7, miR-199, and miR-15, which are three of the most highly expressed miRNA families in Dicer wild-type MSCs. We also observed Dicer loss-dependent up-regulation of targets of miR-202, which is not expressed in Dicer wild-type MSCs but shares a hexamer seed match with let-7. By gene set enrichment analysis (GSEA), we characterized the enrichment of miRNA target sites in the 3’UTRs of differentially expressed genes. Among genes up-regulated in Dicer knockout MSCs, 37 miRNA motifs were enriched at a false discovery rate (FDR) q-value ≤0.1, including let-7 [FDR q-value = 0.01] [Supplemental Fig. S1C], miR-199 [FDR q-value = 0.02], and miR-15 [FDR q-value = 0.03]. Among genes down-regulated in Dicer knockout MSCs, only two miRNA seed-match motifs [miR-339 and miR-517]
were identified at an FDR q-value ≤0.1. Thus, Dicer loss resulted in derepression of miRNA targets in MSCs.

**Dicer knockout fibroblasts retain mesenchymal cell identity**

Tissue-specific deletion of Dicer in vivo during development often results in differentiation defects or lethality, suggesting that miRNAs are critical in establishing or maintaining cellular identity. Therefore, we determined whether gene expression changes in Dicer knockout MSCs in vitro reflected a gross change in cellular identity. We compared the full gene expression profiles of Dicer wild-type and knockout MSCs with those of a diverse array of cell types ranging from ESCs to neural cells (Wu et al. 2009, 2013). Cells generally clustered by cell type, and the profile of Dicer wild-type MSCs correlated most closely with mesenchymal cells such as 3T3-L1 and C3H10(1/2) cells (Fig. 1C). Dicer knockout MSCs remained mesenchymal, correlating most closely with Dicer wild-type MSCs. Thus, MSCs retain their identity independent of miRNAs.

To test whether this observation could be extended to another cell type, we also analyzed Dicer wild-type and knockout ESCs generated from the same Dicer-conditional mouse model (Leung et al. 2011). Both Dicer wild-type and knockout ESCs correlated closely with independently derived, wild-type ESCs in the panel and clustered separately from differentiated cell types (Fig. 1C). In total, loss of miRNAs does not result in gross changes in pre-established cellular identity.

**Dicer knockout fibroblasts exhibit a let-7-regulated oncofetal signature**

Dicer knockout MSCs exhibit signatures for numerous miRNAs. From among these signatures, we focused on let-7 because it is an abundant tumor suppressor that represents the somatic counterpoint to the ESC-specific miR-290 family. In Dicer knockout MSCs, the top two statistically significant up-regulated genes were the oncofetal let-7 targets 
\[\text{Igf2bp1}\] (up 48-fold by microarray) and 
\[\text{Igf2bp2}\] (up 16-fold). By quantitative PCR (qPCR), we confirmed the up-regulation of these genes as well as 
\[\text{Igf2bp3}\] and 
\[\text{Hmga2}\], two additional oncofetal genes and targets of let-7 (Fig. 2A). The large magnitude of up-regulation for all four genes was striking, ranging by qPCR from 11-fold for 
\[\text{Hmga2}\] to 300-fold for 
\[\text{Igf2bp2}\]. Similarly, by Western blot, these genes were largely undetectable at the protein level in Dicer wild-type MSCs but became highly expressed in Dicer knockout MSCs (Fig. 2B). All four genes are let-7 targets, demonstrated by their down-regulation at the mRNA and protein levels by transfection with an siRNA duplex of let-7g (Fig. 2A,B). In contrast, neither 
\[\text{Lin28a}\] nor 
\[\text{Lin28b}\] was expressed in Dicer wild-type or knockout MSCs [Supplemental Table S2], suggesting that factors in addition to miRNA loss contribute to up-regulation of oncofetal genes.

Due to the magnitude of gene expression changes for these genes, we tested whether transcription contributes to this oncofetal signature. We carried out chromatin immunoprecipitation (ChIP) coupled with sequencing (ChIP-seq) for H3K4me3 and H3K36me3, associated with...
transcriptionally active genes, and H3K27me3, associated with transcriptionally inactive genes (Supplemental Table S3). We confirmed that genes associated with H3K4me3 and H3K36me3 marks were expressed highly at the mRNA level relative to all genes, while genes associated with H3K27me3 were expressed lowly relative to all genes (Supplemental Fig. S2A,B).

We inspected oncofetal genes for changes in these histone marks. For Igf2bp1 (Fig. 2C), Igf2bp2 (Supplemental Fig. S2E), and Igf2bp3 (Supplemental Fig. S2F), we observed broad peaks of H3K27me3 in Dicer wild-type MSCs, supporting the observation that these genes are “off” in miRNA-expressing cells. Conversely, in Dicer knockout MSCs, all three genes exhibited a loss of H3K27me3 and concomitant gains in promoter-associated H3K4me3 and gene body-associated H3K36me3, demonstrating transcriptional activation of this family of genes. In comparison, the ChIP-seq density for four flanking genes (Tra2a, Tra2b, Snf8, and Gip) was indistinguishable between Dicer wild-type and knockout cells, consistent with the observation that these control genes are not differentially expressed upon Dicer loss (mean fold change = 1.0, mean adjusted P-value = 0.88). Although Lin28a and Lin28b lost H3K27me3 density in Dicer knockout MSCs, they did not exhibit any detectable peaks of either H3K4me3 near their promoters or H3K36me3 in their gene bodies, suggesting that these two oncofetal genes are not transcriptionally activated, consistent with the absence of their expression by microarray.

A more global comparison of histone marks between genotypes also revealed Dicer-dependent changes in chromatin. Several thousand genotype-specific peaks for
H3K4me3 and H3K36me3 were enriched in one genotype relative to the other (Supplemental Table S3). Specifically, 205 and 981 genes were marked with genotype-specific H3K4me3 peaks in Dicer wild-type and Dicer knockout MSCs, respectively, and generally exhibited higher expression in their respective genotypes (Supplemental Fig. S2C). Similarly, 357 and 147 genes were marked with genotype-specific H3K36me3 peaks in Dicer wild-type and Dicer knockout MSCs, respectively, and generally exhibited higher expression in their respective genotypes (Supplemental Fig. S2D).

To determine whether transcription contributes to up-regulation of let-7 targets generally, we examined the expression of the overlap between predicted let-7 targets (813 genes) and genes enriched for H3K4me3 in Dicer knockout relative to Dicer wild-type MSCs (981 genes). These genes (63 genes) (Supplemental Table S7, “Let-7, H3K4me3 in Dicer KO”), which include the Igf2bp1-3 family but not Hmga2, were more highly expressed than all predicted let-7 targets (Fig. 2D, cf. “Let-7” and “Let-7, H3K4me3 in Dicer KO”) or all genes enriched for H3K4me3 in Dicer knockout relative to Dicer wild-type (Fig. 2D, cf. “H3K4me3 in Dicer KO” and “Let-7, H3K4me3 in Dicer KO”). In total, oncofetal let-7 targets are up-regulated, with transcriptional increases, in Dicer knockout MSCs.

Add-back of let-7 to Dicer knockout fibroblasts identifies direct targets

To expand our analysis of let-7 beyond known oncotel targets and validate additional targets among computatiional predictions, we experimentally identified targets by reconstituting let-7 expression in Dicer knockout MSCs. To do so, we carried out mRNA sequencing (mRNA-seq) on polyA-selected total RNA isolated 48 h after transfection of Dicer wild-type MSCs with non-targeting control siRNA (siCtrl) or transfection of Dicer knockout MSCs with either siCtrl or synthetic let-7g siRNA duplex as a representative member of the let-7 family (Supplemental Table S4).

Comparison of Dicer wild-type and knockout MSCs by mRNA-seq recapitulated the exon microarray results and demonstrated specific gene expression changes with Dicer loss [Supplemental Fig. S3A]. Relative to control-transfected Dicer wild-type MSCs, 1783 genes (FPKM [fragments per kilobase of exon per million fragments mapped] cutoff ≥ 0.1, q-value < 0.05) were up-regulated in control-transfected Dicer knockout MSCs with a median fold change of ~2, while 1746 genes (FPKM cutoff ≥ 0.1, q-value < 0.05) were down-regulated with a median fold change of ~2 (Supplemental Fig. S3A). miRNA targets were largely derepressed in Dicer knockout MSCs relative to Dicer wild-type MSCs (Supplemental Fig. S3B). More genes were differentially expressed by mRNA-seq than by exon microarray, probably due to the sensitivity of sequencing and the use of a single pair of Dicer wild-type and knockout clones in the mRNA-seq experiment, thus eliminating clonal heterogeneity.

Add-back of let-7 to Dicer knockout MSCs induced many changes in gene expression, including down-regulation of 1233 genes (FPKM cutoff ≥ 0.1, q-value < 0.05), with a median fold change of ~2, and the up-regulation of 1993 genes (FPKM cutoff ≥ 0.1, q-value < 0.05), with a median fold change of ~2 (Fig. 3A, Supplemental Fig. S3C). Let-7g-transfected cells clustered more closely with siCtrl-transfected Dicer knockout MSCs than with siCtrl-transfected Dicer wild-type MSCs, indicating that changes in addition to let-7 loss, such as miR-15-loss or miR-199-loss, contribute to the gene expression profile of Dicer loss [Supplemental Fig. S3D].

Analysis of expression changes of all predicted miRNA targets revealed a strong let-7 signature in let-7g-transfected Dicer knockout MSCs. The moving average of the TargetScan score for predicted let-7 targets increased significantly among genes that were most highly expressed in Dicer knockout siCtrl cells relative to siCtrl-transfected Dicer wild-type and let-7g-transfected Dicer knockout MSCs [Fig. 3A, side plot]. In contrast, no signal was observed in this moving average for predicted targets of miR-15, indicating specificity of the let-7g siRNA. Global analysis of median gene expression changes in Dicer knockout MSCs transfected with let-7g revealed statistically significant signals for targets of let-7; miR-202, which shares a 6-mer seed with the let-7 family, and miR-196, whose seed sequence begins at position 2 of the let-7 seed sequence and thus is a seed-shifted relative of let-7 [Supplemental Fig. S3E]. GSEA confirmed down-regulation of let-7 targets upon transfection of let-7g siRNA [Supplemental Fig. S3F].

Add-back of let-7 should result in two types of changes: (1) direct repression of let-7 targets and (2) secondary effects downstream from let-7 targets, including genes induced by let-7 add-back. To identify targets repressed directly by let-7, we overlapped genes that were up-regulated following Dicer loss by mRNA-seq at q-value <0.05 [Fig. 3B, left panel, “Derepressed in Dicer KO”], down-regulated with add-back of let-7g at q-value <0.05 [Fig. 3B, left panel, “Represse[d] with let-7 add-back in Dicer KO”], and predicted by TargetScan to be let-7 targets [Fig. 3B, left panel, “Predicted targets of let-7-7”]. This triple overlap identified 122 genes (listed in Supplemental Table S7) that included Hmga2 and the Igf2bp1-3 family discussed above, an enrichment of 28-fold above background. We randomly selected seven genes from this set for validation by qPCR and, for all seven genes, confirmed up-regulation upon Dicer loss and down-regulation following transfection of let-7g (Fig. 3C). However, repression of predicted targets by add-back of let-7 was only partially reversible and may reflect either transcriptional effects, a requirement for cotargeting by additional seed families, or transfection of synthetic let-7 at a concentration below endogenous levels. Hereafter, we refer to this set of 122 genes as “high-confidence” targets of let-7.

To test whether up-regulation of “high-confidence” targets could be generalized to a second cell type, we analyzed microarray gene expression data [Supplemental Table S5] from sarcoma cell lines that are KrasG12D;Trp53−/−; Dicer1−/− (heterozygous) and in which let-7 is the most abundant seed family or KrasG12D;Trp53−/−; Dicer1−/− (knockout) [Ravi et al. 2012]. Indeed, “high-confidence”
targets of let-7 were largely up-regulated with Dicer loss in sarcoma cells relative to control genes (Supplementary Fig. S3G). The oncofetal genes Igf2bp1 and Igf2bp3 were strongly up-regulated, as were many of the genes that we validated in MSCs by qPCR (Supplementary Fig. S3H, note the logarithmic Y-axis). Up-regulation of Hmgna2 and Igf2bp2 was not observed in sarcoma cells, likely due to already high basal expression in muscle (Li et al. 2012), the tissue type from which these cells were derived.

Thus far, we defined “high-confidence” targets of let-7 based on gene expression. To determine whether these genes are directly bound by the RNA-induced silencing complex (RISC) loaded with let-7, we carried out cross-linked immunoprecipitation (CLIP) sequencing (CLIP-seq) (Chi et al. 2009). To do so, we infected Dicer wild-type MSCs with lentivirus encoding doxycycline-inducible Flag-HA-Ago2 or untagged Ago2 as a negative control, selected transduced cells with hygromycin, and induced transgene expression with doxycycline. We then sequenced UV-cross-linked RNA that was isolated from Flag-HA-Ago2- and untagged Ago2-expressing cells by sequential anti-Flag and anti-HA immunoprecipitation (Supplementary Table S6).

Figure 3. Identification of let-7 targets by let-7 add-back to Dicer knockout (KO) MSCs. (A) Heat map of expression Z-scores in transfected Dicer wild-type (WT) and knockout MSCs for all genes expressed at FPKM ≥0.1. Genes are ranked from highest [top] to lowest [bottom] Z-score in Dicer knockout MSCs. A 10-gene moving average of TargetScan scores is shown in the left plot for let-7 [black] and miR-15 [gray]. [B] Overlap of genes up-regulated at q-value <0.05 in siCtrl-transfected Dicer knockout MSCs relative to siCtrl-transfected Dicer wild-type MSCs (“Derepressed in Dicer KO”), down-regulated at q-value <0.05 in let-7g-transfected Dicer knockout MSCs relative to siCtrl-transfected Dicer knockout MSCs (“Repressed with let-7 add-back in Dicer KO”), and predicted by TargetScan to be conserved targets of let-7 (“Predicted targets of let-7”). Top predicted let-7 targets in TargetScanMouse are shown in the table on the right. (C) qPCR analysis of high-confidence targets of let-7 identified in triple overlap above. Error bars indicate the SEM. (D) Metaplot of CLIP-seq read density. CLIP-seq reads were aligned to TargetScan-predicted sites. The average CLIP-seq read coverage per gene was plotted relative to let-7 target sites within the 3’ UTRs of high-confidence targets of let-7 (“HC, let-7”), let-7 target sites within the 3’ UTRs of all TargetScan-predicted targets of let-7 (“All, let-7”), or miR-124 target sites within the 3’ UTRs of all TargetScan-predicted targets of let-7 (“All, miR-124”). The analysis was carried out for genes with FPKM ≥0.1 in either Dicer wild-type or knockout MSCs transfected with siCtrl. (E) CLIP-seq density along the 3’ UTR of Nr6a1. The blue histogram indicates CLIP-seq density; gray tick marks indicate the location of TargetScan-predicted, conserved let-7 sites; and orange tick marks indicate the location of TargetScan-predicted, nonconserved let-7 sites. The length of the 3’ UTR is indicated. See also Supplemental Figures S3 and S4 and Supplemental Tables S4 and S6.
Let-7 regulates mid-gestation programs

In Flag-HA-Ago2-expressing cells, a metagene plot at TargetScan-predicted, conserved let-7 target sites demonstrated higher CLIP-seq signal per gene for “high-confidence” let-7 targets compared with all conserved TargetScan let-7 targets expressed ≥0.1 FPKM [Wilcoxon rank-sum test P-value = 1.4 × 10−22, Fig. 3D]. In contrast, no peak in signal was observed at target sites of miR-124, a miRNA not expressed in MSCs, within TargetScan-predicted let-7 sites [Fig. 3D] or at let-7 sites in CLIP-seq carried out from cells expressing untagged Ago2 [data not shown]. Inspection of individual 3′ UTRs demonstrated CLIP-seq peaks at both conserved and nonconserved TargetScan-predicted let-7 sites for Nr6a1 [Fig. 3E], a gene characterized in further detail below. We also observed peaks at let-7 target sites within the 3′ UTRs of Igf2bp1 (Supplemental Fig. S4A), Arid3a [Supplemental Fig. S4B], Pbx3 [Supplemental Fig. S4C], and numerous other genes, including those validated in Figure 3C.

Approximately 78% of “high-confidence” targets of let-7 exhibited five or more reads within 5 nt of their predicted let-7 target sites. For all TargetScan targets, ~63% of genes exhibited five or more reads at predicted let-7 target sites. In contrast, for both “high-confidence” let-7 targets and all predicted let-7 targets, only ~30% of genes possessing target sites for miR-124 exhibited five or more reads at predicted miR-124 sites. CLIP-seq from Dicer wild-type MSCs likely underestimates binding of the RISC to those genes that are not transcribed or are very highly repressed than let-7 upon let-7 add-back and are below the five-read cutoff by CLIP-seq in Dicer wild-type MSCs. For multiple genes, CLIP-seq peaks were also observed at predicted sites for miRNAs other than let-7, indicating regulation by multiple miRNA seed families. In total, we experimentally identified numerous let-7 targets functionally by let-7 add-back to Dicer knockout MSCs and independently confirmed that these genes are strongly enriched for Ago2-binding at let-7 sites.

Let-7 targets comprise a mid-gestation embryonic program

Since miRNAs may regulate networks of functionally related genes, we examined whether “high-confidence” targets of let-7 possess any common characteristics. Many of these genes, such as Igf2bp1–3, Hmga2, Pbx3, and Arid3b, are known to be expressed in the embryo [Hirning-Folz et al. 1998; Hansen et al. 2004; Takebe et al. 2006; Vitobello et al. 2011], in line with the hypothesis that mammalian let-7 regulates development. We also identified numerous oncogenes, including Plagl2, which is overexpressed in glioma and colorectal cancer [Zheng et al. 2010], and Arid3a, which drives bypass of RAS-induced senescence [Peep et al. 2002]. Furthermore, several “high-confidence” let-7 targets, including but not limited to Plagl2, Arid3a, and oncopetal genes, are known to peak in expression around mid-gestation (approximately embryonic day 8.5 [E8.5] to E10.5). Importantly, this timing distinguishes these genes from Oct4 and other pluripotency markers that are expressed in ESCs, down-regulated upon differentiation, and inactive in Dicer knockout MSCs.

To systematically determine whether “high-confidence” targets of let-7 exhibit a mid-gestation embryonic signature, we profiled these genes in a published time course of gene expression in whole mouse embryos [Irie and Kuratani 2011]. We restricted our analysis to genes that were expressed in at least one of the time points in the data set. The majority of “high-confidence” targets of let-7, including Hmga2 and the Igf2bp1–3 family, are expressed mid-gestation and steadily decrease as embryogenesis progresses [Fig. 4A]. To quantify these changes, we plotted the mean Z-score for each gene (the expression of each gene at each time point normalized to the mean expression and standard deviation of the gene across all time points). These genes peak in expression at E8.5–E10.5 and subsequently decline [Fig. 4B], thus anti-correlating with the expression of let-7, which becomes detectable by Northern blot in whole mouse embryos at E10.5, plateaus at E14.5, and remains high postnatally [Fig. 4B–E, dashed line; Schulman et al. 2005]. In contrast, the set of all conserved let-7 targets predicted by TargetScan and expressed in at least one time point in the whole mouse embryo data set was largely indistinguishable from background [Fig. 4C]. We confirmed the anti-correlation between let-7 and its “high-confidence” targets in a second, independent published time course of mouse embryonic limb bud development [Taher et al. 2011]. As before, the majority of “high-confidence” targets decreased as the embryonic limb bud developed [Supplemental Fig. S5A]. On average, these genes anti-correlated with let-7 [Supplemental Fig. S5B], while, in contrast, the set of all conserved let-7 targets predicted by TargetScan was largely indistinguishable from background [Supplemental Fig. S5C]. In total, members of the “high-confidence” let-7 target set peak in expression around E8.5 and anti-correlate with let-7 during development.

Add-back of let-7 in Dicer knockout MSCs also induces the expression of many genes, which are likely “indirect targets” downstream from genes repressed directly by let-7. In the whole mouse embryo, we tested whether “indirect targets” of let-7 exhibited a specific pattern. “Indirect targets” were induced as the embryo matured, correlating positively with let-7 and anti-correlating with “high-confidence” targets of let-7 [Fig. 4D]. This positive correlation was also observed in the mouse embryonic limb bud [Supplemental Fig. S5D]. In total, let-7 controls a mid-gestation embryonic program in somatic cells by directly repressing genes that peak mid-gestation and indirectly inducing genes that become activated as the embryo matures.

Nr6a1, an embryonic transcriptional repressor, is a target of let-7 and represses gene expression in Dicer knockout MSCs

The gene Nr6a1, also known as Germ cell nuclear factor (GCNF), was notable among “high-confidence” let-7 targets, since it meets the criteria of a classical developmental target of let-7. Nr6a1 is an embryonically expressed
sequence-specific transcriptional repressor and orphan nuclear receptor (Fuhrmann et al. 2001). Like many targets of let-7, such as the nematode nuclear steroid receptor daf-12 (Hammell et al. 2009), Nr6a1 is critical to development, during which it inhibits the pluripotency factor Oct4 (Gu et al. 2005). Germline deletion of Nr6a1 results in embryonic lethality mid-gestation due to failures in neural tube closure and somitogenesis (Chung et al. 2001). Nr6a1 is the top-predicted target of let-7 by TargetScan and possesses eight let-7 seed matches in its 3' UTR (Fig. 3B, right panel). By qPCR, we confirmed up-regulation of Nr6a1 with miRNA loss and its down-regulation following add-back of let-7 (Fig. 3C, boxed in red). Similarly, Nr6a1 is up-regulated 10-fold in Dicer knockout sarcoma cells (Supplemental Fig. S3H). We also demonstrated its interaction with the RISC at predicted let-7 target sites (Fig. 3E), indicating that Nr6a1 is a direct target of let-7. Finally, Nr6a1 is part of the mid-gestation signature regulated by let-7, with a peak in expression around E8.5 and a subsequent decrease anti-correlating with let-7 in the whole mouse embryo (Fig. 4E) and in the mouse embryonic limb bud (Supplemental Fig. S5E).

Having identified the transcriptional repressor Nr6a1 as part of a mid-gestation program regulated by let-7 in adult fibroblasts, we combined knockdown and overexpression experiments to determine whether Nr6a1 in turn regulates gene expression in Dicer knockout MSCs (Fig. 5A, Supplemental Table S4). We transfected Dicer knockout MSCs, which are “NR6A1-high,” with siRNA against Nr6a1 (Supplemental Fig. S6A) and profiled gene expression by mRNA-seq (Supplemental Fig. S6B). With a knockdown efficiency of ~75%, we observed differential expression of 1601 genes (FPKM cutoff $\geq 0.1$, q-value $<0.05$). Of these, 978 were up-regulated with a median fold change of 1.8 (Supplemental Fig. S6B). The observation that depletion of NR6A1 results in majority up-regulation of genes suggests that NR6A1 acts as a transcriptional repressor in Dicer knockout MSCs.

Next, to identify genes directly regulated by NR6A1 and uncouple it from accompanying gene expression changes observed with Dicer loss, we overexpressed Flag-HA-NR6A1 in Dicer wild-type MSCs (Supplemental Fig. S6C), which are “NR6A1-low.” By mRNA-seq, we observed differential expression of 104 genes (FPKM cutoff $\geq 0.1$, q-value $<0.05$), of which a majority were down-regulated with a median change of 5.5-fold (Supplemental Fig. S6D), consistent with a transcriptionally repressive activity for NR6A1. Next, we characterized the expression of these Flag-HA-NR6A1-responsive genes in Dicer knockout MSCs. To do so, we restricted our analysis to genes differentially expressed at a q-value $<0.05$ within each data set and expressed at an FPKM $\geq 0.1$ across all three data sets (leading to a total of 76 genes responsive to expression of Flag-HA-NR6A1, 1556 genes responsive to knockdown of...
Nr6a1, and 3431 genes responsive to Dicer loss). Flag-HA-NR6A1-responsive genes, listed in Supplemental Table S7, were largely repressed in Dicer knockout MSCs (Fig. 5B) and subsequently induced in Dicer knockout MSCs following knockdown of Nr6a1 (Fig. 5C), demonstrating that Nr6a1 is responsible for their down-regulation with miRNA loss. Thirty-two genes, defined hereafter as “NR6A1-responsive,” were common in the overlap of changes observed upon overexpression of Flag-HA-NR6A1, loss of Dicer, and knockdown of endogenous Nr6a1 (Fig. 5D; Supplemental Table S7), representing an enrichment of 12-fold over background. Of these 32 genes, 13 were also responsive to let-7 (Supplemental Table S7). This partial overlap suggests that the activity of Nr6a1 is not restricted to targets of let-7. Notably, given that Nr6a1 is a transcriptional repressor, it is unlikely to account for the transcriptional induction of let-7 targets, such as Igf2bp1, observed in Dicer knockout MSCs. Furthermore, overexpression of Nr6a1 alone in Dicer wild-type MSCs, which are “Nr6a1-low,” results in the differential expression of ~100 genes, in contrast to the >1000 transcriptional changes that we observed with Dicer loss. These observations suggest that other miRNA-regulated factors in addition to or in place of Nr6a1 contribute to changes in histone marks.

Since Nr6a1 is normally expressed during development, we examined the embryonic expression of NR6A1-responsive genes. These genes correlated inversely with Nr6a1, with low expression mid-gestation that increased as the whole mouse embryo (Fig. 5E) and embryonic limb bud [Supplemental Fig. S6E] matured. In total, we identified Nr6a1 as a let-7 target that mediates secondary transcriptional gene expression changes in Dicer knockout MSCs.

**Genome-wide binding profile identifies direct targets of NR6A1**

Having identified genes functionally responsive to Nr6a1, we next carried out ChIP-seq of Flag-HA-NR6A1 to...
identify genes that are directly bound [Supplemental Fig. S7A; Supplemental Tables S3, S7]. We identified 9223 enriched regions, corresponding to a total of 5210 bound or nearby genes [Fig. 6A]. Approximately 19% of peaks overlapped with promoters, defined as regions spanning 5 kb upstream of to 1 kb downstream from annotated transcription start sites (TSSs); 38% of peaks were in gene bodies; and 43% of peaks were intergenic [Fig. 6A,C]. Peaks in all three categories were strongly enriched for the sequence CAAG(G/T)TCA (Fig. 6B), reported previously to be part of the consensus recognized by NR6A1 [Yan et al. 1997]. We also identified additional enriched sequences [Supplemental Fig. S7B] that may be recognized by interacting partners of NR6A1. Interestingly, Flag-HA-NR6A1 was bound to three oncofetal genes (Hmga2, Igf2bp2, and Igf2bp3), suggesting cross-regulation of genes within mid-gestation programs. However, the mRNA levels of these oncofetal genes did not change upon induction of Flag-HA-NR6A1 or knockdown of Nr6a1, suggesting the presence of additional regulators responsible for their differential expression.

Of the 32 NR6A1-responsive genes that were repressed in Dicer knockout cells and derepressed following knockdown of NR6A1, 26 were bound by NR6A1 [Supplemental Table S7], representing an enrichment of 2.6-fold over background (χ² test, P-value = 6.0 × 10⁻⁷) [Supplemental Fig. S7C], and nine of these genes possessed the NR6A1 consensus motif in either the promoter, gene body, or distal intergenic region. Among the 32 NR6A1-responsive genes [Fig. 6D, “Responsive”], those that were bound by NR6A1 [Fig. 6D, “Responsive, bound”] were more strongly repressed upon overexpression of Flag-HA-NR6A1, particularly if the binding site possessed the NR6A1 consensus motif [Fig. 6D, “Responsive, bound, motif”]. Genes bound by NR6A1, especially in regions containing the consensus site, were more strongly derepressed upon knockdown of NR6A1 in Dicer knockout MSCs [Fig. 6E]. Finally, NR6A1-bound genes that were repressed upon overexpression of Flag-HA-NR6A1 anti-correlated with NR6A1 in the whole mouse embryo [Supplemental Fig. S7D] and in the mouse embryonic limb bud [Supplemental Fig. S7E]. In total, these results identify direct targets of NR6A1 in MSCs and demonstrate the anti-correlation of these targets with NR6A1 expression during embryonic development.

Discussion
We characterized gene expression in immortalized bone marrow-derived Dicer-deficient somatic fibroblasts and...
observed large-magnitude changes in gene expression following loss of miRNAs. A subset of these effects was reinforced transcriptionally, notably for three archetypal oncofetal genes \( \text{Igf2bp1–3} \), indicating that the large fold changes observed in Dicer knockout MSCs for this family are attributable to not only disruption of miRNA–mRNA interactions (often responsible for only modest effects), but also transcriptional feedback that amplifies expression of let-7 targets. Hence, microscale “fine-tuning” activity at the level of miRNA–mRNA interactions masks, in aggregate, macro-level transcriptional effects on gene expression.

Dicer knockout MSCs retain their mesenchymal identity, indicating that miRNAs do not primarily govern pre-established cell identity. This finding, in conjunction with the requirement for miRNAs during active differentiation and development, suggests a role for miRNAs primarily in the transition between cell states. Consistent with this possibility, miRNAs modulate the dynamics of gene expression to regulate cellular transitions and physiological robustness in numerous model systems [Herranz and Cohen 2010; Ebert and Sharp 2012]. If the nearly universally reported stress sensitivity of miRNA-deficient cells is reframed as a general requirement for miRNAs in stimulus response, regardless of the nature of the stimulus, then the observation that miRNA loss results in embryonic lethality is consistent with the notion that miRNA-deficient cells cannot respond properly to developmental stimuli.

In the context of the immortalized cell lines reported here, let-7 represses an embryonic program distinct from pluripotency and related instead to a mid-gestation network that may regulate proliferative and metabolic pathways. In addition to known oncofetal let-7 targets, this program includes oncopgenes that have not previously been associated with embryonic development or let-7, such as \( \text{Plgl2} \) [Zheng et al. 2010] and \( \text{Arid3a} \) [Peeper et al. 2002]. Both genes peak mid-gestation in the whole mouse embryo. Thus, these genes may further populate the let-7-regulated oncofetal network.

The up-regulation of a mid-gestation program in miRNA-deficient somatic cells suggests that the inactivation of these embryonic genes in adult tissue is maintained in a deliberately reversible state. Several of these genes, such as \( \text{Hmga2} \) and \( \text{Igf2bp1} \), play important roles in metabolism and proliferation [Viswanathan et al. 2009; Viswanathan and Daley 2010; Zhu et al. 2010; Frost and Olson 2011] and may be transiently activated, perhaps through down-regulation of let-7 or other miRNAs in somatic tissue, to promote injury repair or growth. This possibility is supported by a recent study demonstrating the importance of an HMGAA2–IGF2BP2 axis in muscle regeneration in adult animals [Li et al. 2012]. The overexpression of these oncofetal genes in tumors may reflect the inappropriate ectopic activation of an otherwise native, context-specific process regulated dynamically by miRNAs. Notably, the up-regulation of this discrete, embryonic program is compatible with mesenchymal cell identity, consistent with previous reports that gene expression consists of separable modules that can be coactivated in various permutations [Wong et al. 2008; Kim et al. 2010].

Our results also extend the current model of let-7 function in mammalian development. Our data suggest a stage in the embryo, from E8.5 to E10.5, in which the embryo has progressed well past a naïve ESC state, thus shutting off the mir-290 family, but has not yet globally activated let-7 [Fig. 7]. This possibility is consistent with recent findings that let-7 promotes development of the emerging mesoderm and ectoderm of mouse and \textit{Xenopus} embryos [Colas et al. 2012]. We postulate that let-7 plays a role beyond inhibition of pluripotency and represses mid-gestation programs to ensure a forward momentum during development, consistent with observations in nematodes that let-7 and lin-4 mutant animals reiterate post-pluripotency larval stages [Lee et al. 1993; Reinhart et al. 2000].

Related to the role of let-7 in developmental timing, we identified \( \text{Nr6a1} \), a potential functional ortholog of nematode \( \text{daf-12} \), as a target of let-7. In our study, \( \text{Nr6a1} \) represses genes that become activated as the embryo matures, suggesting that precise dosage of \( \text{Nr6a1} \) is critical to developmental timing. Consistent with this observation, overexpression of \( \text{Nr6a1} \) results in posterior defects and altered somite formation in \textit{Xenopus} embryos [David et al. 1998]. Thus, \( \text{Nr6a1} \) and let-7 are likely sequentially activated and mutually antagonistic to ensure the proper chronology of gene expression in the mid-gestation embryo.

In conclusion, our findings support the hypothesis that the tumor-suppressive properties of let-7 are coupled to its repression, in somatic tissues, of metabolic embryonic programs. Furthermore, the global loss of miRNAs in adult tissue leads to transcriptional effects that reinforce specific gene expression programs. From a therapeutic standpoint, the inhibition of tumor growth may not be feasible by the inactivation of only a single miRNA target and may instead require delivery of individual synthetic tumor-suppressive miRNAs to inactivate full gene expression networks.

**Figure 7.** Summary of let-7 and target expression in the whole mouse embryo. Let-7 targets peak mid-gestation [around E8.5–E10.5], after down-regulation of the ESC-specific miR-290 family. Expression of mature let-7 becomes detectable in the whole embryo around E10.5 and steadily increases in level, concomitant with down-regulation of high-confidence let-7 targets such as \( \text{Nr6a1} \), identified in this study. Targets of \( \text{Nr6a1} \) identified in this study increase as the embryo matures and positively correlate with let-7 in the whole mouse embryo.
Materials and methods

Complete protocols are provided in the Supplemental Material and are also available on request. Microarray and sequencing data are available under Gene Expression Omnibus accession number GSE44163.

Cell culture conditions

Murine MSCs were passaged in α-MEM supplemented with 10% FBS and penicillin/streptomycin as described previously (Gurtan et al. 2012, Ravi et al. 2012). Sarcoma cells were passaged in MEM supplemented with 10% FBS, L-glutamine, and penicillin/streptomycin as described previously (Ravi et al. 2012).

Transfections and infections

For siRNA transfections, cells were transfected with HiPerfect reagent (Qiagen) and 20 nM siCtrl [Non-Targeting siRNA No. 2, Thermo Scientific Dharmacon], synthetic let-7g siRNA synthesized as perfectly complementary siRNA duplex [custom RNA synthesis service, Thermo Scientific Dharmacon], or gene-specific siRNA against Nr6a1 (Qiagen). Forty-eight hours after transfection, cells were harvested. Total RNA was isolated with an RNaseasy kit (Qiagen) for qPCR or mRNA-seq. Total protein was isolated with radio immunoprecipitation assay (RIPA) buffer supplemented with protease inhibitor tablets (Roche). siRNA experiments were carried out in biological triplicate in a single pair of clonal isogenic MSCs.

For infections, eNX (Phoenix) cells were transfected with Lipofectamine 2000 [Invitrogen] and either pMMP-puro-vector or pMMP-puro-Flag-HA-NR6A1 plasmid. Nr6a1 cDNA [isoform 1] was cloned from Dicer knockout MSCs. Viral supernatant was cleared with a 0.45-μm syringe filter. Subconfluent MSCs were incubated with viral supernatant at a multiplicity of infection ≥1 and 8 μg of polybrene (Sigma) overnight and then selected with 2.5 μg of puromycin. Prior to being used for experiments, transduced MSCs were passaged under selection for 2–4 wk. Infections were carried out on two independent clones of Dicer wild-type MSCs, thus representing biological duplicates.

Western blot

RIPA lysates were diluted twofold with 2× Laemml loading buffer with 5% β-mercaptoethanol and then boiled for 10 min. Samples were separated in 4%–12% Bis-Tris denaturing polybuffer with 5% RIPA lysates were diluted twofold with 2× Laemml loading buffer with 5% β-mercaptoethanol and then boiled for 10 min. Samples were separated in 4%–12% Bis-Tris denaturing poly-buffer with 5% RIPA lysates were diluted twofold with 2× Laemml loading buffer with 5% β-mercaptoethanol and then boiled for 10 min. Samples were separated in 4%–12% Bis-Tris denaturing poly

qPCR

Total RNA was treated with a TURBO DNA-free kit [Ambion]. cDNA was generated with oligo-dT primer using a SuperScript III first strand synthesis system [Invitrogen]. Gene expression was analyzed with Power SYBR Green [Applied Biosystems].

Microarray analysis

For MSCs, two clones of Dicer wild-type and their four derivative clones of Dicer knockout cells were grown to confluence in six-well plates, after which total RNA was prepared with QIAzol (Qiagen). A third Dicer wild-type clone, from which no Dicer knockout cells were available, was also analyzed initially but then subsequently excluded as an outlier. For sarcoma samples, Dicer heterozygous and Dicer knockout cells were grown to confluence in T25s, and total RNA was prepared with an RNeasy kit (Qiagen). Biotinylated cRNA was prepared using the Affymetrix GeneChip WT Sense Target Labeling and Control Reagents kit and hybridized to Affymetrix Mouse 430 2.3 arrays (sarcoma cell lines) or Mouse MoEx-1.0-st exon arrays [MSCs] at the BioMicroCenter at MIT.

RNA-seq

Cloning and sequencing of small RNAs were carried out as described previously (Gurtan et al. 2012, Ravi et al. 2012). mRNA-seq was carried out with total RNA purified with the RNaseasy reagent (Qiagen) and DNase-treated with TURBO DNase (Ambion). Samples were prepared for Illumina sequencing at the BioMicroCenter at MIT.

ChIP-seq

ChIP experiments were carried out as described previously (Rahl et al. 2010). In summary, Dicer wild-type and Dicer knockout MSCs were grown as described above and cross-linked for 10 min at room temperature by the addition of one-tenth of the volume of 11% formaldehyde solution [11% formaldehyde, 50 mM HEPES at pH 7.3, 100 mM NaCl, 1 mM EDTA at pH 8.0, 0.5 mM EGTA at pH 8.0] to the growth medium. Cells were washed twice with PBS, supernatant was aspirated, and the cell pellet was flash-frozen in liquid nitrogen. Frozen cross-linked cells were stored at −80°C. For histone mark immunoprecipitation, the following antibodies were used: Histone H3K4me3, Millipore 07-473, lot number DAM1731494 [7.5 μl per ChIP]; Histone H3K36me3, Abcam ab9050-100, lot number 136352; Histone H3K27me3, Abcam ab6002-100, lot number 49749. For immunoprecipitation of Flag-HA-NR6A1, anti-Flag (Roche, 11867423001) was used.

Acknowledgments

We thank members of the Sharp laboratory for helpful discussions and experimental assistance. Specifically, we thank Timothy Kelly for generating the doxycycline-inducible Flag-HA-Ago2 construct, and Jesse Zamudio and Mohini Jangi for help with Flag-HA-NR6A1 ChIP. We thank Alla Leshinsky, Richard Cook, and the rest of the Swanson Biotechnology Center at the David H. Koch Institute for Integrative Cancer Research at MIT for Illumina and other sequencing services. We thank Stuart Levine and the staff of the BioMicroCenter at MIT for preparation of RNA-seq and ChIP-seq samples. This work was supported by NIH grant R01-GM34277 [to P.A.S.], NCI grant PO1-CA42063 [to P.A.S.], R01-HG002668 [to R.A.Y.], and the NCI Cancer Center Support (core) grant P50-CA140551. A.M.G. acknowledges support from the Leukemia and Lymphoma Society Grant 5198-09. A.M.G., with contributions from A.R., designed the study. A.M.G., A.R., P.B.R., A.D.B., and C.K.J. performed experiments. P.B.R., under the supervision of R.A.Y., carried out ChIP of
histone marks. A.D.B. designed and carried out the CLIP-seq experiments. A.M.G., A.R., A.D.B., A.B., and C.A.W. carried out computational analyses. A.B. and C.A.W. performed the informatics for the microarray and sequencing data. A.M.G. wrote the paper. P.A.S. provided supervision and assisted with manuscript preparation. All authors reviewed and approved the manuscript.

References


