Genome-wide impact of a recently expanded microRNA cluster in mouse

Grace X. Y. Zheng[a,b,c,1,2], Arvind Ravi[a,c,d,1], Genevieve M. Gould[c], Christopher B. Burge[b,c], and Phillip A. Sharp[a,c,3]

[a]MIT Koch Institute for Integrative Cancer Research, Cambridge, MA 02139; [b]Computational and Systems Biology Graduate Program, Cambridge, MA 02139; [c]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139; and [d]Harvard-MIT Health Sciences and Technology Program, Cambridge, MA 02139

Contributed by Phillip A. Sharp, August 10, 2011 (sent for review April 6, 2011)

Variations in microRNA (miRNA) gene and/or target repertoire are likely to be key drivers of phenotypic differences between species. To better understand these changes, we developed a computational method that identifies signatures of species-specific target site gain and loss associated with miRNA acquisition. Interestingly, several of the miRNAs implicated in mouse 3′ UTR evolution derive from a single rapidly expanded rodent-specific miRNA cluster. Located in the intron of *Sfmbt2*, a maternally imprinted polycomb gene, these miRNAs (referred to as the *Sfmbt2* cluster) are expressed in both embryonic stem cells and the placenta. One abundant miRNA from the cluster, miR-467a, functionally overlaps with the mir-290-295 cluster in promoting growth and survival of mouse embryonic stem cells. Predicted novel targets of the remaining cluster members are enriched in pathways regulating cell survival. Two relevant species-specific target candidates, Lats2 and Dedd2, were validated in cultured cells. We suggest that the rapid evolution of the *Sfmbt2* cluster may be a result of intersex conflict for growth regulation in early mammalian development and could provide a general model for the genomic response to acquisition of miRNAs and similar regulatory factors.

The emergence of novel regulatory interactions provides a critical means of evolutionary change (1). By introducing new regulatory elements, or simply rewiring existing ones, organisms can adapt to alterations in their environment. In the case of protein coding genes, a number of precedents for these principles have been established. For instance, the transcriptional cofactor TAFII105 emerged in mammals to specifically direct expression of a subset of genes in ovarian follicle cells (2). Similarly, the remapping of existing transcription factor networks through promoter evolution is thought to be widespread, even between similar species (3, 4). However, given the relatively small increases in protein coding genes across more complex organisms, many have turned their attention to the roles of noncoding RNAs in explaining evolutionary changes (5).

Among noncoding RNAs, miRNAs are thought to be particularly relevant to phenotypic differences between species, with some claiming that miRNA gene number scales roughly with organismal complexity (6). Although only approximately 22 nucleotides in length, miRNAs can repress the expression of hundreds of genes posttranscriptionally, making them ideal candidates for the establishment or alteration of large regulatory networks. Indeed, it has been suggested that miRNAs are in fact more "evolvable" elements than transcription factors because targeting of a novel sequence requires changing only one or a few bases rather than a complex set of amino acids (4). However, the constraints of processing require that precursors be present as hairpin structures in the genome, therefore favoring their emergence via certain evolutionary routes. Three mechanisms in particular have been hypothesized for miRNA generation: (*i*) duplication of existing miRNAs; (*ii*) processing of transposable elements with terminal inverted repeats; or (*iii*) processing of hairpin structures generated by mutation (7). Following gene duplication in the first mechanism, targeting could be altered by subsequent base substitutions, particularly those corresponding to positions 2–7 from the 5′ end of the miRNA, known as the miRNA "seed" (8).

In contrast to miRNA evolution, which can simultaneously introduce multiple novel interactions, mutations in single target sites can provide more modest increments of evolutionary change. Nevertheless, these changes could also be important drivers of organismal differences. For instance, variations in Nodal family targeting by the miR-430/427/302 family guide differences in germ layer specification during development across a range of vertebrates (9). Even within a species, presence or absence of even a single target interaction may have notable effects. In the case of Texel sheep, a forward genetic screen identified a single base change in the myostatin 3′ UTR that creates a miRNA target site and confers muscular hypertrophy (10). In humans, a similar case has been reported for Tourette syndrome, where changes in the SLITRK1 transcript enhance repression by hsa-miR-189 (11). A more global assessment using human SNP datasets identified a number of variants that may alter miRNA binding (12). This study suggested that many target sites in the human genome are under purifying selection to maintain the presence or absence of a miRNA target site but only identified a single instance of positive selection, presumably because few SNPs passed the high heterozygosity thresholds required for informative intraspecies analysis.

Here we describe a method for identifying species-specific changes in miRNA-target relationships and use it to identify miRNA innovations in the mouse genome. We find that many of the mouse-specific changes correspond to a single genomic locus, located in the intron of an imprinted polycomb group gene (13). The miRNAs in this cluster, which we refer to as the *Sfmbt2* cluster, appear to have expanded through a duplication-divergence mechanism, generating both novel seeds and seeds corresponding to earlier miRNA families. Finally, predictions suggest that these miRNAs may in part regulate targets involved in growth and survival, in line with predicted roles for the mir-290-295 cluster, the dominant cluster in mouse ES cells. As the expression patterns of these two clusters appear to mirror one another, we suggest that the *Sfmbt2* cluster may promote proliferation in extraembryonic tissue, serving as a counterpart to the mir-290-295 cluster in early murine development.

Results

Positive Selection Acts on Target Sites of Many Species-Specific miRNAs. To better understand how gene networks respond to miRNA

Author contributions: G.X.Y.Z., A.R., C.B.B., and P.A.S. designed research; G.X.Y.Z., A.R., and G.M.G. performed research; G.X.Y.Z., A.R., G.M.G., C.B.B., and P.A.S. contributed new reagents/analytic tools; G.X.Y.Z. and A.R. analyzed data; and G.X.Y.Z., A.R., C.B.B., and P.A.S. wrote the paper.

The authors declare no conflict of interest.

[1]G.X.Y.Z. and A.R. contributed equally to this work.

[2]Present address: Howard Hughes Medical Institute and Program in Epithelial Biology, Stanford University School of Medicine, Stanford, CA 94305

[3]To whom correspondence should be addressed. E-mail: sharppa@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1112772108/-/DCSupplemental.

introduction, we examined the 3′ UTR landscape of the mouse genome for signatures that represent species-specific responses to different miRNA repertoires, using relative sequence conservation with humans as a background for comparison. Comparative analysis of 3′ UTRs should reveal two groups of genes associated with species-specific miRNAs: (*i*) genes that gained binding sites to novel miRNAs and whose downregulation improved fitness; and (*ii*) genes that were inadvertently targeted by novel miRNAs and that subsequently lost target sites to maintain expression.

To test for the presence of each group of genes, aligned 3′ UTRs from the mouse and human genomes were analyzed for signatures of target site gain and loss (Fig. 1*A*). For target site gain, we considered seed match sites present in mouse 3′ UTRs, whereas for site loss, we considered the reverse scenario, requiring a seed match site in human but not necessarily in mouse UTRs. While these putative assignments of site gain or loss in a given species are the most evolutionarily parsimonious, it is possible that some resulted from loss and gain, respectively, in the second species used for comparison. Despite this caveat, functional species-specific miRNAs would be expected to generate both gain and loss signatures at the genome-wide scale. Although sites with mismatches in the seed region have been described, including compensatory downstream matches (8) and centered pairing sites (14), we defined predicted targets using the traditional method of seed complementarity (7mer-M8 type sites), as it is the most common type of target site and the best characterized to date (8). For miRNA targets that have been positively selected to gain a target site, we would expect higher nucleotide variation between species at the miRNA binding site than at neighboring sequences in the same 3′ UTR (Fig. 1 *A* and *B*, *Left*). To measure this difference, we defined the statistic ***d*** as the difference in average per base mutation frequency between the miRNA target site and the frequency in the adjacent 42 nt upstream and 42 nt downstream of the target site (Fig. 1*A* and *B*, *Middle*). Significance of selection was then assessed by comparing ***d*** for a given seed to that of control oligonucleotides with similar abundance and composition in mouse and human 3′ UTRs (Fig. 1 *A* and *B*, *Right* and *SI Appendix*). Using this approach, 25 of 201 mouse-specific miRNA seeds (12.4%) showed a signal for gain of target sites (Fig. 2*A* and *SI Appendix*), roughly twice the rate observed for all heptamers (6.02%, $p = 0.004$, Chi-square test). Thus, our analysis detected evidence of selection for acquisition of seed matches to these mouse-specific miRNAs in the mouse genome. As an additional control, evaluation of the reverse complements of these seeds failed to show enrichment in any of our tests, consistent with selection being sensitive to the transcribed strand.

We additionally tested for site loss (Fig. 1*C*). Again, we found a number of miRNAs ($16/201 = 7.96\%$) whose target sites were lost at greater frequencies in mouse 3′ UTRs than controls, but this fraction was not significantly higher than the fraction observed for control heptamers (6.02%) (Fig. 2*A* and *SI Appendix*). Of these 16, a significant fraction, 13, overlapped with those seeds

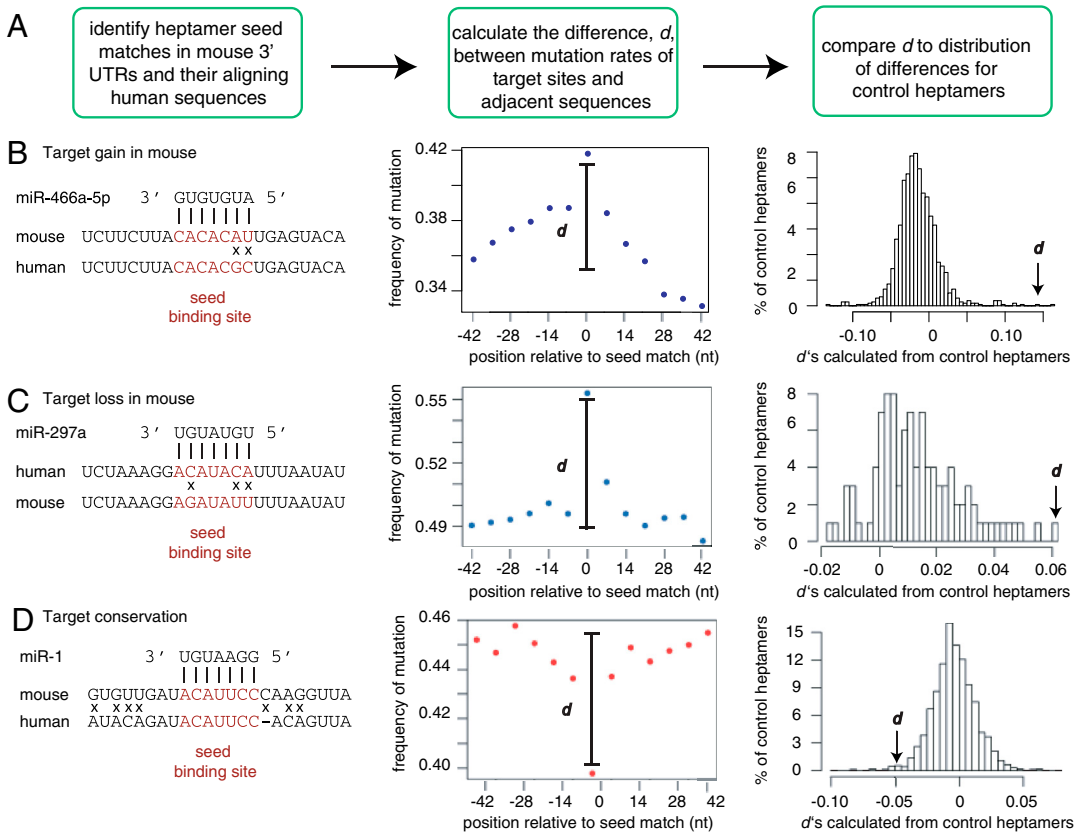

Fig. 1. Positive and purifying selection analysis for miRNAs in mouse and human. (*A*) A flow chart of the analysis: (*i*) seed match sites in aligned mouse and human 3′ UTRs were identified; (*ii*) the difference in average per base mutation rate, noted as ***d***, was compared between the miRNA binding site and adjacent sequences; and (*iii*) ***d*** was compared to that of control heptamers. (*B*) Illustration of the analysis of miR-466a-5p, an example of a mouse-specific miRNA. *Left*: the alignment of mouse and human 3′ UTRs of a target of miR-466a-5p, where the target site is gained in mouse. Seed binding site is colored in dark red. *Middle*: mutation rate of the 84-nt sequence spanning the miR-297a-5p match site. *Right*: distribution of ***d***'s from control heptamers. The black arrow points at the ***d*** for miR-297a-5p match sites. (*C*) Illustration of the analysis of miR-297a-5p, where the target site is lost in mouse. (*D*) Illustration of the analysis of miR-1, a conserved miRNA between mouse and human. Frequency of mutation denotes the fraction of mismatched nucleotides in the 7-nt window. "0" indicates the miRNA binding site, "42" refers to 42-nt downstream of the miRNA binding site, "-42" refers to 42-nt upstream of the miRNA binding site.

showing site gain signatures above (13/201, $p < 0.0001$, Fisher's exact test). Target sites of the 7mer-M8 type generally confer somewhat greater repression at the mRNA level than 7mer-A1 sites (15). Similar results were obtained when considering 7mer-A1 miRNA targets, albeit with a weaker signal (see *SI Appendix*). Human-specific gain/loss events were not statistically enriched over the frequency seen for control heptamers, which may reflect the less refined miRNA annotations presently available in humans (see *SI Appendix*).

As a validation of our method, we tested the behavior of target sites for miRNAs known to be conserved between mouse and human, predicting that these would show trends opposite to that observed for nonconserved miRNAs, namely a decreased mutation rate at the seed binding site relative to adjacent sequence (negative ***d***) (Fig. 1*D*). Using these criteria, of 218 miRNA seeds that are shared between human and mouse, target sites of 44 (20.18%) seeds showed a significant purifying selection signal (see *SI Appendix*), a percentage substantially higher than that observed for all heptamers (4.79%, $p = 0.001$, Chi-square test). These 44 miRNAs have an average signal to background ratio for site conservation of 3.39:1 based on previous calculations by Friedman and colleagues (16), confirming the ability of our method to recognize genomic signatures of selection. As expected, a comparison of the distribution of ***d*** for mouse miRNAs showing a target conservation signature between mouse and human showed a downward skew in seed match mutation rate relative to those with targets undergoing positive selection (see *SI Appendix*).

Target Sites of Many Mouse-Specific *Sfmbt2* Cluster miRNAs Are Under Positive Selection. We next evaluated whether the above 28 seeds (the unique set showing either gain or loss signals above) had any interesting contextual relationships in the mouse genome so that we could better discern evolutionary "hot spots" for miRNA diversity. Interestingly, when the 28 miRNAs with significant target site selection were plotted on a chromosome map of the mouse genome, a prominent clustering was apparent on chromosome 2. Five additional regions on chromosomes 2, 3, 4, 12, and X also contained miRNAs significant for both gain and loss of target sites, suggesting that these miRNAs may be functional (see *SI Appendix*). The chromosome 12 miRNAs are part of a cluster of 40 miRNAs near the imprinted Dlk-Gtl2 region (17). These miRNAs show preferential expression in the placenta and brain and have been suggested to have key roles in embryonic development and neurogenesis (18). Examination of miRNA expression data revealed that the remaining miRNAs outside of chromosome 12 with both gain and loss signals are expressed in germ tissue (miR-511) or during early embryonic development (miR-1198-5p, miR-3094, miR-380, miR-466k) (19), although detailed characterization of their expression and functions is minimal.

A

		count	total	count/total (%)	p-value
Site Gain	all heptamers	987	16,384	6.02	-
	mouse-specific miRNA M2-8 seeds	25	201	12.44	0.004
Site Loss	all heptamers	987	16,384	6.02	-
	mouse-specific miRNA M2-8 seeds	16	201	7.96	0.283

B

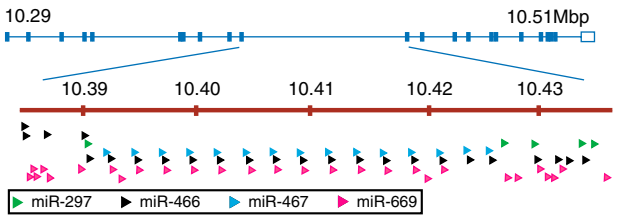

Fig. 2. Target sites of 28 mouse miRNAs show significant positive selection signals, and many fall into the *Sfmbt2* locus. (*A*) Summary statistics of positive selection signals detected from mouse miRNA M8 heptamer binding sites in mouse and human 3′ UTRs. "Count" represents the number of heptamers that were significant in the analysis. "Total" represents the total number of possible heptamers in the specific test category. *P*-values were the result of Chi-square tests. There are 16,384 heptamers in total. There are 201 mouse-specific miRNA seeds (dog and horse miRNAs were used as an outgroup to determine the species specificity). (*B*) Genomic structure of the *Sfmbt2* miRNA cluster. Precursors of *Sfmbt2* miRNAs were mapped to the 10th intron of the *Sfmbt2* gene based on the coordinates of precursors. The intron spans a region from roughly 10.39 Mbp to 10.44 Mbp on Chromosome 2. The *Sfmbt2* miRNAs are color coded as follows: miR-297s, green; miR-466s, black; miR-467s, blue; miR-669s, magenta.

Closer inspection of the largest cluster of hits, located proximally on chromosome 2, revealed that they were derived from the 10th intron of *Sfmbt2*, a polycomb group gene (20). In total, this intron contains 36 distinct miRNAs, which we refer to as the *Sfmbt2* cluster miRNAs (Fig. 2*B* and *SI Appendix*). Although the coding region of the *Sfmbt2* gene is highly conserved among vertebrates, the intron that harbors the miRNA cluster bears little similarity to those outside of rodent species. Two of the 36 mouse *Sfmbt2* miRNAs can be mapped to the homologous intron in rat, but none can be aligned to the corresponding intron in human. Consistent with these findings, expression of the *Sfmbt2* cluster—cloned from both T cells (21) and embryonic stem cells (22)—has only been detected in mouse tissue.

Homology of the intronic sequence to the SINE B4 repeat suggests that the cluster derived from a seeding transposition event followed by segmental duplications (see *SI Appendix*). Consistent with this model, the miRNAs contained within the cluster can be classified into a handful of broad groups based on sequence similarity: miR-297s, miR-466s, miR-467s, and miR-669s (Fig. 2*B*). Of these four classes, members of the miR-297s and miR-467s can be considered miRNA "families" because they largely consist of the seeds "UGUAUG" and "AAGUGC," respectively, whereas the miR-466s are a "superfamily" of the shifted seeds "GAUGUG," "AUGUGU," and "UGUGUG," and the miR-669s contain a diversity of seeds. Plausible reconstruction of the evolutionary history of this cluster suggests that miR-669 elements were most similar to the ancestral sequence and that subsequent duplications gave rise to a miR-466 precursor, from which both the miR-467s and miR-297s are derived (see *SI Appendix*). Sequence variation within these four classes in the *Sfmbt2* cluster is sufficient to generate 23 distinct miRNA seeds, the majority of which are not found in humans.

Experimental characterization of the *Sfmbt2* miRNA cluster confirmed that they function like canonical miRNAs because they are Dicer dependent (see *SI Appendix*), repress luciferase targets (see *SI Appendix*), and appear to globally destabilize messages containing seed complementarity (see *SI Appendix*). Interestingly, these repressive signatures were evident even for miRNAs without significant evolutionary signatures, although those miRNAs that were identified in our earlier analysis with greater seed conservation (see *SI Appendix*) showed increased significance, suggesting they may have a greater number of functional targets.

miR-467a, an Abundant Member of the *Sfmbt2* Cluster, can Promote Cell Proliferation. To dissect the functions of the *Sfmbt2* cluster, we began by examining the represented seeds and their potential targets. Sequencing data revealed the miR-467 family to be the most abundantly expressed miRNAs from the cluster (23–25). Notably, this family shares the hexamer seed "AAGUGC" with the conserved miR-290-295 and miR-302 clusters, expressed predominantly in ES cells. Because of this seed overlap, any species-specific signal was likely masked by a residual conservation signal in our earlier computational analysis. However, given their independent derivation from an ancestral miR-669 sequence, the

miR-467 family is best described as species-specific. Given the previously described cell cycle regulatory roles of miR-290-295 and its human counterpart, miR-371-373 (26, 27), we tested whether miR-467 family members might function similarly. Indeed, miR-467a was able to repress luciferase reporters for previously validated miR-295 targets, such as p21 and Lats2 (see *SI Appendix*) and was comparable to miR-295 in rescuing the G1/S delay of Dicer null cells (Fig. 3*A*). In addition, miR-467a recapitulated a novel antiapoptotic phenotype we have recently described for miR-295 following genotoxic stress (28) (Fig. 3 *B* and *C*). Taken together, these results suggest that miR-467 family members could reinforce proliferative cellular programs in parallel with miR-295.

Examination of available global miRNA expression data suggested that in addition to ES cells, placental tissue shows high expression of *Sfmbt2* miRNAs (29). Comparing the relative expression of miR-467a by Northern blot analysis, there appeared to be a fourfold upregulation in placental tissue relative to ES cells (Fig. 4*A* and *SI Appendix*). Similar expression trends were observed with representatives of the miR-297 (sevenfold increase) and miR-466 (2.5-fold increase) classes (Fig. 4*A* and *SI Appendix*). These findings are consistent with mRNA expression data for *Sfmbt2* (30), which has been shown to be maternally imprinted (and therefore paternally expressed) in the placenta (20). Therefore, it is likely that expression of the intronic miRNA cluster is dependent on transcription of the entire gene, as has been suggested for many miRNAs similarly positioned within host genes (31). In contrast to *Sfmbt2* miRNAs, miR-295, which shares the AAGUGC seed with miR-467a and is highly expressed in ES cells, was barely detectable in the placenta (Fig. 4*A* and *SI Appendix*). The differing expression patterns of these miRNAs, also confirmed by a global miRNA profiling dataset (*SI Appendix*) (29), suggest that the miR-295 family may be principally responsible for the proliferation of ES cells while some members of the *Sfmbt2* miRNAs may regulate placental growth.

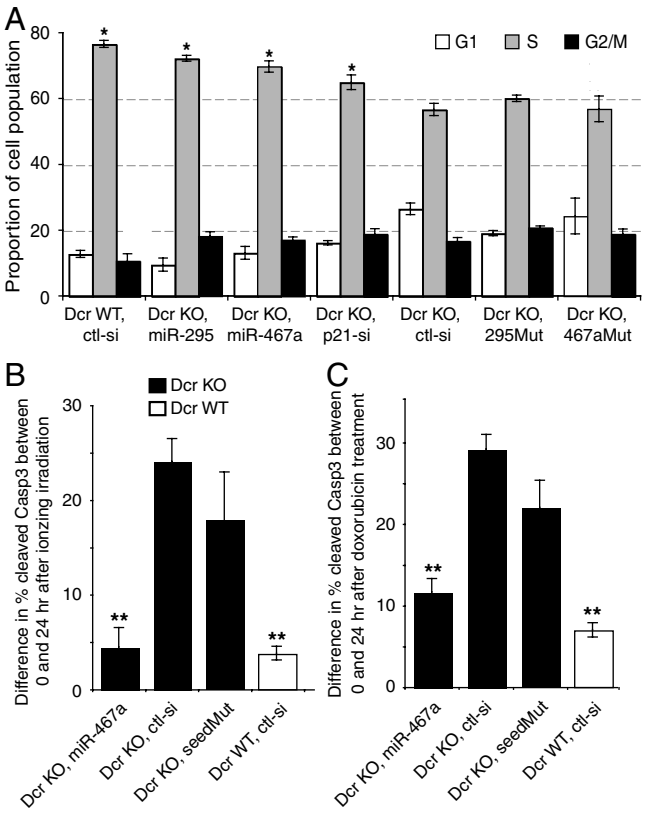

Fig. 3. miR-467 can promote growth and survival of mouse ES cells. (A) *Dcr* KO ES cells were transfected with 50 nM miR-467, siRNA against p21, and other control siRNAs. Twenty-four h after transfection, cells were incubated with BrdU for 10 min, and BrdU positive cells were analyzed with flow cytometry. Assays with miR-467a seed mutant (467aMut), miR-295 seed mutant (295Mut) and control siRNAs served as negative controls. Results are percentages in each stage of the cell cycle and are shown as mean ± S.E.M. (standard error of the mean). $n = 2$ for miR-467a seed mutant and p21 siRNA transfections. $n \geq 3$ for remaining transfections. *P*-values were results of Mann–Whitney tests, and * denotes $p \leq 0.05$. (*B*) ES cells were treated with 5-Gy radiation 24 h after transfection of 50 nM miR-467a or a control siRNA. Casp3 activity was assayed 0 and 24 h after treatment, and the difference in apoptosis rate is shown. Apoptosis rate of *Dcr* KO cells is shown in black bars, and that of WT cells is shown in white bars. (*C*) A similar series of experiments as *B* was performed in *Dcr* KO and WT cells using 100 nM doxorubicin. $n \geq 3$ for all experiments. Results are shown as mean ± S.E.M. *P*-values were results of Mann–Whitney tests, ** denotes $p \leq 0.01$.

Targets of *Sfmbt2* miRNA Families Are Enriched in Pathways that Regulate Growth. Because the remaining members of the *Sfmbt2* cluster that are associated with genome-wide signals of selection have seeds unique to mice, we were able to predict potential mouse-specific target transcripts. As defined above, these target sites should have significantly higher mutation rates than adjacent sequences. We assessed their significance with the binomial test and identified 511 placentally expressed genes whose target sites have probably been specifically selected for in mouse ($p \leq 0.05$, FDR < 0.40) (see *SI Appendix*). Enrichment in annotated pathways and functions was tested for all 511 genes using ingenuity pathway analysis, with the top statistically significant categories including "Cell Death" ($p = 6.7 \times 10^{-5}$ to 4.6×10^{-2}) and "Cell Cycle" ($p = 2.1 \times 10^{-4}$ to 4.7×10^{-2}). Given the similarity between these annotated functions and known functions of the miR-290-295 cluster, we compared this list to AAGUGC targets to determine whether significant cotargeting was evident. Indeed, a statistically significant overlap of 157 common genes ($p = 3.7 \times 10^{-5}$, hypergeometric test) was present, supporting the notion that these two miRNA clusters were selected in part for similar functional roles (Fig. 4*B*).

To explore targets relevant to our pathway analysis results, we tested whether two previously well characterized proliferation-related genes contain functional mouse-specific target sites: Dedd2 ($p = 0.16$, FDR < 65%) and Lats2 ($p = 0.23$, FDR < 57%). Dedd2 is a well conserved inducer of apoptosis across a variety of cell types, known to associate with Caspase 8 and Caspase 10 via its Death Effector Domain (DED) (32). Lats2, a target of the miR-290-295 cluster (26) and its homologous cluster in human (33), participates in both cell cycle progression (34) and apoptosis (35), in part through a positive feedback loop with p53 (36). Both genes show acquisition of target sites for *Sfmbt2* cluster miRNAs specifically in mouse with a mutation profile higher than surrounding 3′ UTR regions (Fig. 4*C*). Reporters containing the 3′ UTRs of these genes fused to luciferase were transfected into *Dicer* KO cells with and without miRNAs predicted to have gained target sites in mice. Of the two potential target sites acquired by Lats2, the miR-466f-3p site showed an approximately 25% repression relative to a site mutant construct, while the miR-669-3p site appeared to be nonfunctional (Fig. 4*D*). In the case of Dedd2, miR-297a led to an approximately 33% repression (Fig. 4*D*). Thus, mouse-specific sequence variations in the 3′ UTRs of these genes gave rise to functional target sites. We additionally tested Caspase 2, a recently described miR-295 target with partially conserved target sites (28) and observed an 80% repression in response to transfection of miR-467a (Fig. 4*D*), indicating that these miRNAs can exert strong as well as moderate repression of targets.

Discussion

Here we present a previously undescribed method of studying species-specific changes in miRNA-target relationships and apply it to uncover the genome-wide response to miRNA acquisition.

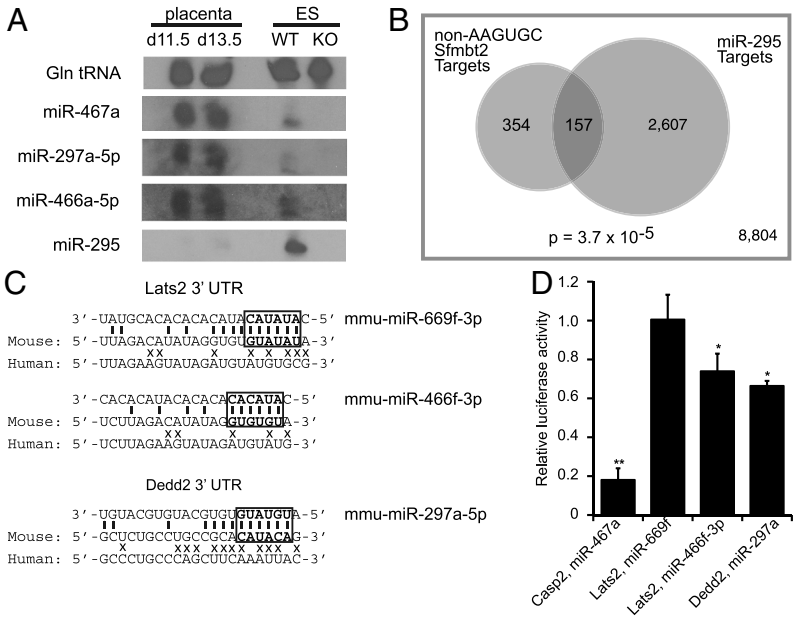

Fig. 4. Expression of *Sfmbt2* miRNAs in ES cells and the placenta, and validation of two targets of the *Sfmbt2* miRNAs. (*A*) Northern blot analysis of miR-467a, miR-297a-5p, and miR-466a-5p from the *Sfmbt2* cluster as well as miR-295 in placental and ES cells. Gln tRNA was probed as a loading control. (*B*) Intersection of predicted Sfmbt2 mouse-specific targets and miR-295 targets. *P*-value represents the significance of the intersection, and was calculated by the hypergeometric distribution. A total of 11,922 placentally expressed genes were considered for the analysis. (*C*) Comparison of mouse and human 3′ UTRs flanking gained miRNA target sites in Lats2 and Dedd2. Bases that differ are marked by X's, and the seed and seed complement are boxed. (*D*) Luciferase reporters with full length Casp2 3′ UTR, Lats2 3′ UTR, Dedd2 3′ UTR, as well as their seed mutant versions were assayed in *Dcr* KO ES cells. 20 nM miR-467a, miR-297a-5p, miR-466f-3p, and miR-669f-3p were transfected in *Dcr* KO cells. $n = 3$ and results are shown as mean ± S.E.M. *P*-values were results of Mann–Whitney tests, * denotes $p \leq 0.05$.

In contrast to many current methods of target prediction, which detect conserved relationships (37–39), our method emphasizes relative divergence to identify positive selection for site gain and loss. Our results indicate that introduction of a novel miRNA cluster can be associated with genome-wide adaptation.

The detection of these genome-wide responses was likely enabled by a number of underlying factors. First, the recency of the cluster expansion and subsequent time window of selection may have aided our ability to observe these responses. In general, evolutionary models of positive selection are likely to have a biphasic mutation profile: Early in the evolutionary window, one would expect an increase in mutation rate as favorable changes are fixed, while afterward a subsequent decrease in mutation rate should be observed as these mutations are under pressure to be maintained. The positive selection signal observed here therefore suggests that the time frame of positive selection was short enough to prevent the second conservation phase from overshadowing the early rate increase (see *SI Appendix*).

In addition to the timing of the selection, the nature of the sites examined may have affected the signals we observed using this method. In general, the statistical significance of signals described here for species-specific miRNAs is less dramatic than signals observed for conserved miRNAs. This difference is likely a consequence of the relatively greater time over which selective pressure has acted on targets of conserved miRNAs, many of which have been conserved quite extensively across mammals and more broadly in vertebrates. Analysis of human-specific miRNAs failed to show a significant enrichment for positively selected targets, which likely results from the differing quality of miRNA annotation in human. Stringent criteria have been used to define mouse miRNAs, unlike the human annotations. As a result the human set contains roughly 300 more miRNAs, some of which may be spurious and therefore reduce the statistical power of this test. Even within the mouse-specific targets, these signals showed some variation, with more significant changes being detected for 7mer-M8 seed matches as compared to 7mer-A1 sites. This result is consistent with the generally stronger activity and improved ability of 7mer-M8 sites to predict authentic targets (15, 40).

While many of the identified target interactions are likely to be novel, a subset may functionally overlap with those of the miR-290-295 cluster which promotes proliferation (26) and survival (28). In some cases, single 3′ UTRs may be targeted by multiple cluster members, as was observed for Lats2 (a target of miR-467a and miR-466f-3p), a gene known to oppose proliferation and growth (32, 34, 35). Because of the less than twofold repression observed by luciferase assay for the targets Lats2 and Dedd2, we favor the hypothesis that these genes are part of a much larger network of targets that together mediate the phenotypic consequences of these miRNAs. Importantly, however, these examples do validate the notion that a subset of relatively rapid mouse-specific mutations has resulted in functional targeting.

Inhibition of Lats2 and our other validated target, Dedd2, fits well with the position of the *Sfmbt2* cluster in a paternally expressed placental gene (20), because such genes are thought to be commonly involved in redistributing resources from mother to offspring under the parental sex conflict model (41). In this model, the father has an evolutionary incentive to promote fetal growth at the expense of maternal resources as only the fetus shares half of his genetic material. On the other hand, the mother has a mixed incentive because her reproductive success is tied to her ability to generate more offspring in the future, all of whom will share half her genome (41, 42). In support of this notion, numerous genes with progrowth properties such as Igf2 tend to be exclusively expressed by the paternal allele, while growth-inhibiting genes such as Igf2r are exclusively maternally expressed (43). Consistent with this idea, paternal duplication of proximal chromosome 2 (which includes the *Sfmbt2* gene) results in placental growth enhancement, whereas maternal disomy results in fetal and placental growth reduction (44). Our data suggest that some of these growth effects may result from contributions of the miRNA cluster in addition to the coding gene. Given the frequently paired relationships between maternally and paternally imprinted

genes, the identification of a maternally expressed gene that counteracts the role of this cluster would not be surprising.

The *Sfmbt2* cluster expansion in mouse may additionally provide a useful model for other species- and lineage-specific miRNA expansions. For instance, in humans, an analogous primate-specific miRNA expansion is present on Chromosome 19 (C19MC), with a total of 46 miRNAs within a 100-kb interval being processed from the highly repetitive intron of a maternally imprinted placental noncoding RNA (45, 46). In addition to a number of novel, primate-specific seeds, this cluster includes miRNAs with the seed "AAGUGC," found in miR-371-373 (the human counterpart of the miR-290-295 cluster) (47). Indeed, recent studies have identified aberrant expression of this cluster in human cancers, where it is thought to enhance oncogenicity by promoting cell survival and growth (48, 49). These observations parallel the results presented here, suggesting that the same pro-survival functions that are advantageous to cancer cells may have spurred the emergence and fixation of these two clusters for their contributions to intersex conflict.

Materials and Methods

Oligos and siRNAs used in all the experiments can be found in *SI Methods* in *SI Appendix*. Microarray data was obtained from and processed according to Zheng et al. (28). Human, dog, and horse mature miRNA sequences were obtained from miRBase Release 15 (50). Mouse mature miRNA sequences were obtained from Chiang et al. (19). Aligned human, mouse, and dog 3′ UTRs were obtained from TargetScan 5.1 (15, 16, 37). For detailed methods, see *SI Methods* in *SI Appendix*.

ACKNOWLEDGMENTS. We thank Joel Neilson for reading of the manuscript and the Sharp and Burge labs for helpful advice and discussions. A.R. was funded by a Fannie and John Hertz Foundation Fellowship. This work was supported by National Institutes of Health grants R01-GM34277, National Cancer Institute (NCI) Grant P01-CA42063, and the NCI Cancer Center Support (core) Grant P30-CA14051. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

1. King MC, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. *Science* 188:107–116.
2. Levine M, Tjian R (2003) Transcription regulation and animal diversity. *Nature* 424:147–151.
3. Bradley RK, et al. (2010) Binding site turnover produces pervasive quantitative changes in transcription factor binding between closely related Drosophila species. *PLoS Biol* 8: e1000343.
4. Chen K, Rajewsky N (2007) The evolution of gene regulation by transcription factors and microRNAs. *Nat Rev Genet* 8:93–103.
5. Mattick JS, Makunin IV (2006) Non-coding RNA. *Hum Mol Genet* 15:R17–29 Spec No 1.
6. Niwa R, Slack FJ (2007) The evolution of animal microRNA function. *Curr Opin Genet Dev* 17:145–150.
7. Nozawa M, Miura S, Nei M (2010) Origins and evolution of microRNA genes in Drosophila species. *Genome Biol Evol* 2:180–189.
8. Bartel DP (2009) MicroRNAs: Target recognition and regulatory functions. *Cell* 136:215–233.
9. Rosa A, Spagnoli FM, Brivanlou AH (2009) The miR-430/427/302 family controls mesendodermal fate specification via species-specific target selection. *Dev Cell* 16:517–527.
10. Clop A, et al. (2006) A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nat Genet* 38:813–818.
11. Abelson JF, et al. (2005) Sequence variants in SLITRK1 are associated with Tourette's syndrome. *Science* 310:317–320.
12. Chen K, Rajewsky N (2006) Natural selection on human microRNA binding sites inferred from SNP data. *Nat Genet* 38:1452–1456.
13. Klymenko T, et al. (2006) A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. *Genes Dev* 20:1110–1122.
14. Shin C, et al. (2010) Expanding the microRNA targeting code: functional sites with centered pairing. *Mol Cell* 38:789–802.
15. Grimson A, et al. (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell* 27:91–105.
16. Friedman RC, Farh KK, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 19:92–105.
17. Seitz H, et al. (2004) A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. *Genome Res* 14:1741–1748.
18. Glazov EA, McWilliam S, Barris WC, Dalrymple BP (2008) Origin, evolution, and biological role of miRNA cluster in DLK-DIO3 genomic region in placental mammals. *Mol Biol Evol* 25:939–948.
19. Chiang HR, et al. (2010) Mammalian microRNAs: Experimental evaluation of novel and previously annotated genes. *Genes Dev* 24:992–1009.
20. Kuzmin A, et al. (2008) The PcG gene Sfmbt2 is paternally expressed in extraembryonic tissues. *Gene Expr Patterns* 8:107–116.
21. Neilson JR, Zheng GX, Burge CB, Sharp PA (2007) Dynamic regulation of miRNA expression in ordered stages of cellular development. *Genes Dev* 21:578–589.
22. Calabrese JM, Seila AC, Yeo GW, Sharp PA (2007) RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. *Proc Natl Acad Sci USA* 104:18097–18102.
23. Leung AK, et al. (2011) Genome-wide identification of Ago2 binding sites from mouse embryonic stem cells with and without mature microRNAs. *Nat Struct Mol Biol* 18:237–244.
24. Ciaudo C, et al. (2009) Highly dynamic and sex-specific expression of microRNAs during early ES cell differentiation. *PLoS Genet* 5:e1000620.
25. Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R (2008) Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* 22:2773–2785.
26. Wang Y, et al. (2008) Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. *Nat Genet* 40:1478–1483.
27. Qi J, et al. (2009) microRNAs regulate human embryonic stem cell division. *Cell Cycle* 8:3729–3741.
28. Zheng GX, et al. (2011) A latent pro-survival function for the mir-290-295 cluster in mouse embryonic stem cells. *PLoS Genet* 7:e1002054.
29. Landgraf P, et al. (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* 129:1401–1414.
30. Lattin J, et al. (2007) G-protein-coupled receptor expression, function, and signaling in macrophages. *J Leukoc Biol* 82:16–32.
31. Baskerville S, Bartel DP (2005) Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* 11:241–247.
32. Alcivar A, Hu S, Tang J, Yang X (2003) DEDD and DEDD2 associate with caspase-8/10 and signal cell death. *Oncogene* 22:291–297.
33. Voorhoeve PM, et al. (2006) A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. *Cell* 124:1169–1181.
34. Li Y, et al. (2003) Lats2, a putative tumor suppressor, inhibits G1/S transition. *Oncogene* 22:4398–4405.
35. Ke H, et al. (2004) Putative tumor suppressor Lats2 induces apoptosis through down-regulation of Bcl-2 and Bcl-x(L). *Exp Cell Res* 298:329–338.
36. Aylon Y, et al. (2006) A positive feedback loop between the p53 and Lats2 tumor suppressors prevents tetraploidization. *Genes Dev* 20:2687–2700.
37. Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120:15–20.
38. Betel D, Wilson M, Gabow A, Marks DS, Sander C (2008) The microRNA .org resource: Targets and expression. *Nucleic Acids Res* 36:D149–153 (Database issue).
39. Krek A, et al. (2005) Combinatorial microRNA target predictions. *Nat Genet* 37:495–500.
40. Nielsen CB, et al. (2007) Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA* 13:1894–1910.
41. Moore T, Haig D (1991) Genomic imprinting in mammalian development: A parental tug-of-war. *Trends Genet* 7:45–49.
42. Haig D, Westoby M (1989) Parent-specific gene-expression and the triploid endosperm. *Am Nat* 134:147–155.
43. Renfree MB, Hore TA, Shaw G, Graves JA, Pask AJ (2009) Evolution of genomic imprinting: Insights from marsupials and monotremes. *Annu Rev Genomics Hum Genet* 10:241–262.
44. Cattanach BM, Beechey CV, Peters J (2004) Interactions between imprinting effects in the mouse. *Genetics* 168:397–413.
45. Bortolin-Cavaille ML, Dance M, Weber M, Cavaille J (2009) C19MC microRNAs are processed from introns of large Pol-II, non-protein-coding transcripts. *Nucleic Acids Res* 37:3464–3473.
46. Noguer-Dance M, et al. (2010) The primate-specific microRNA gene cluster (C19MC) is imprinted in the placenta. *Hum Mol Genet* 19:3566–3582.
47. Lin S, et al. (2010) Computational identification and characterization of primate-specific microRNAs in human genome. *Comput Biol Chem* 34:232–241.
48. Li M, et al. (2009) Frequent amplification of a chr19q13.41 microRNA polycistron in aggressive primitive neuroectodermal brain tumors. *Cancer Cell* 16:533–546.
49. Rippe V, et al. (2010) The two stem cell microRNA gene clusters C19MC and miR-371-3 are activated by specific chromosomal rearrangements in a subgroup of thyroid adenomas. *PLoS One* 5:e9485.
50. Griffiths-Jones S (2006) miRBase: The microRNA sequence database. *Methods Mol Biol* 342:129–138.