RNA BIOLOGY
2017, VOL. 14, NO. 2, 219–235
http://dx.doi.org/10.1080/15476286.2016.1270002

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

Transcriptome profiling of the developing male germ line identifies the miR-29 family as a global regulator during meiosis

Stephanie Hilz [a,b], Elizabeth A. Fogarty[a,*], Andrew J. Modzelewski[c,d,*], Paula E. Cohen[c], and Andrew Grimson[a]

[a]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, USA; [b]Department of Neurological Surgery, University of California San Francisco, San Francisco, CA, USA; [c]Department of Biomedical Sciences, Cornell University, Ithaca, NY, USA; [d]Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, CA, USA

ABSTRACT

MicroRNAs are essential for spermatogenesis. However, the stage-specific requirements for particular miRNAs in the male mammalian germ line remain largely uncharacterized. The miR-34 family is, to date, the only miRNA proven to be necessary for the production of sperm in mammals, though its germline roles are poorly understood. Here, we generate and analyze paired small RNA and mRNA profiles across different stages of germline development in male mice, focusing on time points shortly before and during meiotic prophase I. We show that in addition to miR-34, miR-29 also mediates widespread repression of mRNA targets during meiotic prophase I in the male mouse germline. Furthermore, we demonstrate that predicted miR-29 target mRNAs in meiotic cells are largely distinct from those of miR-34, indicating that miR-29 performs a regulatory function independent of miR-34. Prior to this work, no germline role has been attributed to miR-29. To begin to understand roles for miR-29 in the germ line, we identify targets of miR-29 undergoing post transcriptional downregulation during meiotic prophase I, which likely correspond to the direct targets of miR-29. Interestingly, candidate direct targets of miR-29 are enriched in transcripts encoding extracellular matrix components. Our results implicate the miR-29 family as an important regulatory factor during male meiosis.

ARTICLE HISTORY
Received 15 June 2016
Revised 10 November 2016
Accepted 1 December 2016

KEYWORDS
Fertility; germ line; miR-29; meiosis; microRNA; post transcriptional regulation; spermatogenesis

Downloaded by [New York University] at 11:53 11 October 2017

Introduction

MicroRNAs (miRNAs) are ~21 nt small noncoding RNAs that play a role in almost every known developmental process.[1] They are generated from longer, hairpin forming precursor transcripts, which are encoded at multiple loci throughout the genome.[1-5] Mature miRNAs are loaded onto members of the Argonaute protein family known as AGO proteins.[1] The resulting miRNA-loaded AGO is then guided to specific RNA transcripts via base pairing between the miRNA and the target mRNA (mRNA). In mammals and other animals, specificity is determined primarily by a 7 nt (nucleotide) region at the 5′ end of the miRNA.[6] This 7 nt region, known as the miRNA seed, base pairs with complimentary sequences in target RNAs; additional sequence and context determinants surrounding the target site determine the overall efficacy of the site.[7] Those mature miRNAs sharing the same seed sequence, referred to as miRNA families, therefore target a largely common set of RNAs.[8-11] Effective miRNA target sites are found, almost exclusively, within the 3′UTRs of mRNAs.[1] Recruitment of AGO proteins to target mRNAs results in destabilization of the target transcript, with additional regulation via translational inhibition; importantly, mRNA destabilization is the major consequential form of regulation in most mammalian cells for most target mRNAs,[12] thereby facilitating target identification by transcriptome profiling. miRNAs thus provide the cell with a mechanism for the sequence-specific post-transcriptional repression of cellular mRNAs.

Mice lacking essential miRNA biogenesis factors fail to produce spermatozoa, indicating that miRNAs are necessary for spermatogenesis.[13-20] Prior to spermatogenesis, primordial germ cells (PGCs) migrate to the gonad, where they become prespermatogonia.[21] Spermatogenesis begins in male mice shortly after birth, at about 3 days postpartum (3 days pp), when prespermatogonia have completed a period of reprogramming and transition to undifferentiated spermatogonia.[21,22] From this time onwards, subsets of spermatogonia proliferate further, differentiate, and enter meiosis,[22-24] with the first wave of spermatogonia entering meiosis at about 8-10 days pp. Upon entering meiosis, male germ cells, now called spermatocytes, undergo the process of homologous recombination. This process requires dramatic chromatin-wide induction of DNA double strand breaks leading to crossing over, and is accompanied by changes in chromatin organization and regionalized transcriptional silencing.[25-27] Following meiosis, male germ cells, now known as spermatids, begin to compact their chromatin and to further differentiate, ultimately becoming mature spermatozoa.[28] Importantly, mice with miRNA deficient germ lines exhibit defects both during meiosis, as well as in spermatid formation, suggesting that one or more miRNAs mediate gene regulation critical to the germ line at multiple distinct points.

CONTACT Andrew Grimson agrimson@cornell.edu MBG, Cornell University, Biotechnology Building, 526 Campus Road, Ithaca, NY 14853, USA.
*co-second author.

Supplemental data for this article can be accessed on the publisher's website.

© 2017 Taylor & Francis Group, LLC

Downloaded by [New York University] at 11:53 11 October 2017

To date, only one miRNA family, the miR-34 family, has been shown to be essential for spermatogenesis.[29-32] Mice deficient in miR-34 family members have defects during spermatid elongation,[29-31] possibly due to the role of miR-34 in promoting proper cilia assembly.[30] miR-34 family knockout mice also exhibit a block in spermatogenesis during the pachytene stage of meiosis, when crossing over occurs.[29] One specific germline role attributed to miR-34 derives from negative regulation of Atf1 by miR-34. *Atf1* encodes a protein implicated in germ cell apoptosis.[33] However, repression of *Atf1* was found to promote apoptosis rather than repress it[33]; it is unclear therefore why increased apoptosis is observed in the spermatocytes of miR-34-deficient mice,[29] in which *Atf1* expression is presumably de-repressed. Thus, while miRNAs are essential for the progression of spermatocytes through meiosis, and one of the miRNA families with an important role in meiosis is miR-34, it is unknown what specific meiotic processes miRNAs are crucial for, and if other miRNA families in addition to miR-34 also play consequential roles.

We set out to systematically identify miRNAs that influence gene expression in male meiosis. Our approach was to generate corresponding small RNA and mRNA high-throughput sequencing profiles during germline development, and identify differentially expressed miRNAs whose mRNA target set exhibited reciprocal changes. Using this approach, we identified 2 miRNA families with widespread regulatory roles during mouse germline development – namely the miR-34 and miR-29 families. We further show that mRNA transcripts bearing miR-34 or miR-29 family target sites that are conserved across multiple species show the strongest repression in male mouse spermatocytes. We determine that miRNA regulation by members of the miR-34 and miR-29 families are largely independent of one another, indicating that the repressive effect we discovered for miR-29 is not due to targets common to miR-34. Finally, we identify targets of the miR-29 family whose steady-state mRNA levels are controlled post-transcriptionally, which likely correspond to the direct targets of miR-29. The resulting high-confidence predicted direct targets of the miR-29 family are enriched for mRNAs encoding collagen and other extracellular matrix components, which play roles in mediating signaling in addition to cellular adhesion.[34,35] This suggests that the miR-29 family could play an important role in the germ line by regulating these processes. Our results confirm and extend the importance of the miR-34 family in meiotic cells and, more importantly, we identify the miR-29 family as a novel regulator of mammalian spermatogenesis.

Results

Premeiotic and meiotic germ cells show distinct miRNA profiles

To investigate miRNA profiles leading up to, and during, meiosis in the mammalian testis, we sequenced small RNAs (Fig. S1A) in whole testis from mice at day 1 (D1), day 3 (D3), and day 7 (D7) pp, as well as in pooled leptotene/zygotene (LZ) and pachytene (P) purified meiotic cells. This timecourse, which includes biological replicate samples, is comprised predominantly of prespermatogonia (D1),[21,22] cells transitioning to undifferentiated spermatogonia (D3),[22-24] and germ lines in which spermatogonia are differentiating and moving toward entry into the first wave of meiosis (D7),[22-24] as well as the first distinct stages of meiotic prophase I (LZ and P).[36]

The male germ line possess a dynamic landscape of different small RNA classes, from the shorter, ~21 nt miRNAs found in all germline cell types, to the ~23-29 nt pre-pachytene piRNAs found in PGCs (primordial germ cells) and early spermatogonia, and finally to the ~30 nt pachytene piRNAs found only in spermatocytes.[1,37] To confirm that small RNA profiles of each of our premeiotic and meiotic samples had the expected distribution of small RNA classes, we categorized small RNAs by length as an estimator of small RNA type. The distribution of small RNA lengths for each sample corresponds well to previous reports[38,39] (Fig. S1H). We then identified mature miRNAs as those small RNAs that mapped to miRNA precursor hairpins, grouping them into miRNA families based on a common miRNA seed sequence.[8-11] Thus, in our analysis, all members of a given miRNA family share a common and unique seed (defined in Table SI). We observed strong correlations between replicates for the hundred most abundant miRNA families (Fig. S1B-G). Principal component analysis (PCA) of miRNA family expression showed that differences between premeiotic and meiotic samples were the most pronounced (Fig. 1A), as expected. Together, our small RNA sequencing data provides a precise profile of small RNAs, including miRNAs, leading up to, and during, meiosis in the male germ line.

Highly abundant miRNAs are more likely to have an impact on mRNA target levels, in contrast to miRNAs expressed at lower levels.[40,41] We determined the abundance of each miRNA family by summing the number of sequencing reads for each individual miRNA in that family, and examined the expression patterns of the top 10 most highly abundant families leading up to, and during, meiosis (Fig. 1B and C). As testis mature from D1-D7, levels of the most abundant miRNA family, miR-10, drop relative to those of other testicular miRNAs. Interestingly, one of the other most highly expressed miRNA families during this time is also, like the miR-10 family, derived primarily from the mmu-miR-10a and b hairpins, but represents an alternative processing event from these hairpins (referred to as miR-10-iso; see Table S1). Like the dominant miRNA processed from the miR-10a and b-encoding hairpin transcripts, these isomiRs,[42] which include those with miRbase accession numbers MR0000120462 and MR0000114696[2,3], are derived also from the 5-p strand, but are cleaved at a single nucleotide downstream from the predominant cleavage site, resulting in a different miRNA seed sequence and thus a distinctly different set of targets. In LZ and P spermatocytes, levels of other miRNAs, such as miR-34, 871, and 191, increase. The levels of virtually all highly abundant miRNA families in premeiotic and meiotic samples are dynamic (Fig. 1C). Some, such as the miR-10, miR-10-iso, miR-181, miR-148, and miR-143 families decrease as germ cells near meiosis. Others, such as the miR-125, miR-15, miR-99, let-7, miR-143, miR-26, miR-127, miR-22, and miR-30 families increase leading up to meiosis, then become less abundant in spermatocytes. The miR-34, miR-191, miR-470, and miR-871 families all increase as cells near meiosis, reaching their peak abundance in spermatocytes. Overall, this demonstrates that a number of highly abundant miRNA families in the male germ line, many of which might impact target mRNA abundance, showed dynamic expression patterns during the premeiotic to meiotic transition.

Downloaded by [New York University] at 11:53 11 October 2017

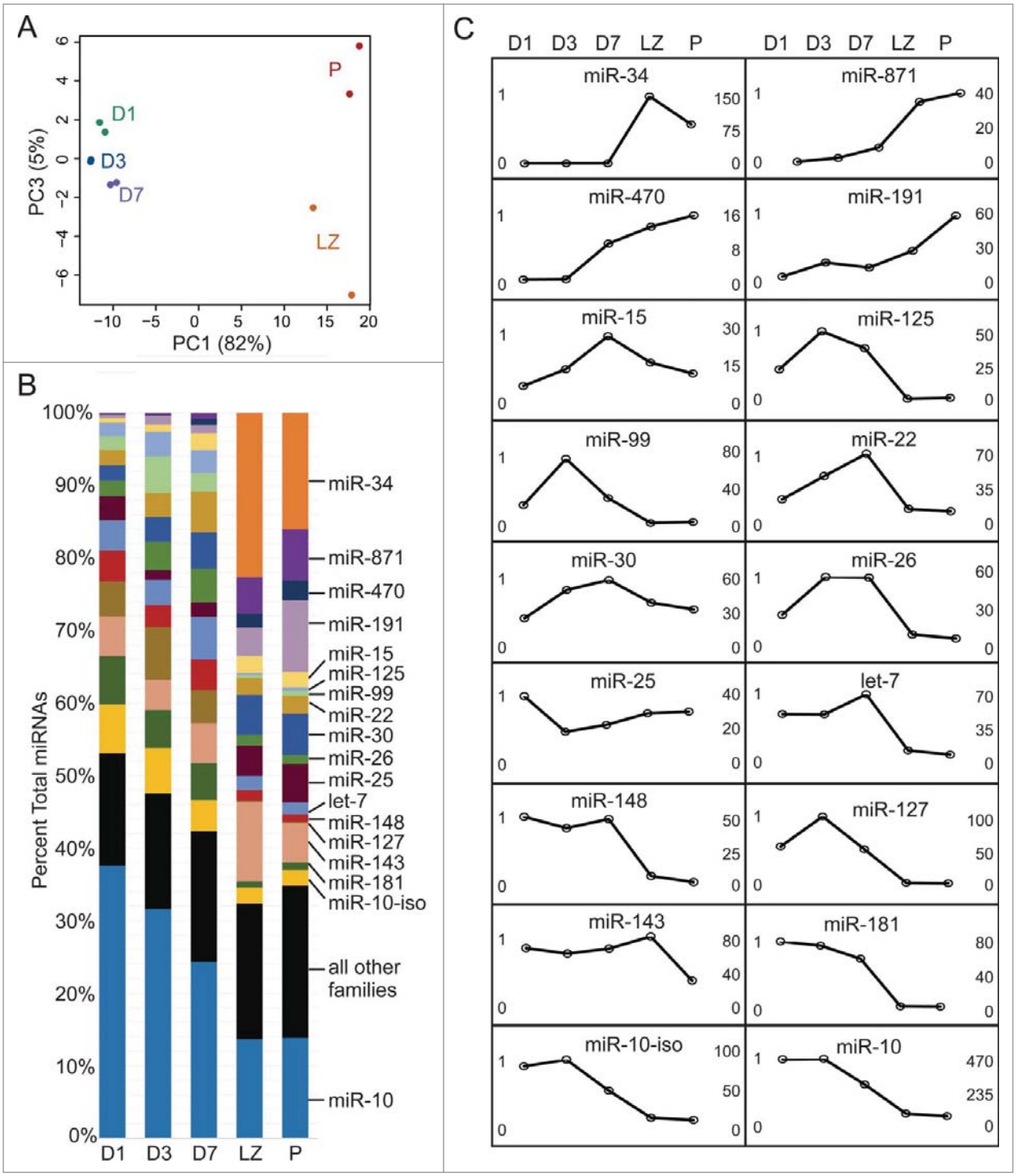


Figure 1. MiRNA abundance is dynamically regulated in spermatocytes. Small RNA sequencing was performed to characterize the abundance of miRNAs in pre and post meiotic samples. (A) Principle component analysis of miRNA family abundance (CPMs) for D1 (green), D3 (blue), D7 (purple), LZ (orange), and P (red) samples. Principal component 2 (PC2) and principal component 3 (PC3) explained comparable proportions of the total variance. PC3 shows separation between LZ and P samples, while PC2 likely reveals batch effects for the LZ and P replicates (Fig. S1). (B) Percent of total miRNA mapping sequencing reads for all miRNA families present within the top 10 most abundant miRNAs in at least one sample. The percent accounted for by all other miRNA families is shown in black. (C) Patterns of abundance across all samples for all miRNA families within the top 10 most abundant miRNA families in at least one sample. Left hand y axis denotes proportion of maximum abundance. Right hand y axis denotes CPM values (1 = 1,000 CPM).

Genes involved in spermatogenesis and cilia formation are upregulated during meiosis, while those involved in cell signaling are downregulated

We next wanted to assess overall changes in mRNA levels between premeiotic and meiotic samples. We used RNAseq to profile mRNAs using the same samples used for small RNA profiling. Reminiscent of our small RNA data, we observed large differences between premeiotic and meiotic samples using PCA (Fig. 2A). To examine changes in expression between samples at different stages of germline development, we compared transcriptome profiles from each successive pair of samples across our time course, as well as D1 to D7 in order to assess changes in premeiotic

Downloaded by [New York University] at 11:53 11 October 2017

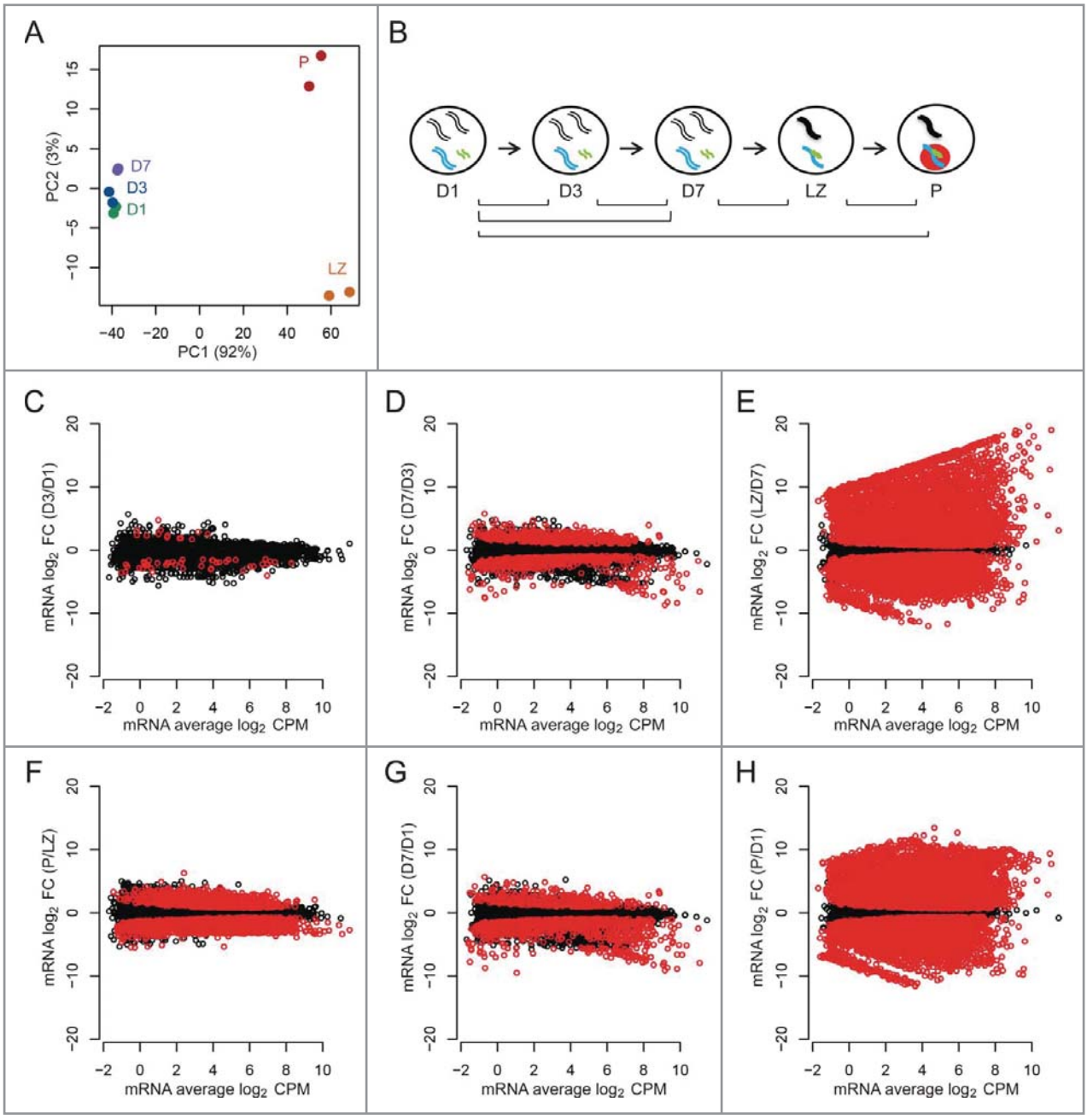


Figure 2. Gene expression patterns are altered between premeiotic and meiotic samples. (A) Principle component analysis of RNAseq transcript abundance (CPMs) for D1 (green), D3 (blue), D7 (purple), LZ (orange), and P (red) samples. (B) The $\log_2$ fold change in abundance of each mRNA transcript was calculated for each of the 6 pairwise comparisons shown. (C-H) Differential abundance of mRNA transcripts was determined in premeiotic testis and early spermatocytes. Transcripts were compared in D1 & D3 (C), D3 & D7 (D), D7 & LZ (E), LZ & Pach (F), D1 & D7 (G), and D1 & P (H) samples. Plots show the average $\log_2$ abundance (as CPM; x axis) for each mRNA transcript and the transcript's $\log_2$ fold change (FC) between the 2 compared samples (y axis). Transcripts that are not differentially abundant at a multiple comparison adjusted p value of $p < 0.05$ are black; those that are differentially abundant are red.

cells leading up to meiosis, and D1 to P to assess changes between early premeiotic and meiotic cells, leading to a total of 6 comparisons (Fig 2B). Data for each of the remaining comparisons (D1 to LZ, D3 to LZ, D3 to P, and D7 to P; Fig. S2A) are provided as Supplemental Figures, and were similar to the 2 premeiotic to meiotic comparisons we choose to focus on (D7 to LZ and D1 to P).

As observed for the small RNA data, we found many more mRNAs with differential expression when compared between meiotic and premeiotic samples than between different premeiotic samples or between different meiotic samples (Fig. 2C-H and Fig. S2B-E). We performed gene ontology analysis in order to identify biological processes enriched among upregulated or downregulated transcripts. For the premeiotic comparisons D3 to D7 and D1 to D7, we observed an upregulation in genes involved in meiosis, DNA methylation, and piRNA metabolism (Table 1); no biological process categories were significantly enriched for the D1 to D3 comparison. These results likely reflect the shift from D1/D3 to D7, when cells prepare for, and enter, meiosis.

Table 1. For each of the 6 comparisons (D1 to D3, D3 to D7, D7 to LZ, LZ to P, D1 to D7, and D1 to P), the top 5 enriched biological processes identified using gene ontology analysis are shown for genes significantly upregulated or downregulated between samples.

Upregulated			Downregulated	
D1 to D3	bone regeneration	1.0E+00	response to carbon dioxide	1.0E+00
	regulation of neuron death	1.0E+00	regulation of type IV hypersensitivity	1.0E+00
	negative regulation of chondrocyte proliferation	1.0E+00	negative regulation of type IV hypersensitivity	1.0E+00
	negative regulation of sensory perception of pain	1.0E+00	lipoprotein transport	1.0E+00
	negative regulation of SMAD protein import into nucleus	1.0E+00	lipoprotein localization	1.0E+00
D3 to D7	piRNA metabolic process	1.7E-16	single-multicellular organism process	2.4E-20
	spermatogenesis	6.1E-15	multicellular organismal process	4.1E-20
	male gamete generation	6.1E-15	multicellular organism development	6.4E-20
	DNA methylation involved in gamete generation	9.6E-15	single-organism developmental process	1.3E-18
	gamete generation	2.1E-13	developmental process	1.6E-18
D7 to LZ	spermatogenesis	2.6E-28	cellular process	7.6E-32
	male gamete generation	3.1E-28	regulation of multicellular organismal process	4.8E-29
	cilium organization	4.1E-24	single-organism process	3.7E-27
	cilium morphogenesis	4.1E-24	biological regulation	4.1E-26
	sexual reproduction	5.7E-23	regulation of biological process	3.6E-25
LZ to P	nucleic acid metabolic process	7.8E-11	spermatogenesis	2.1E-05
	RNA processing	9.7E-10	male gamete generation	2.1E-05
	nucleobase-containing compound metabolic process	1.4E-08	sexual reproduction	2.2E-04
	DNA metabolic process	1.4E-08	gamete generation	7.5E-04
	nitrogen compound metabolic process	1.9E-08	multi-organism reproductive process	1.2E-03
D1 to D7	meiotic cell cycle process	1.2E-13	multicellular organismal process	1.5E-21
	meiotic nuclear division	1.2E-13	single-multicellular organism process	1.4E-17
	meiotic cell cycle	3.4E-13	cell differentiation	3.2E-16
	piRNA metabolic process	3.8E-10	single-organism developmental process	5.7E-16
	DNA methylation involved in gamete generation	5.5E-09	developmental process	6.4E-16
D1 to P	spermatogenesis	2.1E-35	regulation of multicellular organismal process	1.0E-58
	male gamete generation	2.1E-35	signaling	2.7E-58
	cilium morphogenesis	8.5E-27	single organism signaling	5.7E-58
	cilium organization	8.5E-27	cell communication	2.2E-57
	sexual reproduction	1.7E-26	signal transduction	7.1E-56

Downloaded by [New York University] at 11:53 11 October 2017

Genes upregulated in the premeiotic to meiotic comparisons were enriched in functional categories such as spermatogenesis and cilium morphogenesis, reflecting expression in LZ and P spermatocytes of genes necessary for sperm flagellar assembly (Table 1, SII). Downregulated genes for all comparisons, except D1 to D3, in which no categories were significant, were enriched for those involved in multicellular processes and signaling (Table 1), suggesting perhaps that as germ cells transition through spermatogenesis, they communicate less with the cells around them. To identify subgroups of genes with coordinate expression patterns between premeiotic and meiotic samples, we performed *k*-means clustering on the expression of genes differentially expressed between our premeiotic and meiotic samples. This method results in the grouping of genes into clusters with similar expression patterns.[43] We performed gene ontology analysis on each of these clusters, choosing a value of *k* that minimized the number of clusters with highly similar ontologies (Fig. S2F). In addition to confirming our previous observations, namely that cell signaling is downregulated in meiotic samples as compared to premeiotic samples (Table 1, SII), gene ontology analyses of co-regulated sets of genes also revealed a modest upregulation of genes enriched for protein catabolism as cells progress into meiosis, as well as an increase in the expression of genes involved in chemical signaling (Fig. S2F, Table SIII). Together, our analysis reveals both the up- and downregulation of genes belonging to a variety of pathways as male germ cells enter meiosis, with the most pronounced alterations to the transcriptome coincident with entry into meiosis.

Increased expression of both the miR-34 and miR-29 families in spermatocytes results in global downregulation of their mRNA targets

To determine if changes in miRNA expression alter RNA transcript abundance in germ cells entering meiosis, we next investigated the relationship between the expression of miRNA families and their predicted mRNA targets. We chose to focus our analysis on miRNA families with seed sequences that are conserved broadly across vertebrates[11] (referred to as conserved miRNA families hereafter). For our analysis, we integrated changes in miRNA family profiles with changes in their corresponding mRNA target abundance, using TargetScan[8-11] to identify targets based on their predicted response to a given miRNA family (TargetScan context+ score, where we considered only targets with a score < -0.1). The goal of these analyses was to identify miRNA families with the most pronounced regulatory impacts on the transcriptome, and thus likely playing consequential roles during spermatogenesis.

We first wanted to determine if mRNAs that are targets of conserved miRNA families preferentially underwent expression changes during germline development. This determination was necessary because, as a group, mRNAs targeted by miRNAs are distinct from the overall transcriptome; for example, they tend to be depleted in housekeeping genes.[6,44] We began by partitioning mRNAs according to the number of target sites they possessed that corresponded to conserved miRNAs, and compared their expression between samples. We observed that mRNAs that are targets of conserved miRNA families tended to be present at lower levels in meiotic as compared to

Downloaded by [New York University] at 11:53 11 October 2017

premeiotic samples than mRNA transcripts not targeted by these miRNAs (Fig. 3C and F, Fig. S3A-D, all $p = 2.2 \times 10^{-16}$). Furthermore, the more target sites within a given transcript, the more pronounced this tendency became (Fig. 3C and F, Fig. S3A-D). We observed this same phenomenon, though to a much lesser degree, when comparing premeiotic D1 to D3 (Fig. 3A) and D1 to D7 (Fig. 3E). For the meiotic LZ to P comparison, we observed the opposite trend – the targets of conserved miRNA families tended to be increased in P cells as compared to transcripts not targeted by conserved miRNAs (Fig. 3D, $p = 2.2 \times 10^{-16}$), although this difference was smaller than that observed comparing premeiotic to meiotic samples. The same trend, that is a slight overall de-repression of miRNA targets, was also observed for D3 to D7 comparisons, but was even more subtle (Fig. 3B; $p = 1.9 \times 10^{-2}$). While our original assumption was that the overall effects of miRNAs on mRNA transcripts would be neutral, with some miRNA families increasing between samples and thus leading to more repression of their target transcripts while others decreased and repression of their targets was relieved, these data (Fig. 3A-F) argue that in addition to the specific effects of each miRNA family on transcript abundance, there is an overall net miRNA effect, which we needed to consider and model in order to identify specific miRNA families impacting the transcriptome. This is most evident when comparing premeiotic to meiotic populations, where it appears that there is a net repressive effect by miRNAs on mRNA transcripts. However, it is worth noting that these correlations may, in fact, not represent direct effects of miRNA regulation, instead, they could be due to preferential regulation by other mechanisms that simply correlate with the number of miRNA target sites. To begin to unravel whether the increased repression of genes with a greater number of miRNA target sites might be due to regulation by miRNAs, we compared the regression coefficients for repression versus the number of miRNA target sites for targets of 3 different categories of conserved miRNA families (Fig. 3G-L, Fig. S3E-H): 1) those expressed (CPM > 1; red), 2) those highly expressed (CPM > 250; purple), and 3) those both highly expressed and differentially abundant (blue). We found that for premeiotic to meiotic comparisons (D7 to LZ and D1 to P), the regression coefficient was more negative for highly expressed miRNA families as compared to all expressed miRNA families, and even more negative for highly expressed miRNA families that were significantly altered in their abundance (Fig. 3I and L). This trend was not observed for the remaining comparisons (Fig. 3G-H and J-K). As the net downregulation of miRNA targets in meiotic cells increases for targets of highly expressed and differentially abundant miRNAs, this suggests miRNAs are responsible, at least in part, for this net repression, and could indicate an overall increase in the regulatory impact of many miRNAs as germline cells proceed into meiosis.

We next set out to identify specific miRNA families that exhibited a detectable widespread impact on the abundance of their predicted target mRNAs. To differentiate between miRNAs that are significantly altered between samples from those that are not, we determined whether each miRNA family was differentially abundant for each comparison (Fig. 4A-F, Fig. S4A-D). For each miRNA family, we then compared the fold change for all targets of that family to a background distribution, which we generated by randomly sampling a subset of mRNAs that were also miRNA targets, but not targets of that particular family (Fig. 4G-L, Fig. S4E-H). This strategy allowed us to identify significantly altered miRNA families that were having a detectable impact on gene expression in male germ cells, while controlling for the net effects of miRNA targeting described above (Fig 3A-L). For the D1 to D3 (Fig. 4G; Table SIV), D3 to D7 (Fig. 4H; Table SV), and LZ to P (Fig. 4J; Table SVII) comparisons, no significantly altered miRNA families showed a significant impact on mRNA target abundance. For the D1 to D7 comparison, the miR-30 family, which increases in abundance from D1 to D7, shows a slight but significant impact on mRNA target abundance (Fig. 4K; Table SVIII). However, this impact is in the opposite direction from what would be predicted for repressive miRNA-mediated regulation, as target fold change values are greater, rather than less, than the background distribution, indicating that targets of the miR-30 family are more likely to increase in their abundance than all targets of highly conserved miRNAs. Thus, we did not identify miRNAs with a significant impact on gene expression in premeiotic cells.

For the D1 to P comparison, the miR-34 family, which increases dramatically in abundance from D1 to P (~280 fold increase; Fig. 5G), exhibited a significant signature for repression of miR-34 family targets, as compared to background (Fig. 4L; Table SIX). In this case, the change in target abundance is concordant with the large increase in miR-34 family levels, likely indicating the widespread impact of the miR-34 family on the expression of its target mRNAs as cells enter meiosis. Significant repression of miR-34 family targets as compared to background in D1 v P was robust to different expression thresholds cutoffs for mRNA transcript expression (Fig. S4I-P), and was also observed in D1 to LZ, D3 to LZ, and D3 to P (Fig. S4E-G). These observations were expected given that an increase in miR-34 family levels upon meiotic entry has been previously described,[33,45,46] and because miR-34 is the only miRNA family known to be essential for spermatogenesis.[29-31]

In addition to miR-34, the targets of a sole additional miRNA family, the miR-29 family, also showed a significant signature of target repression compared to background in meiotic samples, which we observed for all premeiotic to meiotic comparisons: D7 to LZ (Fig. 4I; Table SVI), D1 to P (Fig. 4L; Table SIX), D1 to LZ (Fig. S4E), D3 to LZ (Fig. S4F), D3 to P (Fig. S4G), and D7 to P (Fig. S4H). The miR-29 family has been shown to play important roles in DNA damage repair and methylation in non-germline tissues,[47,48] together with diverse roles in other tissues[49-53]; however, no role for the miR-29 family in germline development has been reported. The miR-29 family increases in abundance from D1 to P (~21 fold increase as determined by small RNAseq; Fig. 5G). We validated this observed increase in miR-29 family levels using qPCR (Fig. S5D). The repression of mRNA transcripts from premeiotic populations to LZ cells, in which the miR-29 family is maximally abundant, is slightly greater for miR-29 family targets than for targets of the miR-34 family, though this result was not statistically significant (median fold changes for miR-34 and miR-29, respectively, are D1 to LZ: -0.8, -0.9, $p = 0.2$; D3 to LZ: -0.8, -0.9, $p = 0.2$; and D7 to LZ: -0.8, -1.0, $p = 0.2$).

Downloaded by [New York University] at 11:53 11 October 2017

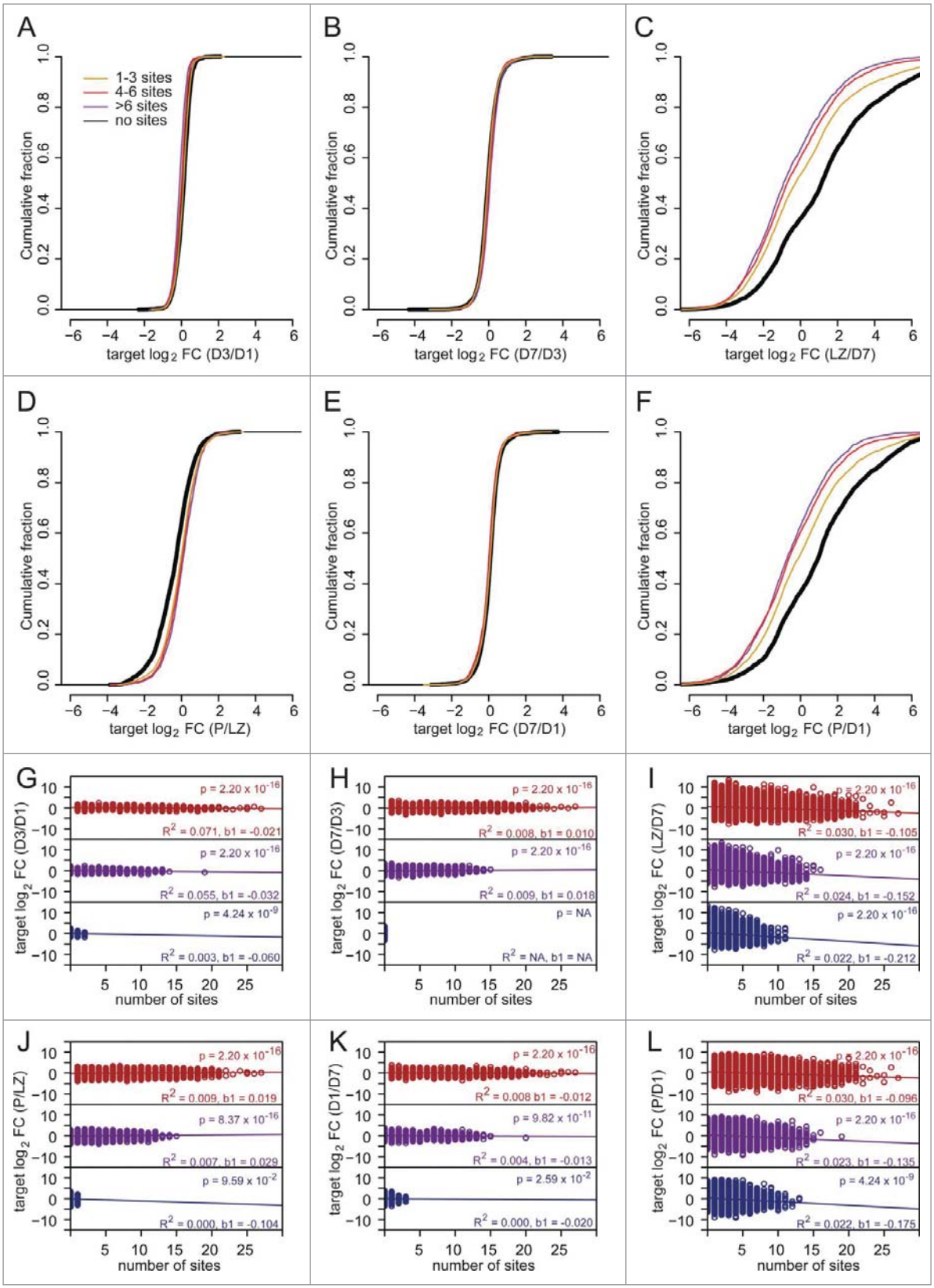


Figure 3. RNAseq reveals downregulation of transcripts targeted by highly conserved miRNAs in meiotic cells. (A-F) Cumulative distributions of mRNA transcript $\log_2$ fold-change for each comparison, for mRNAs not targeted by highly conserved miRNA families (black), mRNAs with 1-3 target sites (yellow), 4-6 target sites (red), and greater than 6 target sites (purple). The mRNA fold change distributions for mRNA that are targets (all colored lines) and are not (black line) were found to be significantly different ($p < 0.05$) for all comparisons, with $p = 2.2\times10^{-16}$ for D1 & D3 (A), $p = 1.9\times10^{-2}$ for D3 & D7 (B), $p = 2.2\times10^{-16}$ for D7 & LZ, (C) $p = 2.2\times10^{-16}$ for LZ & P, (D) $p = 3.0\times10^{-5}$ for D1 & D7 (E), and $p = 2.2\times10^{-16}$ for D1 & P (F). (G-L) The mRNA $\log_2$ fold-change (y axis) compared to the number of target sites (x axis) corresponding to conserved and expressed (CPM > 10) miRNAs (red), conserved and highly expressed miRNAs (CPM > 250) (purple), and conserved, highly expressed, and significantly differentially abundant miRNAs (blue) for D1 & D3 (G), D3 & D7 (H), D7 & LZ (I), LZ and P (J), D1 & D7 (K), and D1 & P (L). The R^2, slope (b1), and p values for each regression are shown.

Downloaded by [New York University] at 11:53 11 October 2017

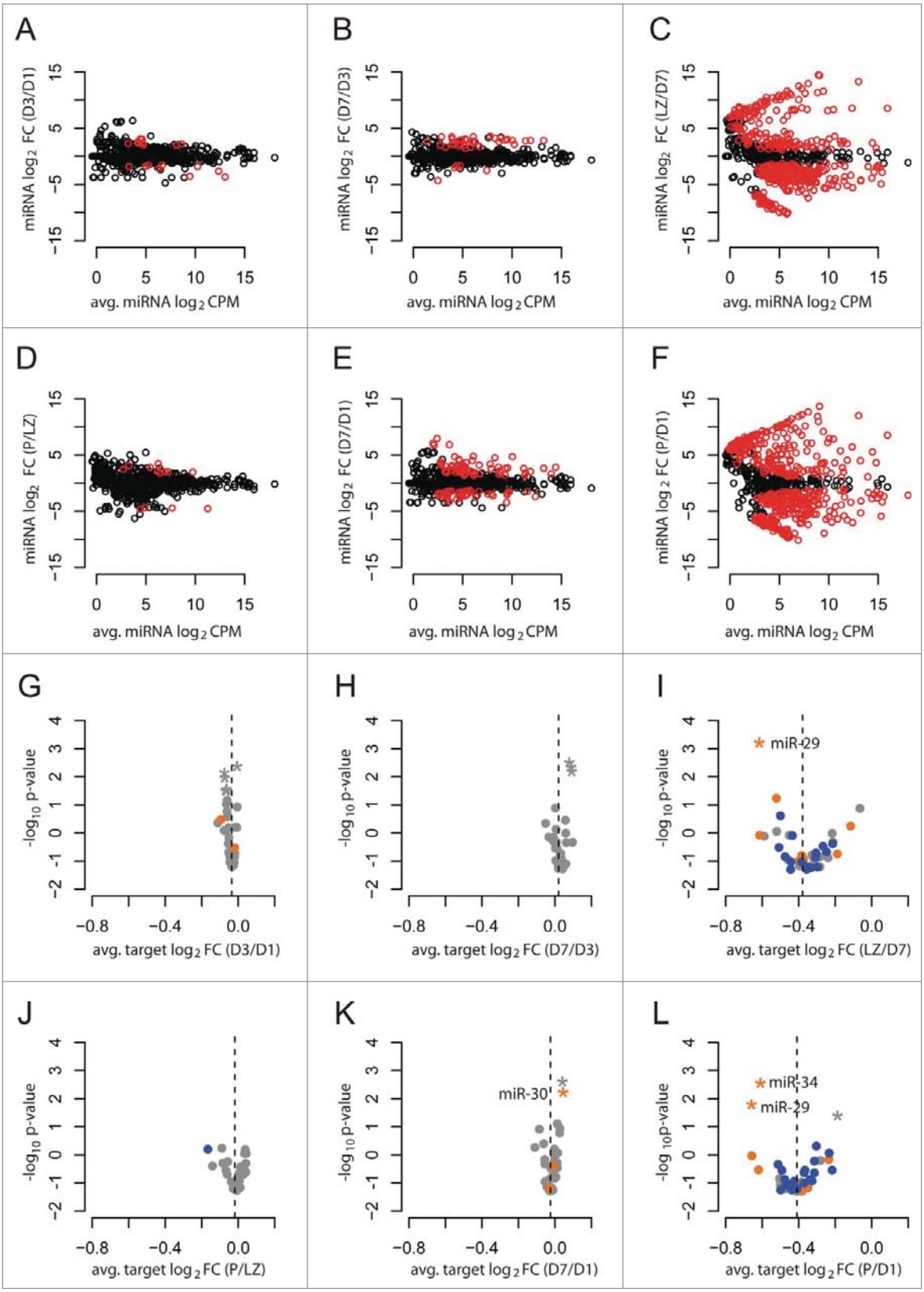


Figure 4. miRNA targeting signature analysis reveals significant downregulation of miR-34 and miR-29 targets in early meiosis. (A-F) Differential abundance of miRNA families was determined in premeiotic testis and early spermatocytes. miRNA families were compared in D1 & D3 (A), D3 & D7 (B), D7 & LZ (C), LZ & P (D), D1 & D7 (E), and D1 & P (F) samples. Plots show the average $\log_2$ abundance (as CPM; x axis) for each miRNA family and the family's $\log_2$ fold change (FC) between the 2 compared samples (y axis). miRNA families that are not differentially abundant at a multiple comparison adjusted p value of $p < 0.05$ are black; those that are differentially abundant are red. (G-L) Targeting signature analysis for the 6 comparisons made in A-F. A Wilcoxon rank test was used to determine significant shifts in fold change distributions of predicted mRNA targets of highly expressed (CPM $\geq$ 250) miRNA families as compared to a background fold-change distribution for targets of all other miRNA families (see Methods). The Bonferroni corrected $\log_{10}$ p value for the Wilcoxon rank test is shown for each miRNA family (y axis), as well as the average $\log_2$ fold change of predicted targets for that miRNA family (x axis), for comparisons between D1 & D3 (G), D3 & D7 (H), D7 & LZ (I), LZ & P (J), D1 & D7 (K), and D1 & P (L) samples. Targets of miRNA families that did not change significantly in their abundance between compared samples are gray. Targets of miRNA families that are significantly more or less abundant are, respectively, orange and blue. miRNA families with no significant targeting signature are represented by dots; those with a signature appear as an asterisk.

Downloaded by [New York University] at 11:53 11 October 2017

Mature miRNAs from the miR-29 family derive from 4 miRNA genes – Mir29a, Mir29b-1, Mir29b-2, and Mir29c.[2,3] Mir29a was the most highly expressed of the miR-29 encoding loci in meiotic cells, with its dominant product, the mature miRNA mmu-mir-29a, accounting for the majority of the increased levels of miR-29 family observed in LZ and P spermatocytes (Fig. S5A; Table SX). To ensure that the targeting signature we observed for the miR-29 family was not specific to the method of target prediction we used (TargetScan), we determined whether the signature was robust when assessed using 2 alternative target prediction methods, PicTar[54] and miRanda,[55,56] both of which predict targets for individual mature miRNAs. We found that expressed targets (CPM > = 1) predicted for mmu-mir-29a were significantly repressed compared to all other expressed mRNAs in D1 to LZ (Fig. S5B-C), when assessed using either PicTar or miRanda. Based on the results of our miRNA targeting analysis, we conclude that targets of the miR-34 and miR-29 families exhibit greater repression than the pooled targets of all other highly conserved miRNA families in meiotic cells, and that this repression is concordant with the upregulation of both the miR-34 and miR-29 families during meiosis.

To investigate further the apparent repression mediated by the miR-34 and miR-29 families in meiotic cells, we asked whether transcripts predicted to be more effectively targeted by either the miR-34 or miR-29 family exhibited stronger repression. The TargetScan context+ score predicts how effectively a given miRNA family targets a particular mRNA based on the nature of the base pairing between the miRNA and target transcript, the target site's contextual environment, and competition with other targets of that same miRNA.[9,10] We found that the median repression from D1 to P for transcripts targeted by miR-29 correlated with the strength of TargetScan context+ scores (Fig. 5B; $p = 8.8\times10^{-3}$), indicating increasingly potent miR-29 mediated repression of targets predicted to be more effectively regulated by the miR-29 family. The repression of transcripts targeted by the miR-34 family were not significantly correlated with the strength of predicted targeting (Fig. 5A; $p = 3.5 \times 10^{-1}$). This difference could be explained by the disparity in miR-29 and miR-34 family abundance in meiotic cells. In the case of the highly abundant miR-34 family, even weak targets will have more opportunities for interaction with miR-34, thus, even if only a small fraction of these interactions are successful, repression will occur. For the less abundant miR-29 family, however, transcripts with weaker target sites have fewer opportunities to interact with miR-29, and thus are less likely to be repressed. Targets of the miR-30 family, which were found to be significantly altered from D1 to D7, albeit in the opposite direction expected, show no discernable relationship between predicted target strength and repression (Fig. 5C; $p = 3.6\times10^{-1}$). These results further indicate that the increase in abundance of miR-30 family targets was coincidental and unrelated to miR-30.

Another predictor of miRNA targeting strength is conservation. Targets with a conserved target site for a given miRNA are more likely to be effectively repressed by that miRNA, moreover, consequential regulation by miRNAs is, presumably, more common for conserved targets.[11] The correlation between observed repression of targets and degree of target site conservation was observed for both the miR-34 and miR-29 family (Fig. 5D and E; $p = 3.5 \times 10^{-3}$ and $p = 1.2 \times 10^{-9}$, respectively), and was more pronounced for miR-29 (Fig. 5D = 4.0×10^{-3}). We did not see this pattern, however, for the miR-30 family (Fig. 5F; $p = 1.5 \times 10^{-1}$). Thus, as expected, more strongly conserved targets of the miR-29 and miR-34 families showed greater repression.

miR-34 and miR-29 family regulatory networks are distinct

We wanted to examine to what degree the regulatory networks of highly expressed miRNA families in the male germ line are independent. We restricted our analysis to mRNAs with conserved predicted target sites to enrich for those most likely to play important biological roles. We calculated the proportion of conserved targets shared between miR-29, miR-34 and all highly expressed germ line miRNA families, for both premeiotic and meiotic samples (Fig. 6A). For most pairwise comparisons between different miRNAs, we found no evidence of co-regulation (Fig. 6A; Fig. S6A-D). On average, about 6% of conserved targets were shared between any given 2 miRNAs. There were, however, a few exceptions in the premeiotic samples. We found more overlap than expected between targets of the miR-26 and miR-148 families in premeiotic samples (Fig. S6A-C). For a subset of premeiotic samples, we also observed a greater than expected overlap in miR-181 and miR-27 family targets, as well as miR-26 and miR-27 family targets (Fig. S6B-C). We also determined the degree of overlap between targets of the miR-34 and miR-29 families to confirm that the repression seen for miR-29 was not due to targets common to miR-34. The miR-29 and the miR-34 families share only 18 conserved targets out of a total of 568 (379 miR-29 targets and 207 miR-34 targets; Fig. 6A, Fig. S6E). Therefore, only ~5% of the miR-29 family's conserved targets are also targets of miR-34 (Fig. 6A). Furthermore, we looked at whether these few shared targets corresponded to mRNAs predicted to be much stronger targets of the miR-34 family (Fig. S6E) and found that they were not. Overall, this result suggests that miRNA targeting networks in meiotic cells are largely distinct, including those of the miR-29 and miR-34 families.

We repeated the targeting analysis for all comparisons in which both the miR-34 and miR-29 families showed a significant targeting signature, with those targets shared between the miR-29 and the miR-34 families removed; in all cases, the targets of the miR-29 family were still significantly more repressed than those of the background distribution sampled from targets of all other miRNA families other than miR-29 (Fig. 6B, only D1 to P comparison shown). These data demonstrate that both the miR-34 and miR-29 families each have independent regulatory functions in spermatocytes, targeting largely distinct sets of mRNA targets.

Direct targets of the miR-29 family are enriched for components of the extracellular matrix

To gain insight into the roles of the mir-29 and miR-34 families in meiotic cells, we examined the conserved predicted targets of these miRNAs that were significantly downregulated in either LZ or P samples as compared to D1. We chose to focus on these

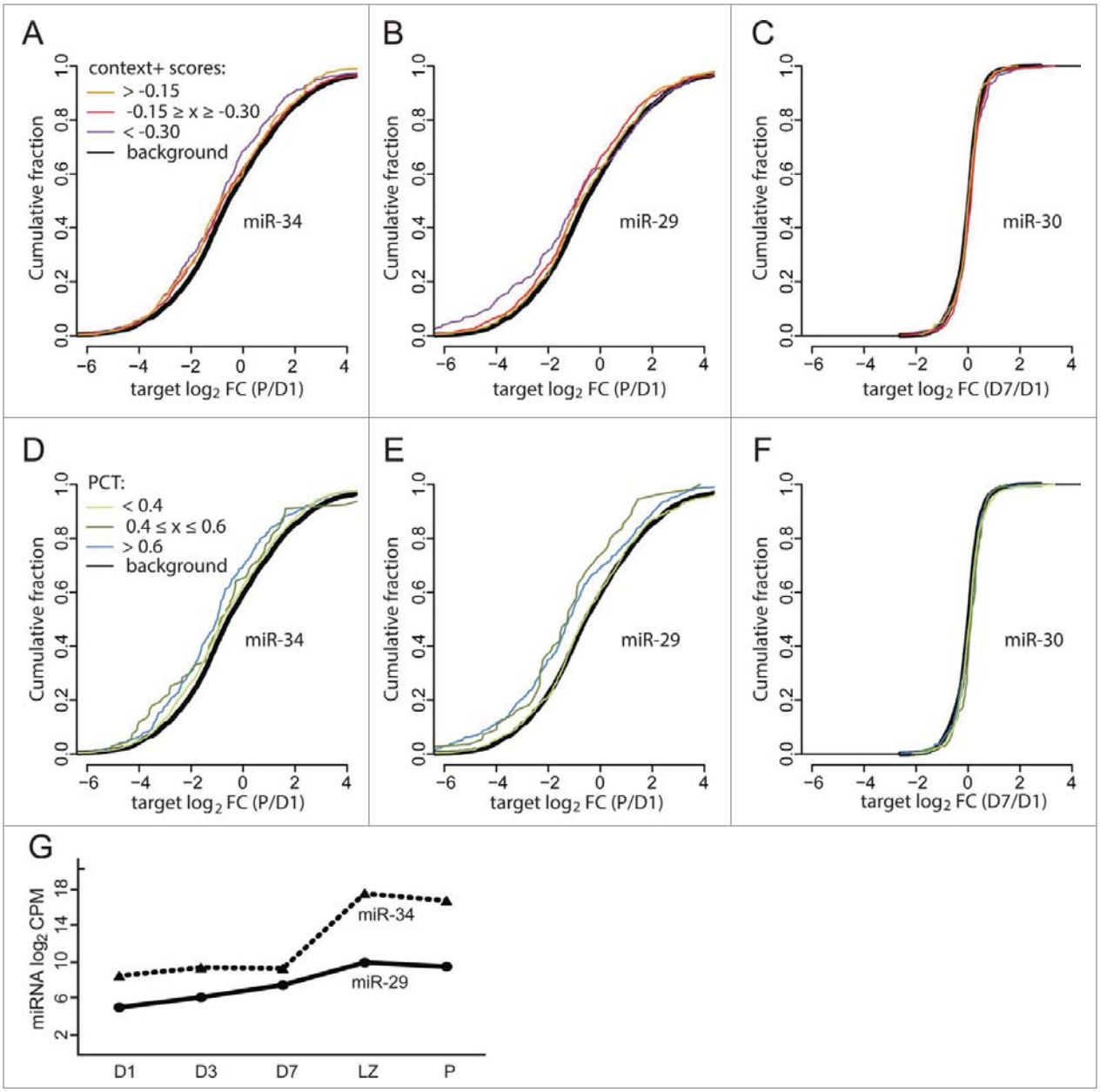


Figure 5. Conserved targets of the miR-34 and miR-29 families are strongly repressed in meiotic cells. (A-C) The cumulative distribution of transcript $\log_2$ fold changes for mRNAs predicted to be weakly (context+ > 0.15; yellow), moderately (0.15 ≥ context+ ≥ 0.30; red), and strongly (context+ < 0.3; purple) targeted by the miR-34 family from D1 to P (A), the miR-29 family from D1 to P (B), and the miR 30 family from D1 to D7 (C). The cumulative distribution of the background set is shown in black. (D-F) The cumulative distribution of transcript $\log_2$ fold-changes for mRNA transcripts predicted to be weakly (PCT < 0.4; light green), moderately (0.4 ≤ PCT ≤ 0.6; dark green), and strongly (PCT > 0.6; blue) conserved targets of the miR-34 family from D1 to P (D), the miR-29 family from D1 to P (E), and the miR 30 family from D1 to D7 (F). The cumulative distribution of the background set appears black. (G) Abundance of miR-34 (dashed line) and miR-29 (solid line) in premeiotic and meiotic samples as quantified by RNA seq.

2 comparisons (D1 to LZ and D1 to P), as both the miR-29 and miR-34 families exhibited significant targeting signatures in each. Predicted targets of both miRNA families were enriched among significantly downregulated mRNAs (miR-29 - D1 to LZ: $p = 7.4 \times 10^{-6}$, D1 to P: $p = 3.7 \times 10^{-16}$; miR-34 - D1 to LZ: $p = 1.4 \times 10^{-4}$, D1 to P: $p = 9.7 \times 10^{-7}$; chi-square test), and were depleted among significantly upregulated mRNAs (miR-29 - D1 to LZ: $p = 2.6 \times 10^{-12}$, D1 to P: $p = 5.5 \times 10^{-7}$; miR-34 - D1 to LZ: $p = 1.2 \times 10^{-4}$, D1 to P: $p = 8.3 \times 10^{-7}$; chi-square test). Gene ontology cellular component analysis showed an enrichment in genes involved in extracellular matrix organization for predicted downregulated targets of the miR-29 family, in both comparisons (Table SXI). No categories were significantly enriched for the miR-34 family. However, while some of the downregulated transcripts are direct targets of the miR-29 or miR-34 family, many are undoubtedly downstream or indirect targets that result from the direct effects of these miRNAs.

To enrich further the list of predicted direct targets of the miR-29 and miR-34 families, we utilized the EISA (exon-intron split read analysis) method,[57] which compares changes in exon to intron levels in RNAseq data in order to infer evidence of post-transcriptional regulation.[58] Specifically, for each gene in our RNAseq data, we determined individual counts for introns,

Downloaded by [New York University] at 11:53 11 October 2017

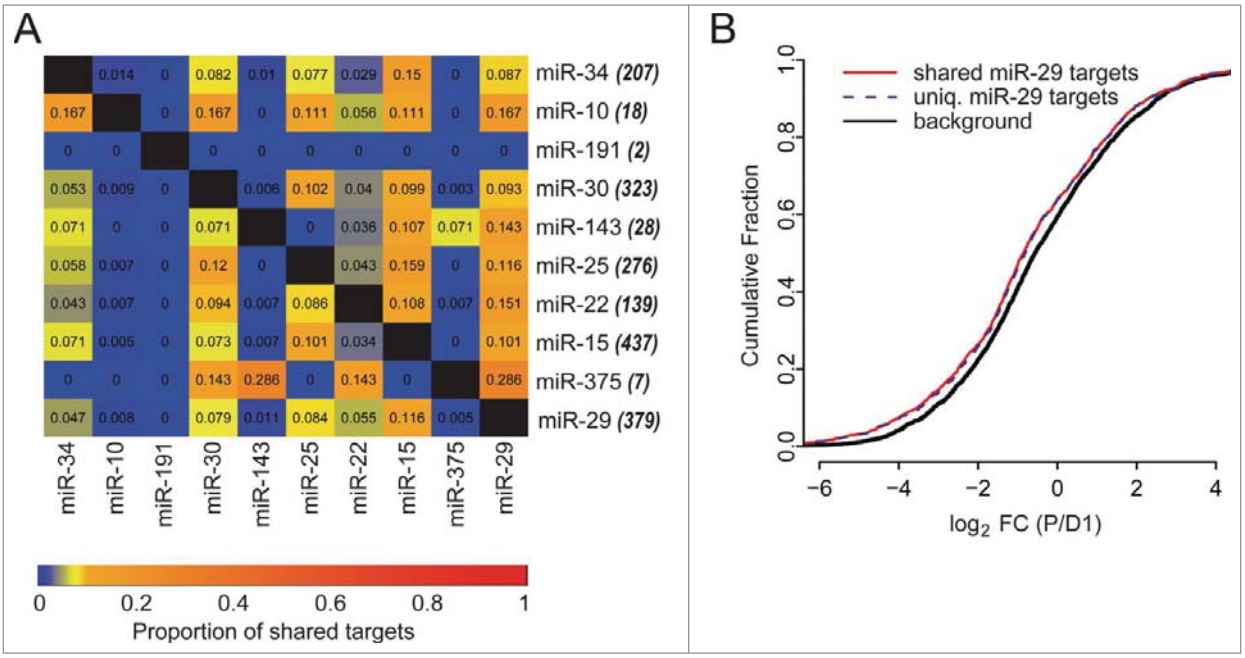


Figure 6. The miR-34 and miR-29 families regulate a distinct set of targets in spermatocytes. (A) Proportion of targets shared between miR-29, miR-34, and other highly expressed miRNAs in purified pachytene spermatocytes. Values shown denote the proportion of conserved targets (PCT > 0.5) for each miRNA (row) that are also targets of another miRNA family (column) in purified pachytene spermatocytes. Those with significantly more overlap than expected (Bonferroni corrected p value < 0.05) appear red. (B) Cumulative distribution of $\log_2$ fold-changes (D1 to P) for mRNAs predicted to be targets of the miR-29 family before (red; Wilcoxon rank $p = 7.5\times10^{-5}$) and after (dotted purple; Wilcoxon rank $p = 1.1\times10^{-3}$) removal of transcripts also targeted by the miR-34 family. The cumulative distribution of the background set appears black.

representing pre-mRNA and serving as a proxy for transcriptional regulation, and exons, representing mature mRNA and reflective of both transcriptional and post-transcriptional regulation. Transcripts with a significantly greater decrease in exon levels, when compared to changes in intron levels, were considered to have been subject to post-transcriptional downregulation via mRNA destabilization. This approach allows us to remove predicted targets of the miR-29 or miR-34 families that did not exhibit evidence of post-transcriptional downregulation, which are less likely to represent true, direct targets in vivo. Out of ~16,000 genes corresponding to mRNA transcripts expressed in premeiotic or meiotic cells, only ~34% had sufficient exonic and intronic coverage for statistical analysis in either of the 2 comparisons, consistent with previous implementations of this method.[58] Of the genes we analyzed, there was an overall correlation between intron and exon changes (Fig. 7A-D; $R^2 = 0.62$ and $p = 2.2\times10^{-16}$ for D1 to LZ, and $R^2 = 0.70$ and $p = 2.2\times10^{-16}$ for D1 to P). 1,715 genes showed a statistically significant signature for posttranscriptional downregulation for D1 to LZ (Fig. 7A and C), and 1,622 for D1 to P (Fig. 7B and D). As expected, predicted targets of the miR-29 and miR-34 families were preferentially found among those genes that had significantly reduced steady-state mRNA levels (i.e. had reduced exon levels during meiosis as determined by RNAseq) and were post-transcriptionally downregulated (i.e., had more significantly reduced exon levels than intron levels during meiosis as determined by EISA; Fig. 7A-D; $p < 2.0\times10^{-11}$ for miR-29 family targets in D1 to LZ; $p < 7.0\times10^{-9}$ for miR-29 family targets in D1 to P; $p < 2.7\times10^{-6}$ for miR-34 family targets in D1 to LZ; $p < 2.1\times10^{-8}$ for miR-34 family targets in D1 to P). miR-29 and miR-34 family predicted target mRNAs were not preferentially enriched, however, in other post-transcriptional categories, such as steady-state upregulated (i.e., had increased exon levels during meiosis as determined by RNAseq) and post-transcriptionally upregulated (i.e., had more significantly increased exon levels than intron levels during meiosis as determined by EISA), or steady-state upregulated but post-transcriptionally downregulated. Notably, absence of miR-29 and miR-34 predicted targets in steady-state upregulated but post-transcriptionally downregulated genes indicates that miRNA-mediated regulation in meiotic cells is typically not antagonistic to transcriptional regulation acting on the same genes, a characteristic of miRNA-mediated regulation that has been observed previously.[59] We found that ~20% of the predicted targets of the miR-29 and miR-34 families showed evidence of post-transcriptional regulation in either of the 2 comparisons (Fig. 7E - purple). There were also downregulated targets that did not show any evidence of post-transcriptional downregulation (Fig. 7E - dark blue); at least some of these are likely downregulated via changes in transcription. In general, the relative frequencies of different types of regulation predicted by EISA for targets of both the miR-34 and miR-29 families were highly similar to one another. These data confirm the targeting signature for the miR-34 and miR-29 families, and identify a high-confidence set of predicted direct targets for both miRNAs during meiosis.

To determine what roles the miR-29 and miR-34 families might play in meiotic cells, we performed gene ontology analysis on predicted direct targets of each miRNA, analyzing genes that were both downregulated at the post-transcriptional level and that exhibited a significant decrease in steady-state mRNA levels. We noted a strong enrichment in functional categories involved in collagen and the extracellular matrix for predicted direct targets of the miR-29 family in D1 to LZ and D1 to P

Downloaded by [New York University] at 11:53 11 October 2017

Downloaded by [New York University] at 11:53 11 October 2017

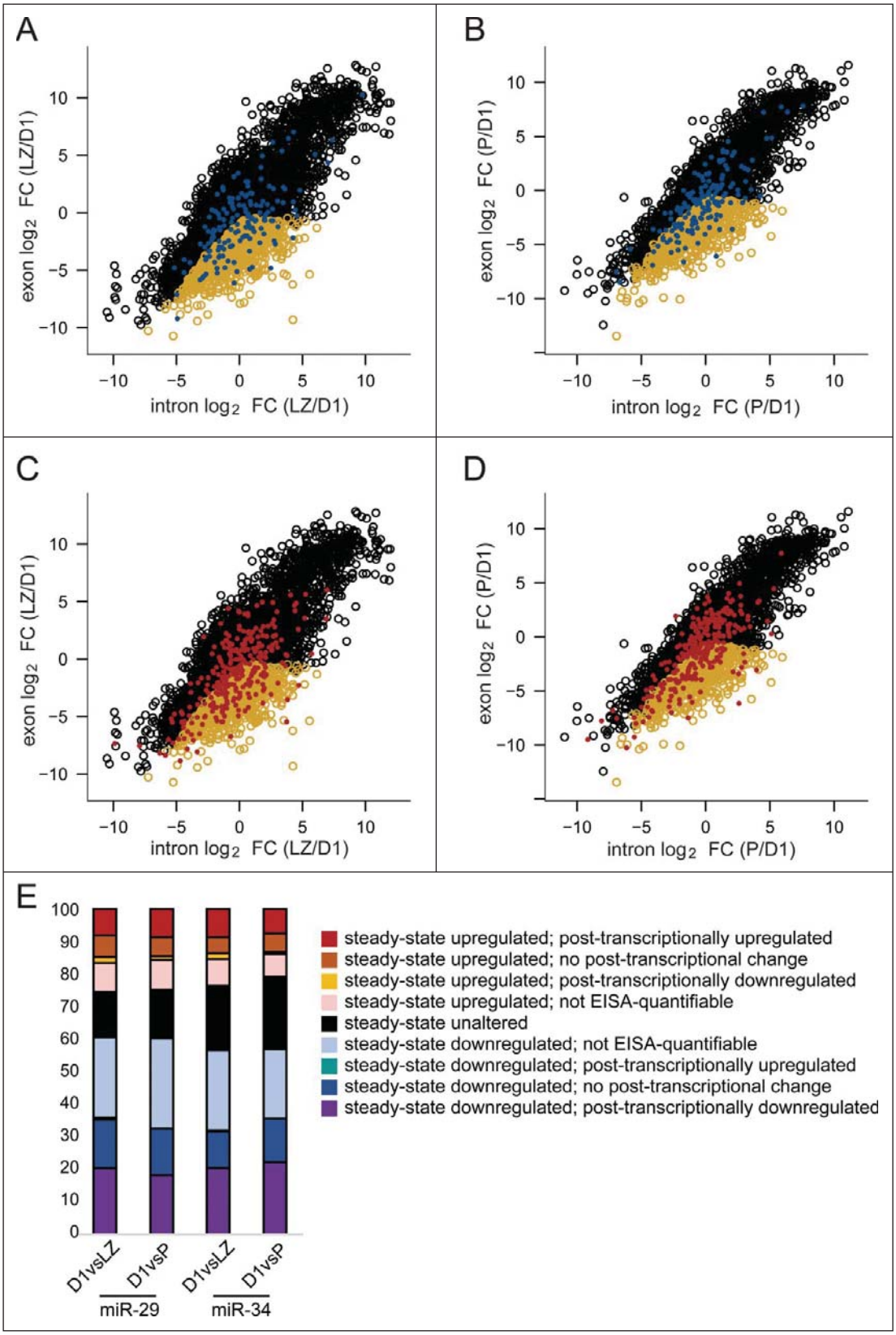


Figure 7. Post-transcriptionally regulated targets of the miR-29 and miR-34 families. (A-B) Log_2 fold change of exon (y axis) vs intron (x axis) levels for all transcripts with sufficient exon and intron coverage for EISA analysis. Transcripts determined to undergo significant post transcriptional downregulation are shown in yellow; those containing conserved miR-34 target sites are shown in blue. Shown for the D1 to LZ (A) and D1 to P (B) comparisons. (C-D) Log_2 fold change of exon (y axis) vs log_2 fold change of intron (x axis) levels for all transcripts with suitable exon and intron coverage for EISA analysis. Transcripts containing conserved miR-29 target sites are shown in red, otherwise as described in panels A and B. (E) Proportion of targets of miR-29 and miR-34 separated by indicated categories of regulation, for the D1 to LZ and D1 to P comparisons.

Downloaded by [New York University] at 11:53 11 October 2017

(Table SXI, Table SXIII – orange-coded genes). Such targets included a number of genes encoding collagen chains, including *Col15a1, Col1a2, Col27a1, Col3a1, Col4a4, Col4a5, Col5a1, Col6a2, and Col6a3* (Table SXIII). Regulation of collagen and other extracellular matrix components by the miR-29 family has been described in multiple non-germline studies.[60,61] We did not identify enrichment in categories such as response to DNA damage or methylation, 2 processes miR-29 has been previously implicated in, but did identify direct targets involved in DNA damage – *Ing4* and Parg (Table SXIII – yellow-coded genes) – and methylation – *Gm17296, Jarid2*, and *Tet3* (Table SXIII – blue-coded genes); the ability of the miR-29 family to regulate Tet3 has been shown *in vitro*.[62] Notably, we did not have enough intron and exon counts to determine if the *de-novo* methyltransferase, *Dnmt3a*, a well-studied target of miR-29,[63,64] was changing post-transcriptionally. It was, however, significantly downregulated in meiotic cells as compared to post-meiotic cells ($\log_2$ fold change $= -5.1$, $p = 4.0 \times 10^{-34}$). As DNA methylation plays a role in the response to DNA damage,[65] it is possible that miR-29 might participate in this process through its downregulation of *Dnmt3a*. To validate this observation, as well as the downregulation of some of the miR-29 family's predicted direct targets, we used qPCR to quantify the fold change in mRNA levels for the D1 to LZ comparison for 11 predicted targets. Of the 11, we confirmed downregulation of 9, including *Dnmt3a* (Fig. S5E). The full list of steady-state downregulated and post-transcriptionally downregulated targets of the miR-29 family in meiotic samples as compared to premeiotic D1 can be found in Table SXIII.

We found no strong gene ontology signatures connected to processes in meiosis for the direct targets of the miR-34 family in the D1 to LZ comparison or the D1 to P comparison. Previous observations have suggested that the miR-34 family plays a role in ciliogenesis, at least in part through the repression of *Ccp110*, a gene involved in the inhibition of ciliogenesis. It was previously shown that *Ccp110* can be regulated by the miR-34 family in vitro, and that *Ccp110* overexpression phenocopied the cilia defects observed in miR-34 family knockout mice.[30] Surprisingly, we did not find evidence for post-transcriptional regulation of *Ccp110* in vivo; instead, it appears to be significantly upregulated in both our D1 to LZ and D1 to P comparisons, and also shows a post-transcriptional upregulation signature (D1 to LZ steady-state mRNA $\log_2$ fold change $= 1.5$, $p = 1.2\times10^{-4}$ and EISA $\log_2$ fold change $= 2.25$, $p = 1.8 \times 10^{-4}$; D1 to P steady-state mRNA $\log_2$ fold change $= 4.3$, $p = 2.7 \times 10^{-25}$ and EISA $\log_2$ fold change $= 3.0$, $p = 9.2 \times 10^{-19}$). Our results argue that either the miR-34 family does not affect *Ccp110* levels until later in spermatogenesis, or, that miR-34 plays a role in ciliogenesis via regulation of a different target mRNA. We identified 2 miR-34 family targets with annotated roles in ciliogenesis that were strongly downregulated and also post-transcriptionally downregulated: *Notch2* and *Prkacb* ($\log_2$ fold changes -4.9 and -3.9, respectively), defective regulation of which might explain the defects in ciliogenesis observed in miR-34 knockout mice. A previous study also noted a decrease in steady-state Notch2 mRNA levels in meiotic cells and suggested it might be regulated by miR-34[45]. We also found no significant post-transcriptional downregulation for another predicted target of miR-34, *Atf1*, which was suggested to play a role in regulating apoptosis in spermatocytes[33] (D1 to LZ steady-state mRNA $\log_2$ fold change $= 1.2$, $p = 3.5 \times 10^{-4}$ and EISA $\log_2$ fold change $= 0.3$, $p = 5.9 \times 10^{-1}$; D1 to P steady-state mRNA $\log_2$ fold change $= 1.3$, $p = 5.7 \times 10^{-5}$ and EISA $\log_2$ fold change $= 0.2$, $p = 7.1 \times 10^{-1}$). Our full list of downregulated targets of the miR-34 family with post-transcriptional downregulation can be found in Table SXIII.

Discussion

Conditional germline knockouts of *Dgcr8*, *Drosha*, and *Dicer* revealed that miRNAs were required for spermatogenesis, and further studies identified the miR-34 family as essential to normal meiosis in the male germline. Our work confirms this observation, revealing widespread control of mRNA levels by the miR-34 family in meiotic cells. In addition, we identify miR-29 as a second miRNA family likely important for spermatogenesis. Indeed, the predicted impact of the miR-29 family on the transcriptome is more widespread than that mediated by the miR-34 family in many premeiotic to meiotic comparisons (Fig. 4I, Fig. S4E-F, H). Finally, we show that targets of the miR-29 family are distinct from those of miR-34, and thus the role of the miR-29 family is independent from that of miR-34.

While roles for the miR-29 family in cellular processes in somatic cells have been investigated, little is known about the role for this microRNA in the germ line, particularly during meiosis. A previous study identified an increase in mmu-mir-29b in murine meiotic as compared to premeiotic whole testis,[66] suggesting that the increase in miR-29 family abundance we observe in our study is not due to contamination present only in our purified meiotic cells. Another study found that exposure of the murine germ line to radiation leads to an increase in germline miR-29 family abundance and downregulation of *Dnmt3a*.[67] *Dnmt3a* plays a role in the methylation and suppression of transposable elements; transposable elements were also found to be hypomethylated in irradiated mice.[67] While we could not confirm that *Dnmt3a* was post-transcriptionally downregulated in meiotic cell populations, we did find that *Dnmt3a* steady-state mRNA levels were significantly decreased. In normal, non-irradiated male germ lines, the miR-29 family was most abundant in leptotene and zygotene spermatocytes (Fig. 5G), which is when DNA double strand breaks have been formed to promote crossing over, and also a time when the cell's DNA damage response pathway has been activated. Interestingly, in addition to the aforementioned germline radiation study, the miR-29 family has been shown to be upregulated upon drug-induced DNA damage.[47] Thus, these results suggest that the miR-29 family might play an important role in the repair of double strand breaks during meiosis. Finally, we found that conserved targets of the miR-29 family exhibiting signatures of post-transcriptional downregulation were enriched for components of the extracellular matrix and collagen; many other studies have noted the connection between the miR-29 family and the extracellular matrix in somatic cells.[60,61] In the male germ line, interactions between germ cells, supporting somatic cells (including Sertoli cells), and the basement membrane (a type of extracellular matrix) of the seminiferous tubules are important for proper spermatogenesis.[34] Our work suggests that the miR-29 family regulates

multiple extracellular matrix components, and thus might play a role in germ cell-basement membrane dynamics.

Taken together, our work both confirms the importance of the miR-34 family in regulating mRNA abundance during male meiosis, as well as implicates the miR-29 family as a post-transcriptional regulator during meiosis. Importantly, our work provides a list of confidently identified predicted direct targets for the miR-29 family, which can be further investigated in order to better define the germline functions of miR-29. Given the previously identified role in DNA damage repair for the miR-29 family and our observation that it also regulates many extracellular matrix genes, a better understanding of the miR-29 family's function may contribute to a better understanding of DNA damage repair and cellular integrity during meiosis.

Materials and methods

Whole testis and spermatocyte preparation

All mouse studies were conducted with the prior approval of the Cornell Institutional Animal Care and Use Committee. To obtain premeiotic whole testis samples, whole testis were removed and decapsulated from C57BL/6 mice at day 1, day 3, and day 7 pp. The proportion of germline cells to somatic cells in these samples is estimated to be ~7:3. To obtain meiotic cells, testes from adult C57BL/6 mice (day 70–80 pp) were removed and decapsulated prior to enrichment of specific spermatogenic cell types using the STA-PUT method based on separation by cell diameter/density at unit gravity.[68,69] The purity of resulting fractions was determined by microscopy based on cell diameter and morphology. Pachytene and pooled leptotene/zygotene cells were approximately 90-95% pure, with potential contamination from spermatocytes of slightly earlier or later developmental timing.

RNA preparation

RNA was extracted from whole testis and STA-PUT purified cells using TRIzol (Life Technologies) and used as the source material for mRNA and small RNA sequencing.

Small RNA library preparation and analysis

Small RNA sequencing libraries were prepared (TruSeq Small RNA, Illumina; 11 cycles of amplification) and sequenced (Illumina HiSeq 2500; 50 bp; single). High-quality reads (those passing Illumina's Y/N quality filter and containing no uncalled bases) were aligned to the genome (mm9) using Bowtie (v0.12.7;[70]). miRNAs were identified requiring perfect alignment to miRNA sequences obtained from miRBase (v21;[2,3]), and grouped together into miRNA families based on the possession of a common seed sequence. The differential abundance of each miRNA between samples was assessed using edgeR (v3.6.8;[71]; R version: 3.1.0). miRNA families are referred to by the first miRNA member of each family based on TargetScan miRNA family nomenclature (for example, the let-7/98/4458/4500 family is referred to here as the let-7 family).

RNA library preparation and analysis

Non-stranded RNAseq libraries were prepared (TRUseq, Illumina; 15 cycles of amplification) and sequenced (Illumina HiSeq 2500; 100 and 50 bp; single). The resulting sequences were mapped to the genome (mm9) using BWA aln (v0.7.8; default settings used;[72]). The number of reads mapping to each mRNA transcript was quantified using HTSeq (v0.6.1; intersection-strict mode and stranded = no settings used, otherwise, default settings were used;[73]), and $\log_2$ fold change abundance of each mRNA between samples was assessed using edgeR (v3.6.8;[71]; R version: 3.1.0).

Accession numbers

Deep sequencing files are available from NCBI GEO (GSE83264).

Targeting analysis

Targets for each broadly conserved[11] miRNA family were identified using the TargetScan Mouse database (v6.2;[8-11]). Only expressed targets (RNAseq CPM > = 1) with a TargetScan context+ score less than -0.1 were considered. A background distribution of miRNA targeting was constructed by sampling 10% of the expressed targets of every miRNA family and pooling their $\log_2$ fold change values together. The distribution of $\log_2$ fold change abundance for mRNAs targeted by each miRNA family were compared to the distribution of background $\log_2$ fold change abundance using a Wilcoxon rank sum test. The custom python script used to perform these comparisons are available on GitHub: https://github.com/SRHilz/2016TargetingAnalysis. For further investigation of miR-29 family targets (Fig. 5), PCT values were used to determine to what degree an mRNA was a conserved target of a miRNA family.

Network analysis

The same strategy to identify targets for each miRNA family in the targeting analysis was used. To enrich for functionally significant targets, only targets with a TargetScan context + score less than −0.1 and that were also conserved (TargetScan PCT > 0.5) were considered. Significance testing using the probability mass function for the binomial distributions (where k = number of shared predicted targets between family A and B, n = number of predicted targets for family A, and p = probability of randomly selecting a predicted target of family B) was used to determine whether 2 miRNA families had a proportion of overlapping targets that significantly deviated from expected values (p-values Bonferroni-corrected). Shared targets for all highly expressed miRNA were visualized using heatmap.2 in R (R version: 3.2.3). The network of shared targets for miR-34 and miR-29 was visualized using the igraph package in R.

Gene ontology analysis

Gene ontology analysis was performed with the Bioconductor package goseq.[74] P-values were corrected for multiple hypothesis testing using false discovery rate (FDR).

EISA analysis

Identification of transcripts with post-transcriptional gene regulatory signatures were identified as described in Gaidatzis et al.

Downloaded by [New York University] at 11:53 11 October 2017

2015[57]. Briefly, the number of reads corresponding to introns and exons were separately counted. Transcripts with significantly different Δ exon - Δ intron counts were identified using edgeR. Significantly different transcripts were then cross-referenced with a list of all significantly downregulated genes for a given comparison in order to specifically identify transcripts that are steady-state downregulated and post-transcriptionally downregulated. edgeR was also used to determine the individual $\log_2$ fold change exon and intron counts shown in Fig. 7.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Acknowledgments

The authors would like to thank Peter Borst for assistance with mouse care, the Cornell Genomics Facility for their help with high-throughput sequencing, and Jen Grenier and Core B of P50HD076210 for helpful suggestions and advice.

Funding

This work was supported by funding from the National Centers for Translational Research in Reproduction and Infertility (NCTRI) (P50HD076210 to P.E.C. and A.G.) and R01GM105668 (to A.G.).

ORCID

Stephanie Hilz http://orcid.org/0000-0002-6995-6719

References

1. Kim N, Han J, Siomi M. Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol 2009; 10(2):126-139; PMID:19165215; http://dx.doi.org/10.1038/nrm2632
2. Kozomara A, Griffiths Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 2014; 42 (Database issue):D68-73; PMID:24275495; http://dx.doi.org/10.1093/nar/gkt1181
3. Griffiths Jones S, Grocock R, van Dongen S, Bateman A, Enright A. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 2006; 34(Database issue):D140-4; PMID:16381832; http://dx.doi.org/10.1093/nar/gkj112
4. Lagos Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. Science (New York, NY) 2001; 294(5543):853-8; PMID:11679670; http://dx.doi.org/10.1126/science.1064921
5. Desvignes T, Batzel P, Berezikov E, Eilbeck K, Eppig JT, McAndrews MS, Singer A, Postlethwait JH. miRNA nomenclature: A view incorporating genetic origins, Biosynthetic Pathways, and sequence variants. Trends Genet 2015; 31(11):613-26; PMID:26453491; http://dx.doi.org/10.1016/j.tig.2015.09.002
6. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell 2009; 136(2):215-33; PMID:19167326; http://dx.doi.org/10.1016/j.cell.2009.01.002
7. Agarwal V, Bell GW, Nam J WW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. Elife 2015; 4:e05005; PMID:26267216; http://dx.doi.org/10.7554/eLife.05005
8. Lewis B, Burge C, Bartel D. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 2005; 120(1):15-20; PMID:15652477; http://dx.doi.org/10.1016/j.cell.2004.12.035
9. Grimson A, Farh K, Johnston W, Garrett Engele P, Lim L, Bartel D. MicroRNA targeting specificity in mammals: Determinants beyond seed pairing. Mol Cell 2007; 27(1):91-105; PMID:17612493; http://dx.doi.org/10.1016/j.molcel.2007.06.017
10. Garcia DM, Baek D, Shin C, Bell GW, Grimson A, Bartel DP. Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. Nat Struct Mol Biol 2011; 18 (10):1139-46; PMID:21909094; http://dx.doi.org/10.1038/nsmb.2115
11. Friedman R, Farh K, Burge C, Bartel D. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 2009; 19(1):92-105; PMID:18955434; http://dx.doi.org/10.1101/gr.082701.108
12. Eichhorn SW, Guo H, McGeary SE, Rodriguez-Mias RA, Shin C, Baek D, Hsu SH, Ghoshal K, Villén J, Bartel DP. mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues. Mol Cell 2014; 56(1):104-15; PMID:25263593; http://dx.doi.org/10.1016/j.molcel.2014.08.028
13. Hayashi K, de Lopes H, Kaneda M, Tang F, Hajkova P, Lao K, O'Carroll D, Das PP, Tarakhovsky A, Miska EA, et al. MicroRNA biogenesis is required for mouse primordial germ cell development and spermatogenesis. PloS One 2008; 3(3):e1738; PMID:18320056; http://dx.doi.org/10.1371/journal.pone.0001738
14. Modzelewski AJ, Hilz S, Crate EA, Schweidenback CT, Fogarty EA, Grenier JK, Freire R, Cohen PE, Grimson A. Dgcr8 and Dicer are essential for sex chromosome integrity during meiosis in males. J Cell Sci 2015; 128(12):2314-27; PMID:25934699; http://dx.doi.org/10.1242/jcs.167148
15. Maatouk D, Loveland K, McManus M, Moore K, Harfe B. Dicer1 is required for differentiation of the mouse male germline. Biol Reprod 2008; 79(4):696-703; PMID:18633141; http://dx.doi.org/10.1095/biolreprod.108.067827
16. Romero Y, Meikar O, Papaioannou MD, Conne B, Grey C, Weier M, Pralong F, De Massy B, Kaessmann H, Vassalli JD, et al. Dicer1 depletion in male germ cells leads to infertility due to cumulative meiotic and spermiogenic defects. PLoS One 2011; 6(10):e25241; PMID:21998645; http://dx.doi.org/10.1371/journal.pone.0025241
17. Greenlee A, Shiao MS, Snyder E, Buaas FW, Gu T, Stearns TM, Sharma M, Murchison EP, Puente GC, Braun RE.. Deregulated sex chromosome gene expression with male germ cell-specific loss of Dicer1. PLoS One 2012; 7(10):e46359; PMID:23056286; http://dx.doi.org/10.1371/journal.pone.0046359
18. Zimmermann C, Romero Y, Warnefors M, Bilican A, Borel C, Smith LB, Kotaja N, Kaessmann H, Nef S. Germ cell-specific targeting of DICER or DGCR8 reveals a novel role for Endo-siRNAs in the progression of mammalian spermatogenesis and male fertility. PLoS One 2014; 9(9):e107023; PMID:25244517; http://dx.doi.org/10.1371/journal.pone.0107023
19. Wu Q, Song R, Ortogero N, Zheng H, Evanoff R, Small CL, Griswold MD, Namekawa SH, Royo H, Turner JM, et al. The RNase III Enzyme DROSHA is essential for MicroRNA production and spermatogenesis. J Biol Chem 2012; 287(30):25173-90; PMID:22665486; http://dx.doi.org/10.1074/jbc.M112.362053
20. Korhonen H, Meikar O, Yadav R, Papaioannou MD, Romero Y, Da Ros M, Herrera PL, Toppari J, Nef S, Kotaja N. Dicer is required for Haploid male germ cell differentiation in mice. PLoS One 2011; 6(9):e24821; PMID:21949761; http://dx.doi.org/10.1371/journal.pone.0024821
21. Saitou M, Kagiwada S, Kurimoto K. Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. Dev (Cambridge, England) 2012; 139(1):15-31; PMID:22147951; http://dx.doi.org/10.1242/dev.050849
22. Tseng Y-TT, Liao H-FF, Yu C-YY, Mo C-FF, Lin S-PP. Epigenetic factors in the regulation of prospermatogonia and spermatogonial stem cells. Reproduction 2015; 150(3):R77-91; PMID:26116003; http://dx.doi.org/10.1530/REP-14-0679
23. De Rooij DG. Proliferation and differentiation of spermatogonial stem cells. Reproduction (Cambridge, England) 2001; 121(3):347-54; PMID:11226060; http://dx.doi.org/10.1530/rep.0.1210347
24. Bellvé AR, Cavicchia JC, Millette CF, O'Brien DA, Bhatnagar YM, Dym M. Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization. J Cell Biol 1977; 74(1):68-85; PMID:874003; http://dx.doi.org/10.1083/jcb.74.1.68
25. Borde V, de Massy B. Programmed induction of DNA double strand breaks during meiosis: setting up communication between DNA and the chromosome structure. Curr Opin Genetics Dev 2013; 23(2):147-55; PMID:23313097; http://dx.doi.org/10.1016/j.gde.2012.12.002

Downloaded by [New York University] at 11:53 11 October 2017

26. Turner J. Meiotic silencing in mammals. Annual Rev Genetics 2015; 49:395-412; PMID:26631513; http://dx.doi.org/10.1146/annurev-genet-112414-055145
27. Hunter N. Meiotic Recombination: The essence of heredity. Cold Spring Harb Perspect Biol 2015; 7(12); PMID:26511629; http://dx.doi.org/10.1101/cshperspect.a016618
28. Braun RE. Post-transcriptional control of gene expression during spermatogenesis. Seminars Cell Dev Biol 1998; 9(4):483-9; PMID:9813196; http://dx.doi.org/10.1006/scdb.1998.0226
29. Comazzetto S, Giacomo M, Rasmussen K, Much C, Azzi C, Perlas E, Morgan M, O'Carroll D. Oligoasthenoteratozoospermia and infertility in mice deficient for miR-34b/c and miR-449 loci. PLoS Genetics 2014; 10(10):e1004597; PMID:25329700; http://dx.doi.org/10.1371/journal.pgen.1004597
30. Song R, Walentek P, Sponer N, Klimke A, Lee JS, Dixon G, Harland R, Wan Y, Lishko P, Lize M, et al. miR-34/449 miRNAs are required for motile ciliogenesis by repressing cp110. Nature 2014; 510(7503):115-20; PMID:24899310; http://dx.doi.org/10.1038/nature13413
31. Wu J, Bao H, Kim M, Yuan S, Tang C, Zheng H, Mastick GS, Xu C, Yan W. Two miRNA clusters, miR-34b/c and miR-449, are essential for normal brain development, motile ciliogenesis, and spermatogenesis. Proc Natl Acad Sci U S A 2014; 111(28):E2851-7; PMID:24982181; http://dx.doi.org/10.1073/pnas.1407777111
32. Hilz S, Modzelewski AJ, Cohen PE, Grimson A. The roles of microRNAs and siRNAs in mammalian spermatogenesis. Development 2016; 143(17):3061-73; PMID:27578177; http://dx.doi.org/10.1242/dev.136721
33. Liang X, Zhou D, Wei C, Luo H, Liu J, Fu R, Cui S. MicroRNA-34c enhances murine male germ cell apoptosis through targeting ATF1. PloS One 2012; 7(3):e33861; PMID:22479460; http://dx.doi.org/10.1371/journal.pone.0033861
34. Siu MK, Cheng CY. Extracellular matrix and its role in spermatogenesis. Adv Exp Med Biol 2008; 636:74-91; PMID:19856163; http://dx.doi.org/10.1007/978-0-387-09597-4_5
35. He Z, Feng L, Zhang X, Geng Y, Parodi DA, Suarez-Quian C, Dym M. Expression of Col1a1, Col1a2 and procollagen I in germ cells of immature and adult mouse testis. Reproduction 2005; 130(3):333-41; PMID:16123240; http://dx.doi.org/10.1530/rep.1.00694
36. Bellvé AR. Purification, culture, and fractionation of spermatogenic cells. Meth Enzymol 1993; 225:84-113; PMID:8231890; http://dx.doi.org/10.1016/0076-6879(93)25009-Q
37. Weick E-MM, Miska EA. piRNAs: from biogenesis to function. Development (Cambridge, England) 2014; 141(18):3458-71; PMID:25183868; http://dx.doi.org/10.1242/dev.094037
38. Tan T, Zhang Y, Ji W, Zheng P. miRNA signature in mouse spermatogonial stem cells revealed by high-throughput sequencing. Biomed Res Int 2014; 2014:154251; PMID:25136556; http://dx.doi.org/10.1155/2014/154251
39. Gan H, Lin X, Zhang Z, Zhang W, Liao S, Wang L, Han C. piRNA profiling during specific stages of mouse spermatogenesis. RNA 2011; 17 (7):1191-203; PMID:21602304; http://dx.doi.org/10.1261/rna.2648411
40. Mukherji S, Ebert MS, Zheng GX, Tsang JS, Sharp PA, van Oudenaarden A. MicroRNAs can generate thresholds in target gene expression. Nature Genetics 2011; 43(9):854-9; PMID:21857679; http://dx.doi.org/10.1038/ng.905
41. Bosson AD, Zamudio JR, Sharp PA. Endogenous miRNA and target concentrations determine susceptibility to potential ceRNA competition. Mol Cell 2014; 56(3):347-59; PMID:25449132; http://dx.doi.org/10.1016/j.molcel.2014.09.018
42. Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu AL, Zhao Y, McDonald H, Zeng T, Hirst M, et al. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Res 2008; 18(4):610; PMID:18285502; http://dx.doi.org/10.1101/gr.7179508
43. Hartigan JA, Wong MA. Algorithm AS 136: A K-Means Clustering Algorithm. Applied Statistics 1979; 28:100-108; http://dx.doi.org/10.2307/2346830
44. Cheng C, Bhardwaj N, Gerstein M. The relationship between the evolution of microRNA targets and the length of their UTRs. BMC Genomics 2009; 10:431; PMID:19751524; http://dx.doi.org/10.1186/1471-2164-10-431
45. Bouhallier F, Allioli N, Lavial F, Chalmel F, Perrard MH, Durand P, Samarut J, Pain B, Rouault JP. Role of miR-34c microRNA in the late steps of spermatogenesis. RNA 2010; 16(4):720-31; PMID:20150330; http://dx.doi.org/10.1261/rna.1963810
46. Bao J, Li D, Wang L, Wu J, Hu Y, Wang Z, Chen Y, Cao X, Jiang C, Yan W, et al. MicroRNA-449 and microRNA-34b/c function redundantly in murine testes by targeting E2F transcription factor-retinoblastoma protein (E2F-pRb) pathway. J Biol Chem 2012; 287(26):21686-98; PMID:22570483; http://dx.doi.org/10.1074/jbc.M111.328054
47. Ugalde A, Ramsay A, de la Rosa J, Varela I, Mariño G, Cadiñanos J, Lu J, Freije JM, López-Otín C. Aging and chronic DNA damage response activate a regulatory pathway involving miR-29 and p53. EMBO J 2011; 30(11):2219-32; PMID:21522133; http://dx.doi.org/10.1038/emboj.2011.124
48. Fabbri M, Garzon R, Cimmino A, Liu Z, Zanesi N, Callegari E, Liu S, Alder H, Costinean S, Fernandez-Cymering C, et al. MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. Proc Natl Acad Sci U S A 2007; 104(40):15805-10; PMID:17890317; http://dx.doi.org/10.1073/pnas.0707628104
49. Kole AJ, Swahari V, Hammond SM, Deshmukh M. miR-29b is activated during neuronal maturation and targets BH3-only genes to restrict apoptosis. Genes Dev 2011; 25(2):125-30; PMID:21245165; http://dx.doi.org/10.1101/gad.1975411
50. Wissink EM, Smith NL, Spektor R, Rudd BD, Grimson A. MicroRNAs and their targets are differentially regulated in adult and neonatal mouse CD8+ T cells. Genetics 2015; 201(3):1017-30; PMID:26416483; http://dx.doi.org/10.1534/genetics.115.179176
51. Li J, Cen B, Chen S, He Y. MicroRNA-29b inhibits TGF-β1-induced fibrosis via regulation of the TGF-β1/Smad pathway in primary human endometrial stromal cells. Mol Med Rep 2016; 13(5):4229-37; PMID:27035110; http://dx.doi.org/10.3892/mmr.2016.5062
52. Kriegel AJ, Liu Y, Fang Y, Ding X, Liang M. The miR-29 family: genomics, cell biology, and relevance to renal and cardiovascular injury. Physiol Genomics 2012; 44(4):237-44; PMID:22214600; http://dx.doi.org/10.1152/physiolgenomics.00141.2011
53. Cushing L, Kuang PP, Qian J, Shao F, Wu J, Little F, Thannickal VJ, Cardoso WV, Lü J. miR-29 is a major regulator of genes associated with pulmonary fibrosis. Am J Respir Cell Mol Biol 2011; 45(2):287-94; PMID:20971881; http://dx.doi.org/10.1165/rcmb.2010-0323OC
54. Krek A, Grün D, Poy M, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, et al. Combinatorial microRNA target predictions. Nat Genetics 2005; 37(5):495-500; PMID:15806104; http://dx.doi.org/10.1038/ng1536
55. Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA.org resource: targets and expression. Nucleic Acids Res 2008; 36(Database issue):D149-53; PMID:18158296; http://dx.doi.org/10.1093/nar/gkm995
56. Betel D, Koppal A, Agius P, Sander C, Leslie C. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. Genome Biol 2010; 11(8):R90; PMID:20799968; http://dx.doi.org/10.1186/gb-2010-11-8-r90
57. Gaidatzis D, Burger L, Florescu M, Stadler MB. Analysis of intronic and exonic reads in RNA-seq data characterizes transcriptional and post-transcriptional regulation. Nat Biotechnol 2015; 33(7):722-9; PMID:26098447; http://dx.doi.org/10.1038/nbt.3269
58. Gosline SJ, Gurtan AM, JnBaptiste CK, Bosson A, Milani P, Dalin S, Matthews BJ, Yap YS, Sharp PA, Fraenkel E. Elucidating MicroRNA regulatory networks using transcriptional, post transcriptional, and histone modification measurements. Cell Rep 2016; 14(2):310-9; PMID:26748710; http://dx.doi.org/10.1016/j.celrep.2015.12.031
59. Shkumatava A, Stark A, Sive H, Bartel DP. Coherent but overlapping expression of microRNAs and their targets during vertebrate development. Genes Dev 2009; 23(4):466-81; PMID:19240133; http://dx.doi.org/10.1101/gad.1745709
60. Maurer B, Stanczyk J, Jüngel A, Akhmetshina A, Trenkmann M, Brock M, Kowal-Bielecka O, Gay RE, Michel BA, Distler JH. MicroRNA 29, a key regulator of collagen expression in systemic sclerosis. Arthritis Rheum 2010; 62(6):1733-43; PMID:20201077; http://dx.doi.org/10.1002/art.27443

Downloaded by [New York University] at 11:53 11 October 2017

61. Wang, Liu J, Chen J, Feng T, Guo Q. MiR-29 mediates TGFβ 1 induced extracellular matrix synthesis through activation of Wnt/β catenin pathway in human pulmonary fibroblasts. Technol Health Care 2015; 23 Suppl 1:S119-25; PMID:26410313; http://dx.doi.org/10.3233/thc-150943
62. Oliveira LH, Schiavinato JL, Fráguas MS, Lucena-Araujo AR, Haddad R, Araújo AG, Dalmazzo LF, Rego EM, Covas DT, Zago M, et al. Potential roles of microRNA 29a in the molecular pathophysiology of T cell acute lymphoblastic leukemia. Cancer Sci 2015; 106(10):1264-77; PMID:26251039; http://dx.doi.org/10.1111/cas.12766
63. Hu W, Dooley J, Chung SS, Chandramohan D, Cimmino L, Mukherjee S, Mason CE, de Strooper B, Liston A, Park CY. miR-29a maintains mouse hematopoietic stem cell self renewal by regulating Dnmt3a. Blood 2015; 125(14):2206-16; PMID:25634742; http://dx.doi.org/10.1182/blood-2014-06-585273
64. Morita S, Horii T, Kimura M, Ochiya T, Tajima S, Hatada I. miR-29 represses the activities of DNA methyltransferases and DNA demethylases. Int J Mol Sci 2013; 14(7):14647-58; PMID:23857059; http://dx.doi.org/10.3390/ijms140714647
65. Cuozzo C, Porcellini A, Angrisano T, Morano A, Lee B, Di Pardo A, Messina S, Iuliano R, Fusco A, Santillo MR, et al. DNA damage, homology directed repair, and DNA methylation. PLoS Genet 2007; 3(7):e110; PMID:17616978; http://dx.doi.org/10.1371/journal.pgen.0030110
66. Yan N, Lu Y, Sun H, Tao D, Zhang S, Liu W, Ma Y. A microarray for microRNA profiling in mouse testis tissues. Reproduction (Cambridge, England) 2007; 134(1):73-9; PMID:17641090; http://dx.doi.org/10.1530/REP-07-0056
67. Filkowski J, Ilnytskyy Y, Tamminga J, Koturbash I, Golubov A, Bagnyukova T, Pogribny IP, Kovalchuk O. Hypomethylation and genome instability in the germline of exposed parents and their progeny is associated with altered miRNA expression. Carcinogenesis 2010; 31(6):1110-5; PMID:19959559; http://dx.doi.org/10.1093/carcin/bgp300
68. Bellvé AR. Purification, culture, and fractionation of spermatogenic cells. Methods Enzymol 1993; 225:84-113; PMID:8231890; http://dx.doi.org/10.1016/0076-6879(93)25009-Q
69. Modzelewski AJ, Holmes RJ, Hilz S, Grimson A, Cohen PE. AGO4 regulates entry into meiosis and influences silencing of sex chromosomes in the male mouse germline. Dev Cell 2012; 23(2):251-64; PMID:22863743; http://dx.doi.org/10.1016/j.devcel.2012.07.003
70. Langmead B, Trapnell C, Pop M, Salzberg S. Ultrafast and memory efficient alignment of short DNA sequences to the human genome. Genome Biol 2009; 10(3):R25; PMID:19261174; http://dx.doi.org/10.1186/gb-2009-10-3-r25
71. Robinson M, McCarthy D, Smyth G. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics (Oxford, England) 2010; 26(1):139-40; PMID:19910308; http://dx.doi.org/10.1093/bioinformatics/btp616
72. Li H, Durbin R. Fast and accurate short read alignment with Burrows Wheeler transform. Bioinformatics (Oxford, England) 2009; 25(14):1754-60; PMID:19451168; http://dx.doi.org/10.1093/bioinformatics/btp324
73. Anders S, Pyl P, Huber W. HTSeq a Python framework to work with high-throughput sequencing data. Bioinformatics (Oxford, England) 2014; PMID:25260700; http://dx.doi.org/10.1093/bioinformatics/btu638
74. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol 2010; 11(2):R14; PMID:20132535; http://dx.doi.org/10.1186/gb-2010-11-2-r14

Downloaded by [New York University] at 11:53 11 October 2017

A

	D1		D3		D7	
Genome Matching Counts	2085622	3796285	13763935	10504860	4683396	11285394
Hairpin Matching Counts	1421741	2200360	11061273	8511133	3392589	8591736

	LZ		P		SV	K
Genome Matching Counts	9023959	3074403	11158190	3835784	17712946	15042447
Hairpin Matching Counts	820464	119746	338443	112967	16838856	14666249

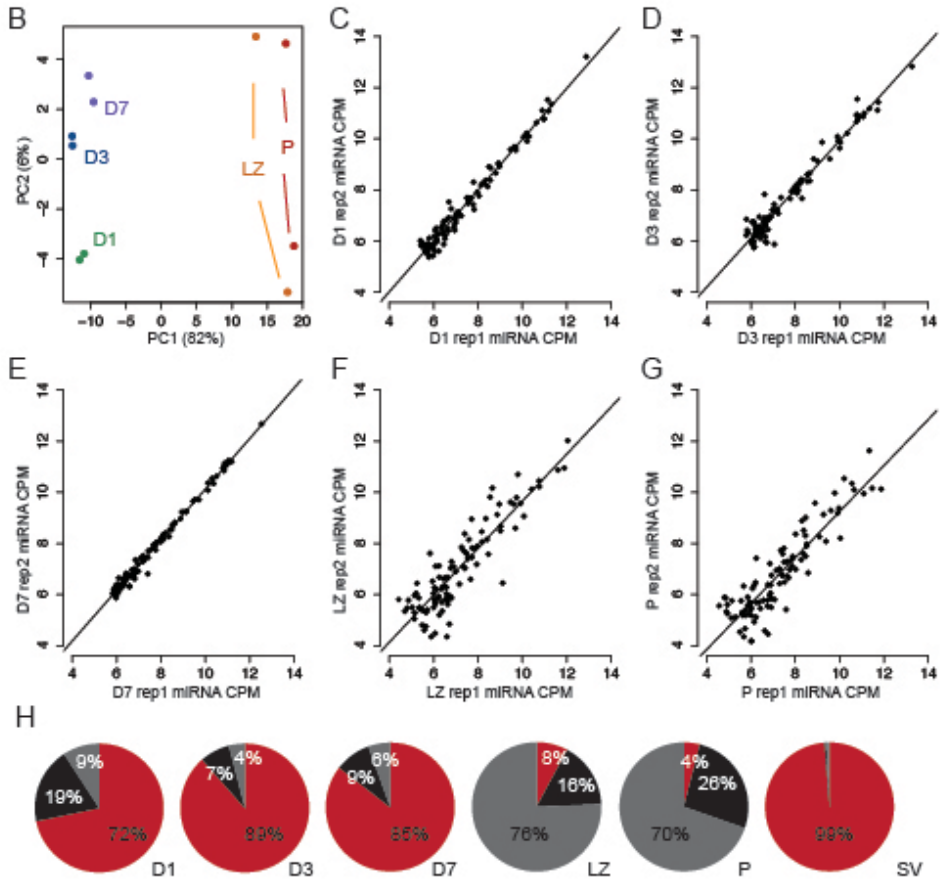


Fig. S1. Characteristics and quality of the small RNA library.

A The number of total genome matching small RNA reads and the subset of those reads that matched to miRNA hairpins are shown for replicates of each germline sample, as well as for somatic control seminal vesicle (SV) and kidney (K) samples.

B Alternative principle component analysis of miRNA family abundance (CPMs) for D1 (green), D3 (blue), D7 (purple), LZ (orange), and P (red) samples.

C-G Correlation of small RNA abundance (CPM) between replicate libraries for D1 (C; R^2 = 0.975), D3 (D; R^2 = 0.960), D7 (E; R^2 = 0.990), LZ (F; R^2 = 0.803), and P (G; R^2 = 0.810).

H Percent of small RNAs 18-22 nts (red), 26-28 nts (black), and 29-31 nts (gray) in length for D1, D3, D7, LZ, P, and SV samples.

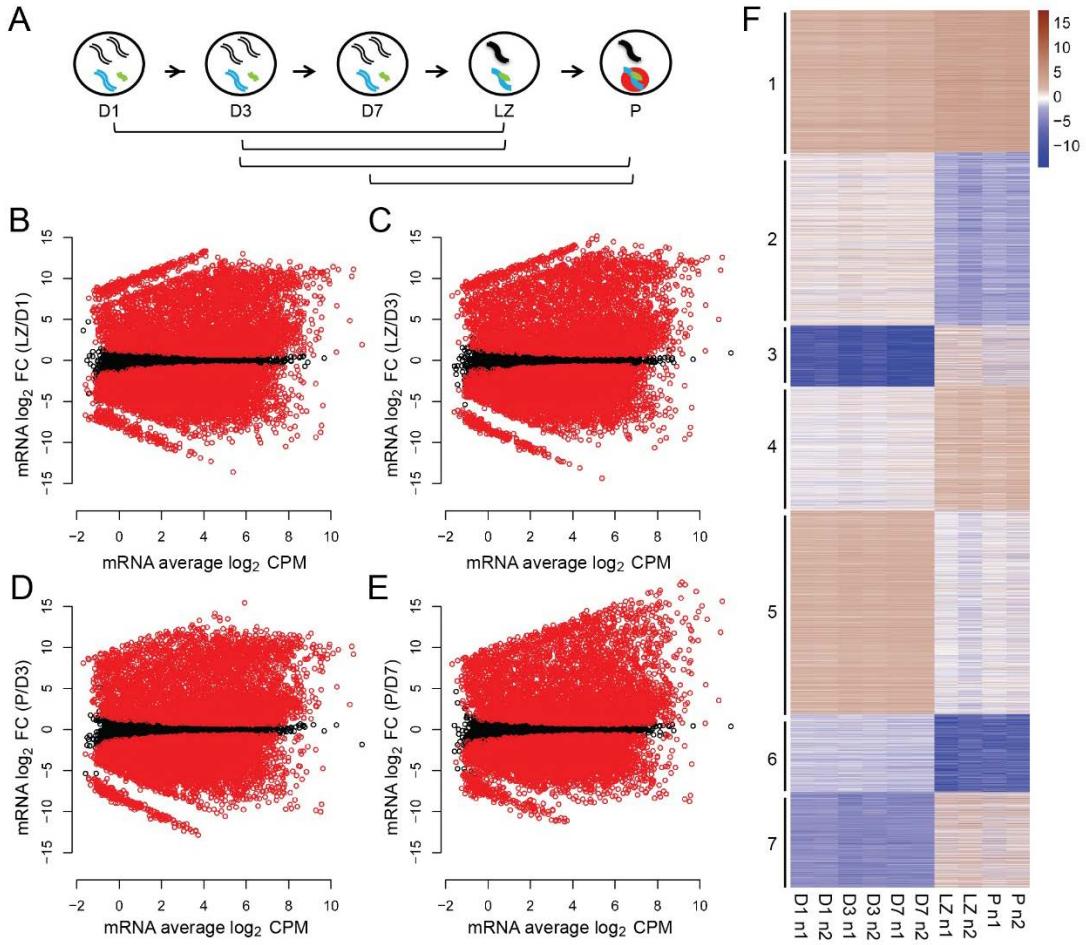


Fig. S2. Gene expression patterns are altered between premeiotic and meiotic samples.

A The $\log_2$ fold change in abundance of each mRNA transcript was calculated for each of the four comparisons shown.

B-E Differential abundance of mRNA transcripts was determined in premeiotic testis and early spermatocytes. Transcripts were compared in D1 & LZ (B), D3 & LZ (C), D3 & P (D), and D7 & P (E) samples. Plots show the average $\log_2$ abundance (as average $\log_2$ CPM; x-axis) for each

mRNA transcript and the transcript's $\log_2$ fold-change (FC) between the two compared samples (y-axis). Transcripts that are not differentially abundant at a multiple-comparison adjusted p-value of $p < 0.05$ are black; those that are differentially abundant are red.

F Differentially abundant transcripts between all premeiotic and meiotic samples were clustered ($k = 7$) by gene expression. Clusters are indicated by number (1-7).

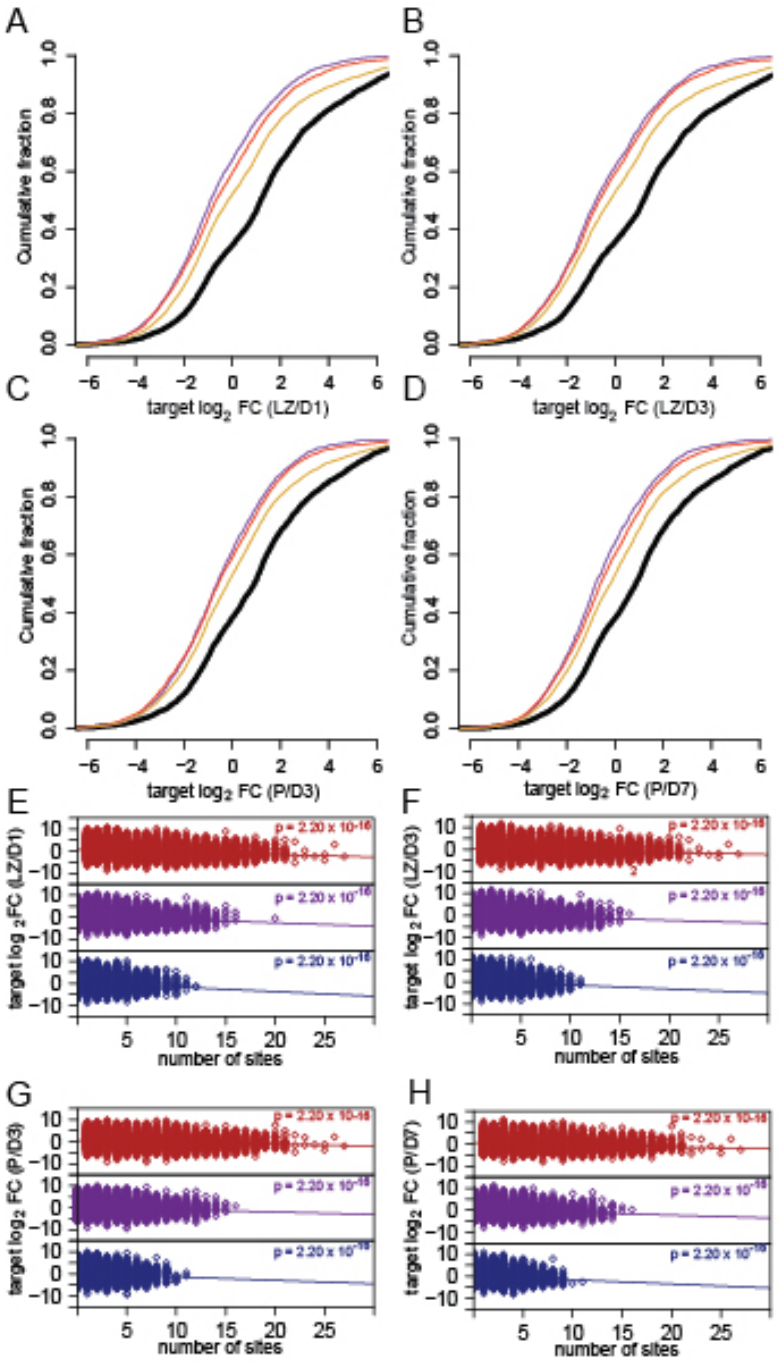


Fig. S3. RNAseq reveals net repression of transcripts targeted by highly conserved miRNAs in meiotic cells.

A-D The cumulative distribution of mRNA transcript $\log_2$ fold-change for each comparison. The cumulative distribution for mRNA transcripts not targeted by conserved miRNA families

appears black. The cumulative distributions for mRNA transcripts that are targets of between 1 and 3, 3 and 6, or greater than 6 conserved miRNA families appear yellow, red, and purple, respectively. The mRNA fold change distributions for mRNA that are targets of miRNAs (all colored lines) and are not (black line) were found to be significantly different ($p < 0.05$) for all comparisons, with $p = 2.2 \times 10^{-16}$ for D1 & LZ (B), $p = 2.2 \times 10^{-16}$ for D3 & LZ (C), $p = 2.2 \times 10^{-16}$ for D3 & P (D), $p = 2.2 \times 10^{-16}$ for D7 & P (E).

E-H The mRNA transcript $\log_2$ fold-change for each comparison compared to the number of sites of highly conserved miRNAs (red), the number of sites of highly conserved and also highly expressed miRNAs (purple), and the number of sites of highly conserved, highly expressed, and differentially abundant miRNAs (blue) for D1 & LZ (I), D3 & LZ (J), D3 & P (K), and D7 & P (L). The p-values for the regression of fold-change vs number of miRNA target sites are also shown for each comparison and category.

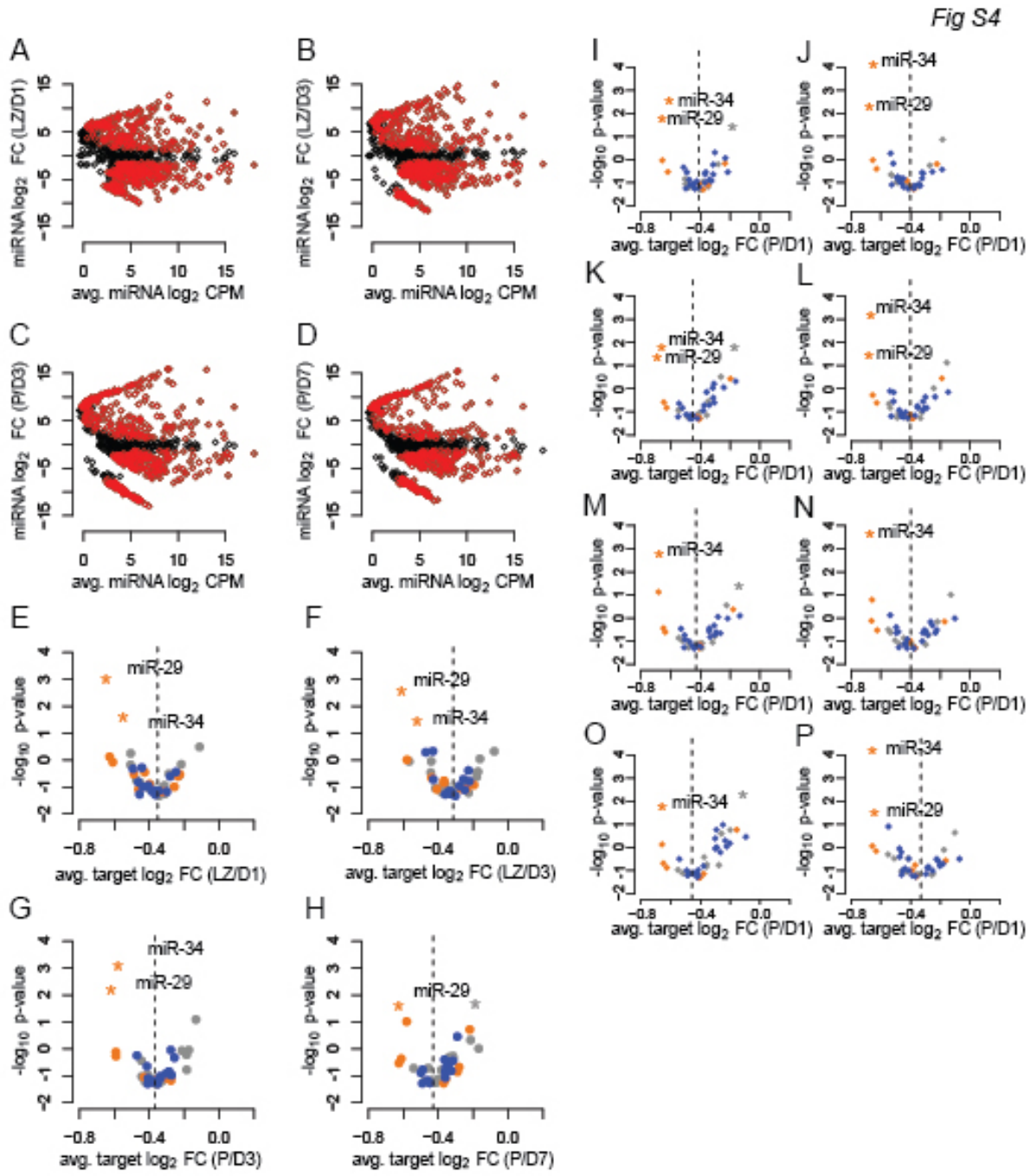


Fig. S4. miRNA targeting signature analysis reveals significant downregulation of miR-34 and miR-29 targets in early meiosis.

A-D Differential abundance of miRNA families was determined in premeiotic testis and early spermatocytes. miRNA families were compared in D1 & LZ (A), D3 & LZ (B), D3 & P (C), and

D7 & P (D) samples. Plots show the average log2 abundance (as CPM; x-axis) for each miRNA family and the families log2 fold-change (FC) between the two compared samples (y-axis). miRNA families that are not differentially abundant at a multiple-comparison adjusted pvalue of $p < 0.05$ are black; those that are differentially abundant are red.

E-H Targeting signature analysis for the four comparisons made in A-D. A Wilcoxon rank test was used to determine significant shifts in fold-change distributions of predicted mRNA targets of highly expressed (CPM ≥ 250) miRNA families as compared to a background fold-change distribution for targets of all other miRNA families (see Methods). The Bonferroni-corrected $-\log_{10}$ p-value for the Wilcoxon rank test is shown for each miRNA family, as well as the average $\log_2$ fold-change of predicted targets for that miRNA family (y-axis), for comparisons between D1 C LZ (I), D3 & LZ (J), D3 & P (K), and D7 & P (L) samples. miRNA families that did not change significantly in their abundance between compared samples are grey. miRNA families that are significantly more or less abundant are, respectively, orange and blue. miRNA families with no significant detectable shifts in target fold-change distribution are represented by dots; those with a detectable, significant shift in target fold-change distribution appear as an asterisk.

I-P Same as in E-H, except that the targeting analysis is performed for the D1 & P comparison at a variety of mRNA target transcript abundance cutoffs: target transcript CPM ≥ 1 (I; same as in Fig. 3L), CPM ≥ 2 (J), CPM ≥ 3 (K), CPM ≥ 4 (L), CPM ≥ 5 (M), CPM ≥ 6 (N), CPM ≥ 7 (O), and CPM ≥ 8 (P).

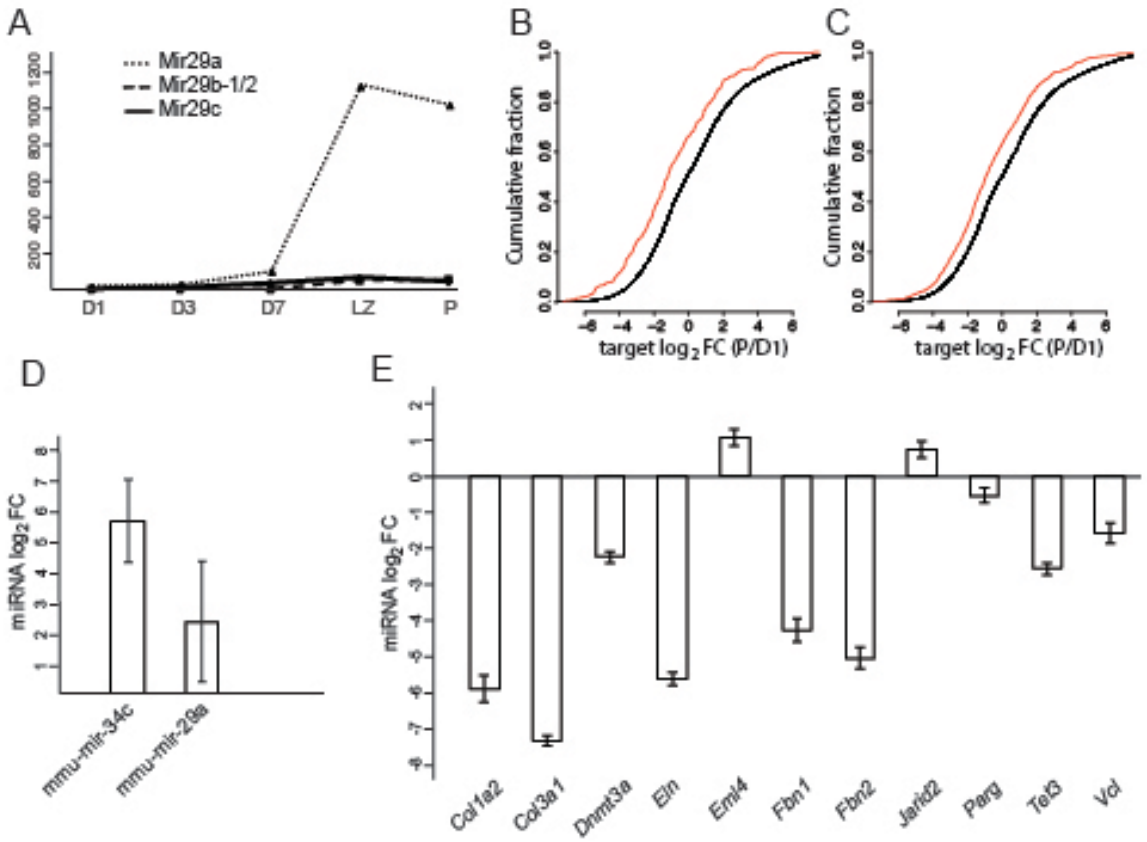


Fig. S5. Validation of miR-29 family expression and targeting signature.

A Levels of individual miR-29 family members bearing the miR-29 family seed sequence leading up to, and during, meiosis. The vast majority of reads with the miR-29 family seed come from the Mir29a, Mir29b-1, Mir29b-2, and Mir29c genes. Mir29b-1 and b-2 genes are highly similar, and in the germ line express identical products. Mir29a is the most abundantly expressed gene producing miR-29 family miRNAs.

B-C Cumulative distribution of $\log_2$ fold-changes (D1 to LZ) for mRNAs predicted to be targets of mmu-miR-29a as predicted by PicTar (B; all targets used were required to have a PicTar cutoff score of >1) or miRanda (C; August 2010 release, from the Mouse good SVR score list of targets) as compared to all other mRNAs expressed in D1 or LZ samples. Targets predicted by

PicTar as well as miRanda show a significant repression in LZ samples as compared to background ($p = 2.90 \times 10^{-12}$ and $p = 2.60 \times 10^{-16}$, respectively; Wilcoxon rank sum test).

D $\log_2$ fold-changes (D1 to LZ) of the most predominant miRNAs, mmu-mir-34c and mmu-mir-29a, belonging to the miR-34 and miR-29 families, respectively, as measured by qPCR. cDNA was prepared using the Universal cDNA Synthesis Kit II (Exiqon) from 10 ng of total RNA. qPCR reactions were run in triplicate on a LightCycler480 (Roche Applied Science). A melt curve for each reaction confirmed amplicon identity, and a standard curve was used to calculate transcript abundance. We assayed mmu-mir-34c using the hsa-miR-34c-5p LNA™ PCR primer set (Exiqon), mmu-mir-29a using the hsa-miR-29a-3p LNA™ PCR primer set (Exiqon), and mmu-mir-16, which we used as our reference sample, using the hsa-miR-16-5p LNA™ PCR primer set (Exiqon). All miRNAs shown were normalized to mmu-mir-16 levels. **E** $\log_2$ fold-changes (D1 to LZ) of eleven targets of the miR-29 family with roles in methylation, DNA damage repair, and formation of the extracellular that were repressed according RNAseq, as measured by qPCR. Assays performed were *Col1A2* (GGTGAGCCTGGTCAAACGG and ACTGTGTCCTTTCACGCCTTT), *Col3a1* (CTGTAACATGGAAACTGGGGAAA and CCATAGCTGAACTGAAAACCACC), *Dnmt3a* (ATGTGGTTCGGAGATGGCAAG and AGATGGCTTTGCGGTACATGG), *Eln* (GCTGGAGGTTTAGTGCCTGG and GCTCCGTATTTGGCAGCTTT), *Eml4* (GAGGTCGGCCAATTACAATGT and AGTCCTTTCCTCGATAACCGTA), *Fbn1* (TGTGGGGATGGATTCTGCTC and AGTGCCGATGTACCCTTTCTG), *Fbn2* (CCACTCCTATTGCTGCCCAG and TTGGGGCGGGAACAGAATC), *Jarid2* (GAAGGCGGTAAATGGGCTTCT and TCGTTGCTAGTAGAGGACACTT), *Parg* (AACGCCACCTCGTTTGTTTTC and CACAGAACTCATCATGGAGTCAA), *Tet3* (TGCGATTGTGTCGAACAAATAGT and TCCATACCGATCCTCCATGAG), *Vcl* (TGGACGGCAAAGCCATTCC and GCTGGTGGCATATCTCTCTTCAG), and *Gapdh* (TGAAGCAGGCATCTGAGGG and CGAAGGTGGAAGAGTGGGAG). All transcripts shown were normalized to *Gapdh*.

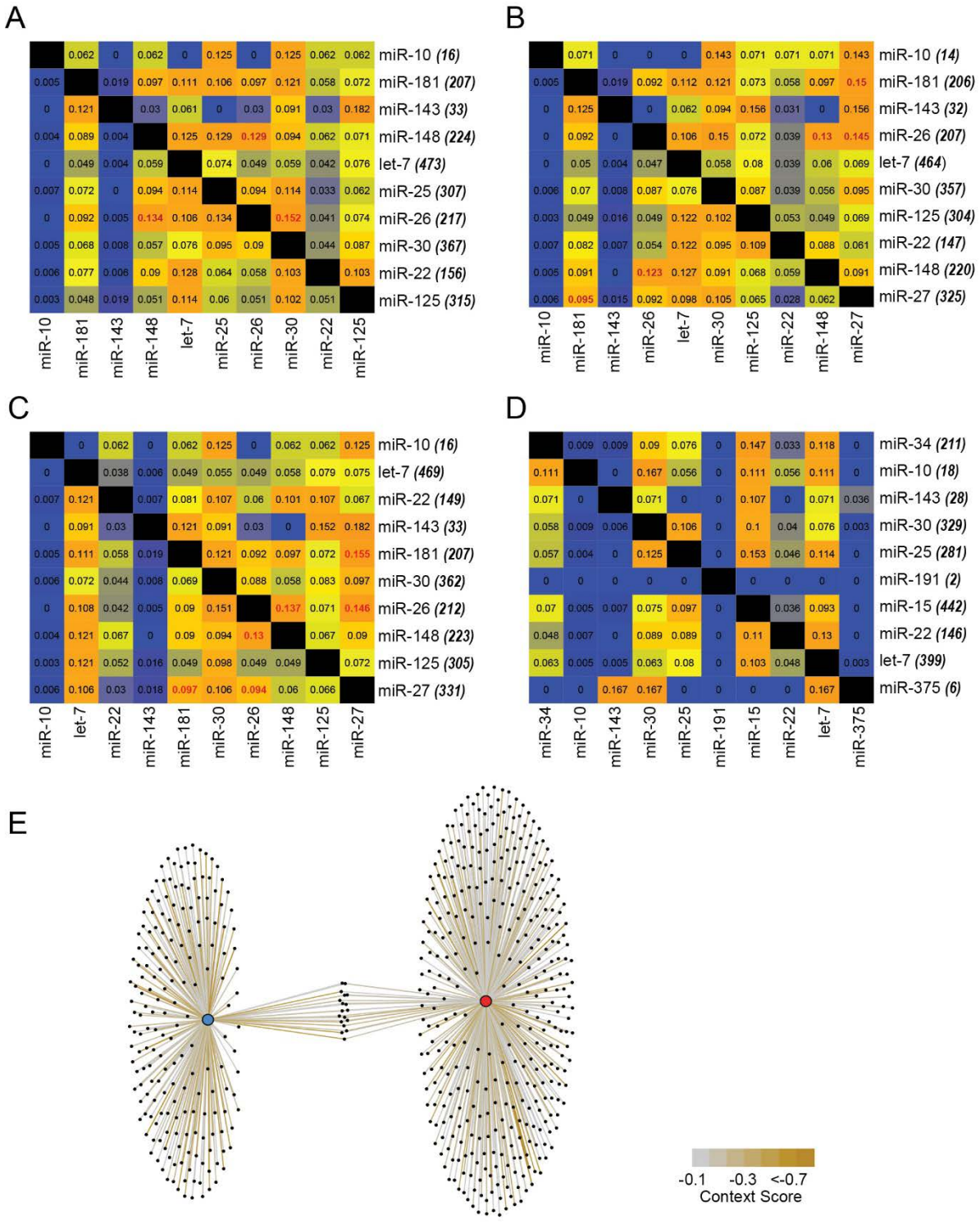


Fig. S6. The miR-34 and miR-29 families regulate a distinct set of targets in spermatocytes.

A-D Values shown denote the proportion of conserved targets (PCT > 0.5) for each miRNA (row) that are also targets of another miRNA family (column) in D1 (A), D3 (B), D7 (C), and LZ (D) samples. miRNA families shown are those within the top ten most abundant for that particular sample. Those with significantly more overlap than expected (Bonferroni corrected p-value < 0.05) appear red. Bonferroni corrected p-values for those with significantly more overlap are as follows – D1: $p = 2.7 \times 10^{-4}$ for targets of miR-148 shared with miR-26, $p = 2.6 \times 10^{-4}$ for targets of miR-26 shared with miR-148, and $p = 4.6 \times 10^{-2}$ for targets of miR-26 shared with miR-30; D3: $p = 1.9 \times 10^{-2}$ for targets of miR-181 shared with miR-27, $p = 1.0 \times 10^{-3}$ for targets of miR-26 shared with miR-148, $p = 4.3 \times 10^{-2}$ for targets of miR-26 shared with miR-27, $p = 1.1 \times 10^{-3}$ for targets of miR-148 shared with miR-26, and $p = 2.3 \times 10^{-2}$ for targets of miR-27 shared with miR-181; D7: $p = 1.3 \times 10^{-2}$ for targets of miR-181 shared with miR-27, $p = 2.5 \times 10^{-4}$ for targets of miR-26 shared with miR-148, $p = 3.9 \times 10^{-2}$ for targets of miR-26 shared with miR-27, $p = 2.6 \times 10^{-4}$ for targets of miR-148 shared with miR-26, $p = 1.6 \times 10^{-2}$ for targets of miR-27 shared with miR-181, and $p = 4.7 \times 10^{-2}$ for targets of miR-27 shared with miR-26.

E Visualization of conserved (PCT > 0.5) targets (black dots) of miR-34 (blue) and miR-29 (red) at P. Targets shared between miR-34 and miR-29 appear as black dots connected to both the red and blue dots; all other targets are unique targets of either miR-34 and miR-29. The predicted strength of the targeting interaction, determined using the TargetScan context+ score, is indicated by the color of the edge between each miRNA and its target; golden edges are predicted to be stronger targets, and greyer edges are predicted to be weaker targets.

Supplemental Tables

miRNA family name	*seed*
let-7	GAGGUAG
miR-10	ACCCUGU
miR-10-iso	CCCUGUA
miR-125	CCCUGAG
miR-126	CGUACCG
miR-127	CGGAUCC
miR-128	CACAGUG
miR-143	GAGAUGA
miR-145	UCCAGUU
miR-148	CAGUGCA
miR-15	AGCAGCA
miR-17	AAAGUGC
miR-181	ACAUUCA
miR-183	AUGGCAC
miR-19	GUGCAAA
miR-191	AACGGAA
miR-204	UCCCUUU
miR-22	AGCUGCC
miR-23	UCACAUU
miR-25	AUUGCAC
miR-26	UCAAGUA
miR-30	GUAAACA
miR-34	GGCAGUG
miR-375	UUGUUCG
miR-470	UCUUGGA
miR-871	GACUGGC
miR-99	ACCCGUA

Table SI. List of miRNA family names used in this manuscript and the corresponding miRNA seed sequence.

	Upregulated		*Downregulated*	
D1 to LZ	spermatogenesis	3.2E-39	regulation of multicellular organismal process	6.1E-45
	male gamete generation	6.0E-39	response to stimulus	1.5E-38
	sexual reproduction	2.1E-30	biological regulation	1.0E-37
	cilium morphogenesis	6.1E-30	single-organism process	2.7E-37
	cilium organization	3.5E-29	response to chemical	4.1E-37

D3 to LZ	spermatogenesis	4.8E-40	regulation of multicellular organismal process	2.0E-40
	male gamete generation	9.1E-40	cellular process	2.5E-38
	sexual reproduction	6.1E-32	single-organism process	9.4E-37
	cilium morphogenesis	1.6E-29	response to stimulus	3.2E-35
	gamete generation	2.0E-28	single-organism cellular process	3.0E-34
D3 to P	spermatogenesis	5.1E-40	regulation of multicellular organismal process	1.7E-61
	male gamete generation	9.8E-40	signaling	3.3E-51
	sexual reproduction	3.2E-30	single-organism process	4.4E-51
	gamete generation	1.1E-28	single organism signaling	4.5E-51
	cilium organization	2.3E-28	cell communication	1.3E-50
D7 to P	spermatogenesis	2.6E-36	regulation of multicellular organismal process	9.8E-50
	male gamete generation	4.8E-36	system development	9.2E-44
	cilium organization	6.2E-31	single-organism process	6.8E-43
	cilium morphogenesis	6.9E-31	single-multicellular organism process	1.4E-42
	cilium assembly	8.3E-29	cell communication	1.8E-42

Table SII. For each of the four supplemental comparisons (D1 to LZ, D3 to LZ, D3 to P, and D7 to P), the top five enriched biological processes identified using gene ontology analysis are shown for genes significantly upregulated or downregulated between samples.

Premeiotic vs Meiotic GO Analysis By Cluster (k = 7)

Cluster 1	ubiquitin-dependent protein catabolic process	1.4E-28
	modification-dependent protein catabolic process	4.2E-28
	modification-dependent macromolecule catabolic process	4.8E-28
	proteolysis involved in cellular protein catabolic process	1.5E-27
	macromolecule catabolic process	1.9E-27
Cluster 2	cell surface receptor signaling pathway	4.2E-09
	positive regulation of response to stimulus	1.6E-06
	single organism signaling	4.1E-06
	immune system process	4.9E-06
	signaling	6.9E-06
Cluster3	sensory perception of chemical stimulus	2.1E-02
	sensory perception of smell	3.9E-02
	detection of stimulus involved in sensory perception	6.8E-01
	sperm motility	6.8E-01
	protein desumoylation	8.7E-01
Cluster 4	cilium organization	1.8E-16
	cilium assembly	3.2E-16
	cilium morphogenesis	6.2E-16
	cellular component assembly involved in morphogenesis	1.5E-09
	cilium movement	7.6E-08
Cluster 5	cellular process	6.0E-33
	cellular component organization or biogenesis	3.4E-22
	cellular component organization	1.7E-19

	regulation of biological process	2.3E-19
	biological regulation	1.2E-18
Cluster 6	regulation of multicellular organismal process	3.8E-10
	defense response	1.4E-09
	G-protein coupled receptor signaling pathway	8.5E-09
	inflammatory response	1.5E-07
	single-multicellular organism process	1.6E-07
Cluster 7	sexual reproduction	5.1E-08
	spermatogenesis	1.0E-07
	male gamete generation	1.0E-07
	multi-organism reproductive process	3.6E-07
	gamete generation	1.9E-06

Table SIII. Premeiotic (D1, D3, and D7) and meiotic (LZ and P) samples were grouped together, and differential expression analysis was performed to identify differentially expressed genes. These differentially expressed genes were then clustered based on their expression, and gene ontology analysis was performed on each cluster.

Tables SIV-IX. For each of the six comparisons (D1 to D3, D3 to D7, D7 to LZ, LZ to P, D1 to D7, and D1 to P), the tables provide information about and the results of the targeting analysis for each miRNA family. These tables include the sequence for each miRNA family seed, the name of each miRNA family, whether that family was differentially abundant (DA) or not (NotDA) in the given comparison, the number of expressed targets of that miRNA family, as well as the number of target for the background list (BG), the Wilcoxon rank sum test p-value, the Bonferroni-corrected Wilcoxon rank sum test p-value, the average $\log_2$ fold change (FC) of the targets of as well as the background list for that miRNA family, the maximum level of expression for the miRNA family in either of the two samples compared, and the $\log_2$ fold change in that miRNA family in the second sample as compared to the first.

Seed	*Family*	*Status*	*# of Targets*	*# of BG Targets*	*P-value*	*Corr. p-value*	*Avg. log FC Targets*	*Avg. log FC BG Targets*	*Max miRNA CPM*	*miRNA log2 FC*
CAGCAGG	miR-214	NotDA	1953	2773	1.15E-04	4.27E-03	-0.01	-0.04	598	0.11

UGUGCUU	miR-218	NotDA	1022	3124	1.92E-04	7.11E-03	-0.07	-0.03	428	0.60
AUUGCAC	miR-25	NotDA	706	3235	2.94E-04	1.09E-02	-0.07	-0.03	41346	-1.02
AGUGCAA	miR-130	NotDA	917	3108	7.45E-04	2.76E-02	-0.06	-0.03	4201	0.09
UCACAGU	miR-27	NotDA	1300	3002	7.88E-04	2.91E-02	-0.06	-0.03	29105	0.87
UUGGCAA	miR-182	NotDA	977	3129	8.62E-04	3.19E-02	-0.07	-0.03	6137	-0.97
ACAGUAC	miR-101	NotDA	977	3131	1.86E-03	6.90E-02	-0.06	-0.03	493	-0.12
ACAUUCA	miR-181	NotDA	803	3194	2.49E-03	9.22E-02	-0.06	-0.04	83263	-0.09
GGCUCAG	miR-24	NotDA	1698	2889	3.18E-03	1.18E-01	-0.01	-0.04	602	2.13
GUAAACA	miR-30	NotDA	869	3169	4.59E-03	1.70E-01	-0.06	-0.03	52056	0.96
AUGGCAC	miR-183	NotDA	1011	3124	6.79E-03	2.51E-01	-0.06	-0.03	397	0.70
UCAAGUA	miR-26	NotDA	660	3264	7.20E-03	2.67E-01	-0.06	-0.03	58103	1.11
ACCCGUA	miR-99	DA	194	3464	9.01E-03	3.33E-01	-0.10	-0.04	75951	1.69
CGUACCG	miR-126	NotDA	81	3500	1.17E-02	4.32E-01	-0.11	-0.04	588	0.44
ACCCUGU	miR-10	NotDA	1218	3041	1.70E-02	6.30E-01	-0.01	-0.04	474287	0.01
CCCUGAG	miR-125	NotDA	1511	2973	1.70E-02	6.30E-01	-0.02	-0.04	51466	1.13
GCUACAU	miR-221	NotDA	987	3104	1.77E-02	6.56E-01	-0.05	-0.03	256	-0.45
AACGGAA	miR-191	NotDA	319	3392	2.13E-02	7.87E-01	-0.07	-0.04	18625	1.60
UGUGCGU	miR-210	NotDA	305	3412	2.20E-02	8.13E-01	-0.08	-0.04	266	-0.07
GUGCAAA	miR-19	NotDA	953	3151	3.93E-02	1.45E+00	-0.05	-0.03	810	0.21
AAAGUGC	miR-17	NotDA	765	3201	6.39E-02	2.36E+00	-0.05	-0.04	2641	0.31
CCAGUGU	miR-199	DA	1330	2996	8.91E-02	3.30E+00	-0.02	-0.04	5348	2.11
AGCUUAU	miR-21	NotDA	517	3325	1.32E-01	4.88E+00	-0.05	-0.04	5694	0.25
GAGGUAG	let-7	NotDA	979	3133	1.41E-01	5.21E+00	-0.02	-0.04	52501	-0.01
UCACAUU	miR-23	NotDA	809	3181	1.72E-01	6.35E+00	-0.05	-0.04	751	1.30
GAGAUGA	miR-143	NotDA	704	3249	1.88E-01	6.97E+00	-0.02	-0.04	68469	-0.13
GAGAACU	miR-146	NotDA	767	3226	2.28E-01	8.42E+00	-0.04	-0.04	756	1.42
AACCGUU	miR-451	NotDA	244	3436	2.60E-01	9.64E+00	-0.04	-0.04	353	1.22
AAUACUG	miR-200	NotDA	622	3258	2.85E-01	1.05E+01	-0.04	-0.04	372	0.03
UCCAGUU	miR-145	NotDA	1172	3058	2.89E-01	1.07E+01	-0.02	-0.04	581	0.88
AGGUAGU	miR-196	NotDA	632	3268	2.93E-01	1.08E+01	-0.05	-0.04	636	0.20
UGACCUA	miR-192	NotDA	697	3251	3.01E-01	1.11E+01	-0.02	-0.04	1302	0.08
GCAGCAU	miR-103	NotDA	1415	2969	3.33E-01	1.23E+01	-0.04	-0.04	3730	-0.45
CAGUGCA	miR-148	NotDA	1368	2940	3.56E-01	1.32E+01	-0.03	-0.03	53489	-0.24
GGCAGUG	miR-34	NotDA	1637	2891	3.85E-01	1.42E+01	-0.03	-0.03	647	0.89
AGCUGCC	miR-22	NotDA	1425	2938	3.94E-01	1.46E+01	-0.03	-0.03	48463	0.88
AGCAGCA	miR-15	NotDA	1853	2766	4.19E-01	1.55E+01	-0.03	-0.03	14632	0.93

Table SIV – D1 to D3

Seed	*Family*	*Status*	*# of Targets*	*# of BG Targets*	*P-value*	*Corr. p-value*	*Avg. log FC Targets*	*Avg. log FC BG Targets*	*Max miRNA CPM*	*miRNA log2 FC*
GUAAACA	miR-30	NotDA	866	3383	2.01E-08	7.85E-07	0.11	0.01	60878	0.23
AAAGUGC	miR-17	NotDA	762	3403	2.38E-07	9.30E-06	0.10	0.01	3762	0.51
ACAGUAC	miR-101	NotDA	979	3306	7.88E-05	3.07E-03	0.08	0.01	774	0.77
UCAAGUA	miR-26	NotDA	660	3452	1.16E-04	4.53E-03	0.09	0.02	58103	-0.02
ACAUUCA	miR-181	NotDA	801	3401	1.67E-04	6.52E-03	0.09	0.02	78438	-0.32
GGCAGUG	miR-34	NotDA	1622	3086	3.32E-03	1.30E-01	0.00	0.02	647	-0.10
GUGCAAA	miR-19	NotDA	955	3295	9.03E-03	3.52E-01	0.06	0.02	933	0.20
AACCGUU	miR-451	NotDA	238	3639	1.15E-02	4.49E-01	-0.05	0.02	353	-1.34
CAGCAGG	miR-214	NotDA	1931	2948	1.89E-02	7.36E-01	0.01	0.02	598	-0.73
AGUGCAA	miR-130	NotDA	914	3290	2.60E-02	1.01E+00	0.06	0.02	4201	-0.06
UGUGCGU	miR-210	NotDA	300	3613	3.59E-02	1.40E+00	-0.03	0.02	254	-0.36
GAGAUGA	miR-143	NotDA	695	3436	3.66E-02	1.43E+00	-0.01	0.03	67939	0.11
UGUGCUU	miR-218	NotDA	1012	3304	5.07E-02	1.98E+00	-0.01	0.02	428	-0.86
UGACCUA	miR-192	NotDA	686	3451	5.30E-02	2.07E+00	0.01	0.02	1302	-0.12
ACCCGUA	miR-99	NotDA	194	3640	5.46E-02	2.13E+00	0.10	0.02	75951	-1.27
UCACAUU	miR-23	NotDA	796	3407	5.59E-02	2.18E+00	0.06	0.02	751	-0.22
GGCUCAG	miR-24	NotDA	1687	3013	6.78E-02	2.64E+00	0.00	0.02	602	-0.68
UCCCUUU	miR-204	NotDA	1067	3275	8.77E-02	3.42E+00	0.01	0.03	317	1.16
ACCCUGU	miR-10	NotDA	1210	3218	8.89E-02	3.47E+00	0.00	0.02	474287	-0.66
CAGUGCA	miR-148	NotDA	1364	3140	1.35E-01	5.27E+00	0.03	0.02	51723	0.19
AGCAGCA	miR-15	NotDA	1842	2923	1.56E-01	6.09E+00	0.03	0.02	28399	0.96
AGCUUAU	miR-21	NotDA	519	3524	2.07E-01	8.08E+00	0.03	0.02	6205	0.12
UCACAGU	miR-27	NotDA	1292	3156	2.16E-01	8.42E+00	0.03	0.02	33708	0.21
AGCUGCC	miR-22	NotDA	1420	3123	2.48E-01	9.66E+00	0.00	0.02	69738	0.53
GAGAACU	miR-146	NotDA	761	3421	2.73E-01	1.06E+01	0.04	0.02	756	-0.85
CCAGUGU	miR-199	NotDA	1333	3162	2.76E-01	1.08E+01	0.02	0.02	5348	-1.03
UUGGCAA	miR-182	NotDA	976	3308	2.84E-01	1.11E+01	0.02	0.02	4102	0.39
CCCUGAG	miR-125	NotDA	1498	3113	3.18E-01	1.24E+01	0.03	0.01	51466	-0.40
CGUACCG	miR-126	NotDA	81	3706	3.20E-01	1.25E+01	0.06	0.02	588	-0.15
AUGACAC	miR-425	NotDA	716	3419	3.27E-01	1.27E+01	0.03	0.02	257	0.30
CACAGUG	miR-128	NotDA	1351	3115	3.48E-01	1.36E+01	0.02	0.02	409	0.84
AGGUAGU	miR-196	NotDA	636	3473	3.60E-01	1.40E+01	0.05	0.02	636	-0.14
AUUGCAC	miR-25	NotDA	699	3437	3.61E-01	1.41E+01	0.02	0.02	24449	0.26
AACGGAA	miR-191	NotDA	321	3601	4.18E-01	1.63E+01	0.02	0.02	18625	-0.40
GCAGCAU	miR-103	NotDA	1397	3133	4.25E-01	1.66E+01	0.02	0.02	3399	0.32
GAGGUAG	let-7	NotDA	976	3294	4.40E-01	1.72E+01	0.03	0.02	73187	0.49
AUGGCAC	miR-183	NotDA	999	3326	4.50E-01	1.75E+01	0.02	0.02	397	-0.08
UCCAGUU	miR-145	NotDA	1171	3233	4.72E-01	1.84E+01	0.02	0.02	581	-0.23
AAUACUG	miR-200	NotDA	622	3469	4.83E-01	1.88E+01	0.04	0.02	462	0.31

Table SV – D3 to D7

Seed	*Family*	*Status*	*# of Targets*	*# of BG Targets*	*P-value*	*Corr. p-value*	*Avg. log FC Targets*	*Avg. log FC BG Targets*	*Max miRNA CPM*	*miRNA log2 FC*
AGCACCA	miR-29	DA	1361	3621	1.50E-05	6.14E-04	-0.62	-0.31	945	2.43
GGCAGUG	miR-34	DA	1760	3482	1.41E-03	5.80E-02	-0.52	-0.34	167618	8.11
GAGAUGA	miR-143	NotDA	797	3905	3.24E-03	1.33E-01	-0.06	-0.39	81055	0.25
CCCUGAG	miR-125	DA	1641	3526	5.90E-03	2.42E-01	-0.50	-0.34	39052	-4.29
AACCGUU	miR-451	DA	253	4116	1.38E-02	5.66E-01	-0.12	-0.40	363	1.39
ACAGUAC	miR-101	NotDA	1051	3762	2.12E-02	8.70E-01	-0.52	-0.36	774	-1.10
UCACAUU	miR-23	NotDA	887	3825	2.52E-02	1.03E+00	-0.22	-0.40	646	-0.86
GGACGGA	miR-184	DA	458	4022	2.93E-02	1.20E+00	-0.61	-0.38	2383	7.73
UUGGCAA	miR-182	NotDA	1065	3733	2.94E-02	1.20E+00	-0.45	-0.35	4102	-0.46
ACCCUGU	miR-10	DA	1343	3655	3.04E-02	1.24E+00	-0.43	-0.37	299646	-1.57
AACGGAA	miR-191	NotDA	335	4066	3.17E-02	1.30E+00	-0.59	-0.37	29189	1.05
GAGAACU	miR-146	NotDA	839	3879	5.12E-02	2.10E+00	-0.21	-0.40	1433	1.78
AGCUUAU	miR-21	DA	581	3960	5.89E-02	2.42E+00	-0.21	-0.39	6205	-0.88
CCAGUGU	miR-199	DA	1467	3580	7.18E-02	2.94E+00	-0.27	-0.40	2616	-4.44
AAAGUGC	miR-17	DA	827	3846	7.90E-02	3.24E+00	-0.51	-0.38	3762	-1.10
GAGGUAG	let-7	DA	1101	3753	1.15E-01	4.70E+00	-0.25	-0.39	73187	-2.35
UCAAGUA	miR-26	DA	723	3902	1.24E-01	5.10E+00	-0.31	-0.38	57402	-2.41
UUGUUCG	miR-375	DA	211	4129	1.35E-01	5.53E+00	-0.19	-0.38	11706	8.28
AGCAGCA	miR-15	NotDA	2009	3325	1.45E-01	5.93E+00	-0.33	-0.40	28399	-0.71
AUGACAC	miR-425	DA	785	3875	1.52E-01	6.21E+00	-0.38	-0.38	1093	2.08
AAUACUG	miR-200	DA	683	3921	1.69E-01	6.92E+00	-0.47	-0.39	462	-2.46
AUGGCAC	miR-183	NotDA	1083	3758	1.73E-01	7.11E+00	-0.32	-0.41	375	-0.20
CACAGUG	miR-128	NotDA	1495	3540	1.74E-01	7.14E+00	-0.29	-0.39	472	0.20
UCACAGU	miR-27	DA	1427	3586	1.79E-01	7.32E+00	-0.31	-0.39	33708	-2.26
UCCCUUU	miR-204	NotDA	1188	3703	1.85E-01	7.60E+00	-0.24	-0.39	401	0.35
UGACCUA	miR-192	DA	749	3880	2.00E-01	8.20E+00	-0.37	-0.37	6035	2.33
GUGCAAA	miR-19	NotDA	1017	3758	2.19E-01	8.97E+00	-0.46	-0.38	1373	0.56
CUUUGGU	miR-9	DA	750	3885	2.35E-01	9.63E+00	-0.40	-0.37	279	1.89
AGUGCAA	miR-130	DA	970	3768	2.55E-01	1.05E+01	-0.44	-0.39	4026	-1.10
UCCAGUU	miR-145	NotDA	1286	3645	2.61E-01	1.07E+01	-0.38	-0.37	496	-0.07
ACAUUCA	miR-181	DA	867	3831	2.64E-01	1.08E+01	-0.38	-0.38	62874	-3.27
GGCUCAG	miR-24	NotDA	1840	3453	2.85E-01	1.17E+01	-0.32	-0.38	374	-0.50
AUUGCAC	miR-25	NotDA	766	3877	3.59E-01	1.47E+01	-0.40	-0.38	31231	0.35
CAGUGCA	miR-148	DA	1462	3572	3.75E-01	1.54E+01	-0.36	-0.37	51723	-2.26
CGUACCG	miR-126	NotDA	89	4184	3.85E-01	1.58E+01	-0.28	-0.39	530	-0.37

AGGUAGU	miR-196	DA	701	3904	3.95E-01	1.62E+01	-0.30	-0.38	575	-4.82
GUAAACA	miR-30abcdef	NotDA	936	3801	3.95E-01	1.62E+01	-0.35	-0.37	60878	-0.58
CAGCAGG	miR-214	DA	2110	3334	4.22E-01	1.73E+01	-0.33	-0.36	362	-4.48
AGCUGCC	miR-22	DA	1563	3535	4.65E-01	1.91E+01	-0.35	-0.38	69738	-2.02
GCAGCAU	miR-103a	DA	1515	3572	4.75E-01	1.95E+01	-0.36	-0.37	3399	-1.71
ACCCGUA	miR-99ab	DA	214	4130	4.82E-01	1.97E+01	-0.44	-0.39	31531	-3.23

Table SVI – D7 to LZ

Seed	*Family*	*Status*	*# of Targets*	*# of BG Targets*	*P-value*	*Corr. p-value*	*Avg. log FC Targets*	*Avg. log FC BG Targets*	*Max miRNA CPM*	*miRNA log FC*
UCAAGUA	miR-26	NotDA	639	2998	4.51E-01	1.67E+01	-0.02	-0.02	10779	-0.51
GGCUCAG	miR-24	NotDA	1594	2633	4.53E-01	1.68E+01	-0.01	-0.02	259	-0.96
GGCAGUG	miR-34	NotDA	1514	2623	4.74E-02	1.75E+00	-0.06	-0.02	167618	-0.78
CUUUGGU	miR-9	NotDA	648	2978	1.32E-01	4.89E+00	0.03	-0.02	279	-0.42
GAGAACU	miR-146	NotDA	737	2961	2.22E-01	8.21E+00	0.01	-0.02	1433	-0.39
GGACGGA	miR-184	NotDA	385	3074	1.33E-01	4.92E+00	0.00	-0.02	2383	-0.50
AACGGAA	miR-191	NotDA	298	3106	3.82E-01	1.41E+01	-0.02	-0.02	60428	1.05
ACAUUCA	miR-181	NotDA	747	2919	1.76E-02	6.53E-01	0.04	-0.03	6534	-0.05
AGCUGCC	miR-22	NotDA	1356	2711	1.08E-01	3.98E+00	0.01	-0.02	17188	-0.19
AGCAGCA	miR-15	NotDA	1768	2518	2.43E-01	9.01E+00	-0.04	-0.01	17368	-0.43
AGCACCA	miR-29	NotDA	1155	2771	1.60E-01	5.92E+00	0.00	-0.02	945	-0.43
AGUGCAA	miR-130	NotDA	852	2860	5.00E-01	1.85E+01	-0.01	-0.02	1874	-0.71
ACCCGUA	miR-99	NotDA	185	3145	4.29E-01	1.59E+01	-0.02	-0.01	4403	0.39
GAGAUGA	miR-143	NotDA	680	2968	1.58E-02	5.85E-01	-0.09	-0.01	81055	-1.28
AUGACAC	miR-425	NotDA	690	2952	6.56E-02	2.43E+00	0.03	-0.02	1260	0.21
UCACAUU	miR-23	NotDA	787	2940	5.42E-02	2.01E+00	0.04	-0.02	351	-0.56
UGACCUA	miR-192	NotDA	645	2985	1.30E-01	4.80E+00	0.02	-0.02	6035	-0.72
AUUGCAC	miR-25	NotDA	677	2970	9.63E-02	3.56E+00	-0.07	-0.01	32313	0.05
AGCUUAU	miR-21	NotDA	518	3013	3.90E-01	1.44E+01	0.01	-0.02	3372	-0.67
AUGGCAC	miR-183	NotDA	943	2847	4.05E-01	1.50E+01	-0.02	-0.02	323	-0.09
CGUACCG	miR-126	NotDA	80	3207	6.76E-02	2.50E+00	-0.14	-0.01	424	0.06
GCAGCAU	miR-103	NotDA	1324	2700	2.95E-01	1.09E+01	-0.02	-0.02	1151	0.15
GUAAACA	miR-30	NotDA	831	2900	2.42E-02	8.97E-01	0.04	-0.03	40655	-0.23
CCCUGAG	miR-125	NotDA	1413	2691	7.23E-02	2.67E+00	0.01	-0.02	2716	0.44
AAAGUGC	miR-17	NotDA	729	2937	4.59E-01	1.70E+01	-0.02	-0.02	2058	0.23
CACAGUG	miR-128	NotDA	1290	2701	3.36E-01	1.24E+01	-0.02	-0.01	472	-0.43
ACAGUAC	miR-101	NotDA	932	2853	1.11E-01	4.12E+00	0.04	-0.02	357	-1.03
CAGUGCA	miR-148	NotDA	1281	2713	3.43E-01	1.27E+01	0.00	-0.02	10770	-0.74
UCACAGU	miR-27	NotDA	1233	2721	4.41E-01	1.63E+01	-0.01	-0.01	7052	-0.88

GUGCAAA	miR-19	NotDA	878	2855	4.13E-01	1.53E+01	0.00	-0.02	1373	-0.56
ACCCUGU	miR-10	NotDA	1140	2766	2.45E-02	9.07E-01	0.04	-0.02	101266	-0.26
AACCGUU	miR-451	DA	214	3150	1.72E-02	6.37E-01	-0.17	-0.01	363	-4.60
UUGUUCG	miR-375	NotDA	195	3155	5.33E-02	1.97E+00	-0.08	-0.02	11706	-0.11
GAGGUAG	let-7	NotDA	942	2874	1.34E-01	4.94E+00	-0.06	-0.02	14386	-0.53
UCCAGUU	miR-145	NotDA	1120	2796	2.35E-01	8.68E+00	-0.01	-0.02	477	0.02
UCCCUUU	miR-204	NotDA	1018	2818	4.20E-01	1.56E+01	-0.03	-0.02	626	0.63
UUGGCAA	miR-182	NotDA	920	2845	1.28E-01	4.72E+00	0.01	-0.03	2981	-0.42

Table SVII – LZ to P

Seed	*Family*	*Status*	*# of Targets*	*# of BG Targets*	*P-value*	*Corr. p-value*	*Avg. log FC Targets*	*Avg. log FC BG Targets*	*Max miRNA CPM*	*miRNA log2 FC*
AAAGUGC	miR-17	NotDA	782	3528	6.43E-05	2.51E-03	0.04	-0.03	3762	0.82
GUAAACA	miR-30	DA	880	3511	1.55E-04	6.03E-03	0.04	-0.03	60878	1.18
CCCUGAG	miR-125	NotDA	1532	3252	2.01E-03	7.85E-02	0.01	-0.03	39052	0.73
UCAAGUA	miR-26	NotDA	671	3597	3.00E-03	1.17E-01	0.03	-0.03	57402	1.09
UGUGCUU	miR-218	NotDA	1036	3431	3.16E-03	1.23E-01	-0.09	-0.02	282	-0.25
ACAUUCA	miR-181	NotDA	816	3523	4.28E-03	1.67E-01	0.03	-0.03	83263	-0.41
UCCCUUU	miR-204	NotDA	1102	3430	1.03E-02	4.00E-01	-0.06	-0.02	317	0.38
UGUGCGU	miR-210	NotDA	305	3740	1.39E-02	5.43E-01	-0.11	-0.02	266	-0.43
CCAGUGU	miR-199	NotDA	1349	3309	1.62E-02	6.33E-01	0.00	-0.03	2616	1.08
ACAGUAC	miR-101	NotDA	989	3440	2.18E-02	8.50E-01	0.01	-0.02	774	0.65
GUGCAAA	miR-19	NotDA	971	3446	2.37E-02	9.26E-01	0.00	-0.02	933	0.41
CAGCAGG	miR-214	NotDA	1974	3083	2.54E-02	9.91E-01	0.00	-0.02	555	-0.62
CAGUGCA	miR-148	NotDA	1385	3279	3.74E-02	1.46E+00	0.00	-0.03	53489	-0.05
AUUGCAC	miR-25	NotDA	714	3567	3.81E-02	1.48E+00	-0.06	-0.02	41346	-0.76
GGCUCAG	miR-24	NotDA	1721	3145	4.79E-02	1.87E+00	0.00	-0.03	374	1.45
GAGGUAG	let-7	NotDA	991	3450	4.84E-02	1.89E+00	0.01	-0.03	73187	0.48
UUGGCAA	miR-182	NotDA	992	3463	5.37E-02	2.10E+00	-0.06	-0.02	6137	-0.58
AGCAGCA	miR-15	DA	1879	3073	6.14E-02	2.40E+00	-0.01	-0.02	28399	1.89
UCACAUU	miR-23	NotDA	821	3531	7.27E-02	2.84E+00	0.01	-0.02	646	1.08
GAGAACU	miR-146	NotDA	776	3552	7.34E-02	2.86E+00	0.00	-0.03	418	0.57
GCAGCAU	miR-103	NotDA	1428	3253	7.73E-02	3.02E+00	-0.02	-0.02	3730	-0.13
UCCAGUU	miR-145	NotDA	1191	3339	1.01E-01	3.96E+00	-0.01	-0.03	496	0.65
ACCCUGU	miR-10	NotDA	1242	3346	1.53E-01	5.97E+00	-0.01	-0.03	471116	-0.65
AGGUAGU	miR-196	NotDA	647	3598	1.54E-01	6.00E+00	0.00	-0.02	575	0.06
AAUACUG	miR-200	NotDA	633	3599	1.59E-01	6.18E+00	0.00	-0.03	462	0.34
ACCCGUA	miR-99	NotDA	199	3792	1.75E-01	6.81E+00	0.00	-0.03	31531	0.42
GGCAGUG	miR-34	NotDA	1654	3190	1.81E-01	7.07E+00	-0.03	-0.02	605	0.79

CGUACCG	miR-126	NotDA	83	3849	2.32E-01	9.06E+00	-0.05	-0.03	530	0.29
AUGGCAC	miR-183	NotDA	1025	3433	2.76E-01	1.07E+01	-0.04	-0.03	375	0.62
AACGGAA	miR-191	NotDA	325	3744	3.52E-01	1.37E+01	-0.04	-0.03	14137	1.20
GAGAUGA	miR-143	NotDA	715	3592	3.92E-01	1.53E+01	-0.04	-0.02	68469	-0.01
AGCUGCC	miR-22	DA	1453	3262	3.93E-01	1.53E+01	-0.03	-0.02	69738	1.40
UGACCUA	miR-192	NotDA	701	3593	3.95E-01	1.54E+01	-0.02	-0.02	1230	-0.04
AGCUUAU	miR-21	NotDA	528	3662	4.49E-01	1.75E+01	-0.02	-0.02	6205	0.38
CACAGUG	miR-128	NotDA	1382	3298	4.50E-01	1.75E+01	-0.02	-0.02	409	0.73
UCACAGU	miR-27	NotDA	1322	3297	4.57E-01	1.78E+01	-0.03	-0.01	33708	1.09
AGUGCAA	miR-130	NotDA	932	3440	4.58E-01	1.79E+01	-0.01	-0.02	4026	0.03
GCUACAU	miR-221	NotDA	999	3453	4.79E-01	1.87E+01	-0.02	-0.03	256	-0.81
AUGACAC	miR-425	NotDA	728	3578	4.97E-01	1.94E+01	-0.03	-0.02	257	0.80

Table SVIII – D1 to D7

Seed	*Family*	*Status*	*# of Targets*	*# of BG Targets*	*P-value*	*Corr. p-value*	*Avg. log FC Targets*	*Avg. log FC BG Targets*	*Max miRNA CPM*	*miRNA log FC*
GGCAGUG	miR-34	DA	1759	3448	6.87E-05	2.82E-03	-0.61	-0.35	97806	8.13
AGCACCA	miR-29	DA	1366	3587	3.99E-04	1.64E-02	-0.66	-0.38	703	4.42
UCACAUU	miR-23	NotDA	887	3833	9.92E-04	4.07E-02	-0.19	-0.42	306	-0.35
CCAGUGU	miR-199	DA	1455	3549	1.20E-02	4.91E-01	-0.30	-0.44	1242	-4.49
AGCUUAU	miR-21	DA	574	3935	2.11E-02	8.67E-01	-0.23	-0.42	4775	-1.17
AACGGAA	miR-191	DA	337	4027	2.63E-02	1.08E+00	-0.66	-0.41	60428	3.29
GAGAACU	miR-146	DA	842	3824	3.58E-02	1.47E+00	-0.24	-0.41	1093	1.96
GUAAACA	miR-30	NotDA	941	3769	3.86E-02	1.58E+00	-0.28	-0.42	34728	0.37
UCAAGUA	miR-26	DA	715	3875	4.05E-02	1.66E+00	-0.31	-0.43	26970	-1.84
UUGGCAA	miR-182	DA	1064	3701	5.35E-02	2.19E+00	-0.51	-0.40	6137	-1.47
GGACGGA	miR-184	DA	454	3993	8.37E-02	3.43E+00	-0.62	-0.41	1682	6.86
GAGAUGA	miR-143	DA	776	3870	8.53E-02	3.50E+00	-0.22	-0.41	68469	-1.04
CAGCAGG	miR-214	DA	2096	3341	8.59E-02	3.52E+00	-0.37	-0.43	555	-5.74
CCCUGAG	miR-125	DA	1636	3496	8.64E-02	3.54E+00	-0.49	-0.39	23580	-3.12
CAGUGCA	miR-148	DA	1461	3534	1.05E-01	4.31E+00	-0.37	-0.43	53489	-3.05
GAGGUAG	let-7	DA	1085	3705	1.07E-01	4.39E+00	-0.31	-0.41	52501	-2.40
AUUGCAC	miR-25	NotDA	774	3848	1.71E-01	7.00E+00	-0.50	-0.39	41346	-0.36
ACAGUAC	miR-101	DA	1049	3707	1.82E-01	7.47E+00	-0.48	-0.40	493	-1.49
ACAUUCA	miR-181	DA	855	3800	1.85E-01	7.57E+00	-0.34	-0.43	83263	-3.72
AGGUAGU	miR-196	DA	697	3863	2.05E-01	8.41E+00	-0.33	-0.42	554	-4.54
ACCCUGU	miR-10	DA	1321	3606	2.14E-01	8.76E+00	-0.42	-0.39	471116	-2.48
AGCAGCA	miR-15	NotDA	1994	3303	2.25E-01	9.21E+00	-0.38	-0.41	12846	0.74
UCACAGU	miR-27	DA	1418	3556	2.46E-01	1.01E+01	-0.37	-0.41	15880	-2.05

CGUACCG	miR-126	NotDA	86	4136	2.53E-01	1.04E+01	-0.48	-0.41	433	-0.02
AUGACAC	miR-425	DA	788	3835	2.68E-01	1.10E+01	-0.41	-0.41	1260	3.10
CACAGUG	miR-128	NotDA	1485	3528	2.68E-01	1.10E+01	-0.35	-0.42	348	0.51
UGUGCUU	miR-218	DA	1122	3668	2.78E-01	1.14E+01	-0.45	-0.40	282	-4.37
AAAGUGC	miR-17	NotDA	824	3791	2.78E-01	1.14E+01	-0.50	-0.41	2130	-0.05
UGUGCGU	miR-210	DA	336	4033	2.96E-01	1.21E+01	-0.46	-0.41	266	-3.57
AGUGCAA	miR-130	DA	977	3699	3.05E-01	1.25E+01	-0.47	-0.41	3940	-1.78
UUGUUCG	miR-375	DA	210	4092	3.10E-01	1.27E+01	-0.35	-0.41	10834	6.01
UCCCUUU	miR-204	NotDA	1195	3673	3.66E-01	1.50E+01	-0.35	-0.42	626	1.37
GCAGCAU	miR-103	DA	1514	3523	3.85E-01	1.58E+01	-0.41	-0.41	3730	-1.69
AUGGCAC	miR-183	NotDA	1085	3721	3.92E-01	1.61E+01	-0.40	-0.42	303	0.32
GUGCAAA	miR-19	NotDA	1008	3726	3.99E-01	1.64E+01	-0.48	-0.41	936	0.42
AAUACUG	miR-200	DA	675	3874	4.27E-01	1.75E+01	-0.44	-0.40	365	-1.92
GCUACAU	miR-221	DA	1065	3719	4.31E-01	1.77E+01	-0.44	-0.42	256	-2.64
ACCCGUA	miR-99	DA	207	4079	4.34E-01	1.78E+01	-0.50	-0.42	23604	-2.42
AGCUGCC	miR-22	NotDA	1552	3513	4.79E-01	1.96E+01	-0.40	-0.41	26338	-0.81
UGACCUA	miR-192	DA	761	3843	4.81E-01	1.97E+01	-0.38	-0.40	3651	1.57
UCCAGUU	miR-145	NotDA	1278	3653	4.89E-01	2.00E+01	-0.41	-0.41	477	0.60

Table SIX – D1 to P

Gene	*5p or 3p*	*Sequence*	*D1 Average*	*LZ Average*	*P Average*
Mir29a	3p	UAGCACCAUCUGAAAU	0.00	0.61	0.00
Mir29a	3p	UAGCACCAUCUGAAAUCGG	0.35	24.38	39.86
Mir29a	3p	UAGCACCAUCUGAAAUCGGU	2.75	232.48	262.83
Mir29a	3p	UAGCACCAUCUGAAAUCGGUU	6.72	396.97	317.47
Mir29a	3p	UAGCACCAUCUGAAAUCGGUUA	10.17	389.70	348.45
Mir29a	3p	UAGCACCAUCUGAAAUCGGUUAU	0.23	75.63	44.31
Mir29b-1/2	3p	UAGCACCAUUUGAAAUCAG	0.00	0.61	1.48
Mir29b-1/2	3p	UAGCACCAUUUGAAAUCAGU	0.00	1.22	4.43
Mir29b-1/2	3p	UAGCACCAUUUGAAAUCAGUG	0.00	0.00	4.43
Mir29b-1/2	3p	UAGCACCAUUUGAAAUCAGUGU	0.00	4.88	4.43
Mir29b-1/2	3p	UAGCACCAUUUGAAAUCAGUGUU	0.35	44.55	41.32
Mir29c	3p	UAGCACCAUUUGAAAUCGG	0.35	0.00	1.48
Mir29c	3p	UAGCACCAUUUGAAAUCGGU	1.03	14.48	5.91
Mir29c	3p	UAGCACCAUUUGAAAUCGGUU	0.93	28.59	23.61
Mir29c	3p	UAGCACCAUUUGAAAUCGGUUA	3.37	18.21	10.34
Mir29c	3p	UAGCACCAUUUGAAAUCGGUUAU	0.23	6.63	0.00

Table SX. All individual miRNAs and isomiRs of the miR-29 gene family which bear the miR-29 family seed sequence. The gene of origin, hairpin arm, and sequence for each are given. Counts shown are average counts per million (CPM).

		sig. downregulated		*sig. downregulated + sig. PT regulation*	
miR-29 targets	*D1 to LZ*	extracellular matrix component	1.2E-10	proteinaceous extracellular matrix	1.3E-08
		complex of collagen trimers	1.2E-10	extracellular matrix component	1.9E-08
		collagen trimer	1.2E-10	extracellular matrix	1.9E-08
		proteinaceous extracellular matrix	7.0E-09	collagen trimer	8.2E-05
		extracellular matrix	1.5E-08	basement membrane	8.2E-05
	D1 to P	collagen trimer	1.2E-12	extracellular matrix component	1.6E-07
		extracellular matrix component	1.5E-11	proteinaceous extracellular matrix	1.0E-05
		proteinaceous extracellular matrix	5.4E-11	extracellular matrix	1.8E-05
		complex of collagen trimers	1.0E-10	basement membrane	3.8E-04
		extracellular matrix	1.1E-10	complex of collagen trimers	7.3E-04
miR-34 targets	*D1 to LZ*	intrinsic component of plasma membrane	9.2E-02	extracellular membrane-bounded organelle	1.0E+00
		synaptobrevin 2-SNAP-25-syntaxin-1a-complexin II	1.9E-01	extracellular exosome	1.0E+00
		membrane part	2.3E-01	extracellular organelle	1.0E+00
		plasma membrane part	2.6E-01	extracellular vesicle	1.0E+00
		anchored component of plasma membrane	2.8E-01	intrinsic component of plasma membrane	1.0E+00
	D1 to P	organelle outer membrane	4.9E-01	extracellular membrane-bounded organelle	1.6E-01
		outer membrane	4.9E-01	extracellular exosome	1.6E-01
		mitochondrial outer membrane	4.9E-01	extracellular organelle	1.6E-01
		synaptobrevin 2-SNAP-25-syntaxin-1a-complexin II	6.7E-01	extracellular vesicle	1.6E-01
		membrane part	8.0E-01	extracellular region part	1.0E+00

Table SXI. Gene ontology analysis for all conserved (PCT > 0.5) mRNA transcript targets of either the miR-29 or miR-34 family that are significantly downregulated (left) or both significantly downregulated as well as showing a significant signature of post-transcriptional

(PT) regulation (right) for the D1 vs LZ and D1 v P comparisons. Significantly downregulated targets were compared to a background list of all targets with high enough abundance for RNAseq analysis to determine enrichment. Significantly downregulated post-transcriptionally downregulated targets were compared to a background list of all targets with high enough abundance for EISA analysis to determine enrichment.

miR-29		***miR-34***	
D1 vs LZ	*D1 vs P*	*D1 vs LZ*	*D1 vs P*
1700025G04Rik	1700025G04Rik	2510009E07Rik	2510009E07Rik
4931406P16Rik	4931406P16Rik	5730455P16Rik	5730455P16Rik
9530068E07Rik	9530068E07Rik	Acsl4	Acsl4
Adam12	Adam12	Adipor2	Adipor2
Adam19	Adam19	Alcam	Aff4
Adamts9	Adamts9	Ankrd52	Alcam
Agpat4	Agpat4	Anp32a	Ankrd52
Agtr2	Agtr2	Arhgap26	Anp32a
Ahr	Ahr	Arhgap36	Arhgap26
Ak3	Ak3	Arl15	Arhgap36
Akap17b	Akap17b	Asic2	Arl15
Akap5	Akap5	AW549877	Asic2
Amer1	Amer1	Bcl2	AW549877
Amot	Amot	Btbd11	Bcl2
Ankrd13b	Ankrd13b	Capn5	Bnc2
Ankrd52	Ankrd52	Casp2	Btbd11
Ap4e1	Ap4e1	Ccnd1	Capn5
Apc	Apc	Cntnap1	Casp2
Arf2	Arf2	Col12a1	Ccnd1
Arhgap28	Arhgap28	Coro1c	Cntnap1
Arhgap36	Arhgap36	Cplx2	Col12a1
Asxl3	Asxl3	Crhr1	Coro1c
Atad2b	Atad2b	Crkl	Cplx2
Atp1b1	Atp1b1	Crtc1	Crkl
Atrn	Atrn	Csf1r	Crtc1
Bach2	Bach2	Ddx17	Csf1r
Bahd1	Bahd1	Dpysl4	Dcp1a
Bcl2l2	Bcl2l2	E2f3	Ddx17
Bmf	Bmf	Elf4	Dmwd

Bmpr1a	Bmpr1a	Eml5	Dpysl4
C1qtnf6	Brwd3	Ergic1	E2f3
C77370	C1qtnf6	Fam63b	Elf4
Camk1d	C77370	Fat3	Eml5
Capn7	Camk1d	Fbln5	Erc1
Cav2	Capn7	Fbxo10	Ergic1
Cbx2	Cav2	Fbxo41	Fat3
Cbx6	Cbx2	Fndc3b	Fbln5
Ccdc28b	Cbx6	Foxj2	Fbxo10
Ccnd2	Ccdc28b	Foxn2	Fbxo41
Ccnj	Ccnd2	Frmd4a	Fndc3b
Ccnl2	Ccnj	Gabra3	Foxj2
Cd276	Ccnl2	Gas1	Foxn2
Cdc42bpa	Cd276	Glce	Frmd4a
Cenpb	Cdc42bpa	Gnai2	Gabra3
Cep68	Cenpb	Hapln1	Gas1
Cep97	Cep68	Hcn3	Glce
Cldn1	Cldn1	Hecw2	Gnai2
Clec2l	Clec2l	Hook3	Grm7
Cmpk1	Cmpk1	Igfbp3	Hcn3
Cnot6	Col11a1	Ing5	Hecw2
Col11a1	Col15a1	Inhbb	Hook3
Col15a1	Col1a2	Kalrn	Igfbp3
Col1a2	Col24a1	Kcnd3	Inhbb
Col27a1	Col27a1	Kcnk3	Jag1
Col3a1	Col3a1	Kitl	Kalrn
Col4a1	Col4a1	Lgr4	Kcnd3
Col4a2	Col4a2	Lima1	Kcnk3
Col4a3	Col4a3	Met	Kitl
Col4a4	Col4a4	Mllt3	Lgr4
Col4a5	Col4a5	Mpp2	Lima1
Col4a6	Col4a6	Mras	Met
Col5a1	Col5a1	Mta2	Mex3c
Col5a2	Col5a2	Myh9	Mgat4a
Col5a3	Col5a3	Myrip	Mllt3
Col6a2	Col6a2	Nampt	Mpp2
Col6a3	Col6a3	Nav1	Mras
Col7a1	Col7a1	Nav3	Mta2
Creb5	Col8a1	Nceh1	Myh9
Crispld1	Creb5	Ndst1	Myrip
D430041D05Rik	Crispld1	Nfe2l1	Nampt
D630045J12Rik	D430041D05Rik	Notch1	Nav1
Dbt	D630045J12Rik	Notch2	Nav3

Ddx3x	Dbt	Nrn1	Nceh1
Ddx3y	Ddx3x	Nrxn2	Ndst1
Dennd6a	Ddx3y	Nsd1	Nfe2l1
Dicer1	Dennd6a	Nsmce4a	Notch1
Dip2c	Dicer1	Numbl	Notch2
Dpp4	Dip2c	Pdgfra	Nrn1
Dpysl2	Dpp4	Pea15a	Nrxn2
Drp2	Dpysl2	Phf19	Nsd1
Dtx4	Drp2	Pitpnc1	Nsmce4a
E2f7	Dtx4	Pkia	Numbl
Edc3	Dusp2	Plekhg3	Pdgfra
Eln	Eln	Plod1	Pea15a
Elovl4	Elovl4	Ppp2r3a	Phf19
Eml4	Eml4	Prex2	Pitpnc1
Eml5	Eml5	Prkacb	Pkia
Enho	Enho	Prkd1	Plekhg3
Entpd7	Entpd7	Ptprm	Plod1
Epha1	Epha1	Rab43	Ppp2r3a
Ercc6	Erc1	Rarb	Prex2
Etv4	Ercc6	Ret	Prkacb
Etv6	Etv4	Rnf165	Prkd1
F11r	Etv6	Rragd	Rarb
Fam134c	F11r	Rras2	Reln
Fam184b	Fam134c	Rtn4rl1	Ret
Fam193b	Fam184b	Rufy2	Rnf165
Fam57b	Fam193b	Sash1	Rragd
Fam76b	Fam57b	Satb2	Rras
Fbn1	Fam76b	Scn2b	Rras2
Fbn2	Fbn1	Sema4b	Rtf1
Fbxl20	Fbn2	Sema4c	Rtn4rl1
Fbxo42	Fbxl20	Shank3	Sash1
Fchsd1	Fchsd1	Slc25a27	Scn2b
Foxj2	Foxj2	Slc27a4	Sema4b
Fras1	Fras1	Slc2a13	Sema4c
Frem1	Frem1	Slc44a2	Shank3
Frmd4a	Frmd4a	Slc6a17	Slc25a27
Fstl1	Fstl1	Smim15	Slc27a4
Fyn	Fyn	Snx30	Slc2a13
Glis2	Gfod1	Snx4	Slc44a2
Gm17296	Glis2	Stac2	Slc6a17
Golga7	Gm17296	Stat6	Smim15
Gpr37	Gpr37	Syvn1	Snx30
Gpx7	Gpx7	Tfrc	Srpr

Hip1	Hdac4	Tgif2	Stac2
Hmcn1	Hip1	Thumpd1	Stat6
Hook3	Hmcn1	Tm9sf3	Stk38l
Hs3st3b1	Hook3	Tmcc3	Tanc2
Hspg2	Hs3st3b1	Tmed8	Tgif2
Iffo2	Hspg2	Tmem109	Tm9sf3
Ifi30	Iffo2	Tmem167b	Tmcc3
Il1rap	Ifi30	Tmem255a	Tmed8
Impdh1	Il1rap	Tmem55a	Tmem109
Ina	Impdh1	Tob2	Tmem55a
Ing4	Ina	Trank1	Trank1
Insig1	Ing4	Trim67	Trim67
Insrr	Insig1	Txn1	Txn1
Irs1	Insrr	Ube2g1	Ube2g1
Itgb1	Ireb2	Uhrf2	Uhrf2
Jarid2	Irs1	Usf1	Usf1
Kctd15	Itga11	Usp31	Vamp2
Kctd3	Itgb1	Vamp2	Vat1
Kctd5	Jarid2	Vcl	Vcl
Kdelc1	Jmy	Wasf1	Wasf1
Kdm2a	Kctd15	Wtap	Wscd2
Kdm5c	Kctd20	Yy1	Wtap
Kdm6b	Kctd5	Zc3h4	Zc3h4
Kif26a	Kdelc1	Zdhhc16	Zdhhc16
Kif3c	Kdm5c	Zfp282	Zfp282
Klhl9	Kdm6b	Zfp553	Zfp553
Lamc1	Kif26a		
Ldlrad3	Kif26b		
Ldlrap1	Kif3c		
Lox	Lamc1		
Lpl	Ldlrad3		
Lrp6	Ldlrap1		
Lsm11	Lox		
Ltbr	Lpl		
Mafb	Lrp6		
Map2k6	Lsm11		
Mapre2	Ltbr		
Mast4	Mafb		
Mblac2	Map2k6		
Med12l	Mapre2		
Megf6	Mast4		
Mga	Med12l		
Morf4l2	Megf6		

Mtss1l	Mex3b			
Mycn	Mga			
N4bp2l1	Morf4l2			
Narf	Mtss1l			
Nav1	Mycn			
Nckap5l	N4bp2l1			
Nfatc3	Naa60			
Nid2	Narf			
Nkrf	Nav1			
Nova1	Nckap5l			
Nsd1	Nfatc3			
Nudt11	Nid2			
Orai3	Nkrf			
Oxct1	Nsd1			
Parg	Nudt11			
Pcgf3	Orai3			
Pcsk5	Oxct1			
Pdgfa	Palm2			
Pdgfb	Pcgf3			
Per3	Pcsk5			
Pi15	Pdgfa			
Pik3r1	Pdgfb			
Plod1	Pgap2			
Plp1	Pi15			
Plxna1	Pik3r1			
Pmp22	Plod1			
Ppard	Plp1			
Ppic	Plxna1			
Ppp1r15b	Pmp22			
Ppp1r3d	Ppard			
Prkra	Ppic			
Proser1	Ppm1e			
Ptprd	Ppp1r15b			
Ptx3	Ppp1r3d			
Purg	Prelp			
Pxdn	Prkra			
R3hdm4	Proser1			
Rab30	Ptprd			
Rab40c	Ptx3			
Rapgefl1	Purg			
Rarb	Pxdn			
Reps2	R3hdm4			
Rere	Rab30			

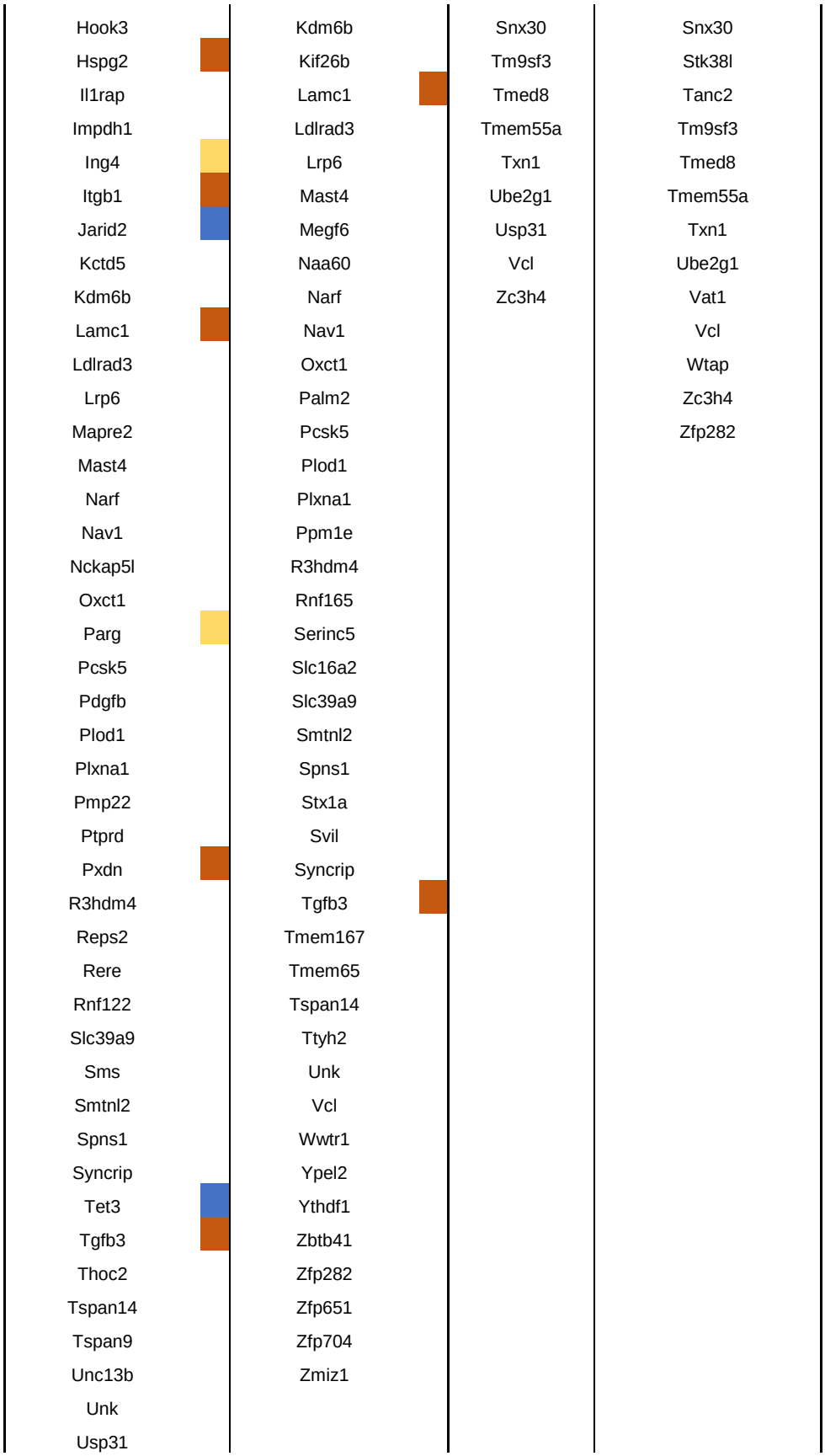
Rest	Rapgefl1			
Rlf	Rarb			
Rnf122	Reps2			
Rnf150	Rere			
Rnf165	Rest			
Scamp5	Rev3l			
Schip1	Rlf			
Serinc5	Rlim			
Serpinh1	Rnf122			
Sestd1	Rnf150			
Shroom2	Rnf165			
Slc16a1	Scamp5			
Slc16a2	Schip1			
Slc31a1	Serinc5			
Slc39a9	Serpinh1			
Slc7a1	Sestd1			
Sms	Shroom2			
Smtnl2	Slc16a2			
Snx24	Slc39a9			
Snx4	Slc7a1			
Sparc	Sms			
Spns1	Smtnl2			
Spry1	Snx24			
Spsb4	Sparc			
Ss18l1	Spns1			
Stmn2	Spry1			
Stx1a	Ss18l1			
Svil	Stmn2			
Syn3	Stx1a			
Syncrip	Svil			
Syt9	Syn3			
Tbc1d13	Syncrip			
Tet2	Syt9			
Tet3	Tet2			
Tgfb3	Tet3			
Thoc2	Tgfb3			
Tll1	Thoc2			
Tmem167	Tll1			
Tmem33	Tmem167			
Tmem65	Tmem65			
Tmem86a	Tmem86a			
Tmod3	Tmod3			
Tmtc3	Tmtc3			

Tnfrsf1a	Tnfrsf1a		
Traf4	Tnrc18		
Trib2	Traf4		
Trove2	Trib2		
Tspan14	Trove2		
Tspan4	Tspan14		
Tspan9	Tspan4		
Ttc14	Tspan9		
Ttyh2	Ttc14		
Tubb2b	Ttyh2		
U2surp	Tubb2b		
Ubfd1	U2surp		
Ubtf	Ubfd1		
Unc13b	Ubtf		
Unk	Unc13b		
Upk1b	Unk		
Urm1	Upk1b		
Usp31	Vcl		
Vamp3	Wwtr1		
Vcl	Xkr4		
Wwtr1	Ypel2		
Xkr4	Ythdf1		
Ypel2	Zbtb41		
Ythdf1	Zc4h2		
Yy1	Zfp282		
Zbtb41	Zfp362		
Zc4h2	Zfp366		
Zdhhc21	Zfp36l1		
Zfp113	Zfp651		
Zfp282	Zfp704		
Zfp362	Zmiz1		
Zfp366			
Zfp36l1			
Zfp651			
Zfp704			
Zmiz1			

Table SXII. Names of all mRNA transcript targets of either the miR-29 or miR-34 family that are significantly downregulated for the D1 vs LZ and D1 v P comparisons.

miR-29	***miR-34***

D1 vs LZ	*D1 vs P*	*D1 vs LZ*	*D1 vs P*
1700025G04Rik	9530068E07Rik	Alcam	5730455P16Rik
9530068E07Rik	Adam12	Anp32a	Alcam
Adam12	Agpat4	Arhgap26	Anp32a
Ak3	Ak3	Arl15	Arhgap26
Arhgap28	Apc	Btbd11	Arl15
Atad2b	Arhgap28	Col12a1	Bnc2
Atrn	Atrn	Crkl	Btbd11
Camk1d	Ccdc28b	Ergic1	Col12a1
Ccdc28b	Col15a1	Fam63b	Coro1c
Cep97	Col1a2	Fat3	Crkl
Col15a1	Col3a1	Fbln5	Erc1
Col1a2	Col4a4	Fbxo10	Ergic1
Col27a1	Col4a5	Fndc3b	Fat3
Col3a1	Col5a1	Frmd4a	Fbln5
Col4a4	Col5a2	Gnai2	Fbxo10
Col4a5	Dbt	Hecw2	Fndc3b
Col5a1	Dennd6a	Hook3	Frmd4a
Col6a2	Dip2c	Kalrn	Gnai2
Col6a3	Dtx4	Kcnd3	Hecw2
Creb5	Eln	Kcnk3	Hook3
D430041D05Rik	Eml4	Lgr4	Kalrn
Dbt	Erc1	Mllt3	Lgr4
Ddx3x	Etv6	Mras	Mgat4a
Dennd6a	Fam134c	Myh9	Mllt3
Dip2c	Fam184b	Nav1	Mras
Dtx4	Fam193b	Ndst1	Myh9
Eln	Fam57b	Notch2	Nampt
Eml4	Fbn1	Nrxn2	Nav1
Etv6	Fbn2	Numbl	Notch2
Fam184b	Fbxl20	Pitpnc1	Nrxn2
Fam193b	Fras1	Plod1	Plod1
Fam57b	Frem1	Ppp2r3a	Ppp2r3a
Fbn1	Frmd4a	Prex2	Prkacb
Fbn2	Fstl1	Prkacb	Prkd1
Fbxl20	Gm17296	Prkd1	Reln
Fras1	Hip1	Ptprm	Rnf165
Frem1	Hook3	Rab43	Rras
Frmd4a	Hspg2	Rtn4rl1	Rtf1
Fstl1	Ing4	Sash1	Rtn4rl1
Gm17296	Itgb1	Shank3	Sash1
Hip1	Jmy	Slc25a27	Slc25a27
Hmcn1	Kctd20	Slc2a13	Slc2a13


Hook3
Hspg2
Il1rap
Impdh1
Ing4
Itgb1
Jarid2
Kctd5
Kdm6b
Lamc1
Ldlrad3
Lrp6
Mapre2
Mast4
Narf
Nav1
Nckap5l
Oxct1
Parg
Pcsk5
Pdgfb
Plod1
Plxna1
Pmp22
Ptprd
Pxdn
R3hdm4
Reps2
Rere
Rnf122
Slc39a9
Sms
Smtnl2
Spns1
Syncrip
Tet3
Tgfb3
Thoc2
Tspan14
Tspan9
Unc13b
Unk
Usp31
Kdm6b
Kif26b
Lamc1
Ldlrad3
Lrp6
Mast4
Megf6
Naa60
Narf
Nav1
Oxct1
Palm2
Pcsk5
Plod1
Plxna1
Ppm1e
R3hdm4
Rnf165
Serinc5
Slc16a2
Slc39a9
Smtnl2
Spns1
Stx1a
Svil
Syncrip
Tgfb3
Tmem167
Tmem65
Tspan14
Ttyh2
Unk
Vcl
Wwtr1
Ypel2
Ythdf1
Zbtb41
Zfp282
Zfp651
Zfp704
Zmiz1
Snx30
Tm9sf3
Tmed8
Tmem55a
Txn1
Ube2g1
Usp31
Vcl
Zc3h4
Snx30
Stk38l
Tanc2
Tm9sf3
Tmed8
Tmem55a
Txn1
Ube2g1
Vat1
Vcl
Wtap
Zc3h4
Zfp282

Vcl			
Wwtr1			
Ypel2			
Ythdf1			
Zbtb41			
Zfp362			
Zfp651			
Zfp704			
Zmiz1			

Table SXIII. Names of all mRNA transcript targets of either the miR-29 or miR-34 family that both significantly downregulated as well as showing a significant signature of post-transcriptional (PT) regulation for the D1 vs LZ and D1 v P comparisons. For miR-29 targets, an orange box appears to the right of those genes that are components of the extracellular matrix (GO:0005578), a yellow box to the right of those involved in the biological process cellular response to DNA damage (GO:0006974), and a blue box next to those involved in the biological process of methylation (GO:0032259).