SUPPLEMENTAL MATERIALS AND METHODS:

Linker ligation

In order to determine the precise nature of the small RNAs and verify that the miRNAs are generated by an endonucleolytic processing event rather than non-specific degradation, total RNA, or small RNA isolated using the miRvana kit (Ambion) was ligated on the 3' end to Linker1 (IDT) using T4 RNA ligase 2 truncated (NEB) as per the manufacturer's instructions. The ligated RNA was analyzed by RT-PCR using a combination of adapter (3' linker) and miRNA-specific primers (miR-1225 Forward, miR-1228 forward and miR-877 forward). Products were subsequently cloned and sequenced. This strategy only allows detection of molecules that have a 3' hydroxyl, like known miRNAs. Using this methodology we verified the predicted miR-877 mature miRNA (Supplemental Figure 2C), both mature miR-1228 and mature miR-1228* (Supplemental Figure 2B), and several species of mature miR-1225-5p with two to four guanines on the 3' end (Supplemental Figure 2A). Although the versions of miR-1225-5p previous sequenced from humans had six to seven guanines on the 3' end, the shorter miR-1225 sequence was previously detected in macaque (40).This data indicates 3' heterogeneity of miR-1225.

Sequencing

For mRNA transcripts, RT-PCR amplicons were sequenced directly using gene specific primers (See Supplemental Table 1). For miRNAs, stemloop RT-PCR products were inserted into the pGem-T Easy Vector (Promega) and sequenced using a T7 promoter primer. All sequencing was preformed at the Northwestern University Genomics Core Facility.

In vitro Transcription

Wild-type and Δss plasmids were used as templates in PCR reactions. The forward primers were T7ABCF1, T7LRP1, T7PKD1, T7DHX30 that are specific to the 5' end of the host gene and contain a T7 promoter sequence. The reverse primers were targeted to the exon following the mirtron/simtron (Gene specific reverse) or to the 3' end of the predominate miRNA species (miRNA specific reverse) in order to generate substrates that should be recognized by gene specific amplification or stemloop RT-PCR, respectively. The PCR products were used as templates for *in vitro* transcription with a T7 RNA polymerase (Promega) to make RNA transcripts. Transcription reactions were carried out in the presence of ^{32}P-UTP and 7Me-GpppG cap analog (NEB) to make pre-mRNA. RNA was purified by gel electrophoresis (80).The resulting synthetic RNA was reverse transcribed as described above and subjected to radiolabelled RT-PCR or stemloop RT-PCR.

qRT-PCR

Total RNA collected from HEK-293T cells transiently transfected with wild-type, Δss, or empty vector control (-) was reverse transcribed using the Taqman microRNA reverse transcription kit (Applied Biosystems) and Taqman microRNA assay kit (Applied

Biosystems PN4427975: miR-1225 TM2764, miR-1228 TM002919 and snoRNA48 TM001006). qPCR was carried out with Taqman universal master mix II, no UNG (Applied Biosystems) on an Applied Biosystems (ABI) 7500 Real-Time PCR System using the ABI 7500 detection software. For each experiment, samples were analyzed in triplicate. Values greater than +/- one standard deviation within the triplicate measurements were not considered.

Construction of plamids

For construction of the miRNAs expressed in an intergenic context, pcDNA-1228, the miRNA containing introns were PCR amplified with Phire polymerase (NEB) using the previously constructed minigenes as templates and primers with restriction sites incorporated on the termini. PCR products were digested with BamHI and EcoR1 and inserted into the similarly digested pcDNA3.1+ plasmid (Invitrogen) using T4 DNA ligase.

For pmiRGLO-luciferase reporter plasmid construction, oligonucleotides miR-1228 Target 3' UTR and miR-1228 Target 3' UTR R were annealed and inserted into an XbaI- and XhoI-digested pmiRGLO plasmid (Promega) as per manufacturer's instructions.

Immunoprecipitation

Immunoprecipitation was performed as described in the methods of the main text

Luciferase assays

Assays were performed as described in the methods of the main text. Where HEK-293T cells were transfected with 0.4 µg, 1 µg or 1.5 µg of minigene plasmid and 10 ng of pmiRGLO (Promega) using Optifect (Invitrogen). miR-877 was transfected with the luciferase reporter plasmid for miR-1228 as a mismatch control.

SUPPLEMENTAL REFERENCES:

40. Berezikov, E., Chung, W.J., Willis, J., Cuppen, E. and Lai, E.C. (2007) Mammalian mirtron genes. *Mol Cell*, 28, 328-336.
80. Jodelka, F.M., Ebert, A.D., Duelli, D.M. and Hastings, M.L. (2010) A feedback loop regulates splicing of the spinal muscular atrophy-modifying gene, SMN2. *Hum Mol Genet*, 19, 4906-4917.
81. Cikos, S., Bukovska, A. and Koppel, J. (2007) Relative quantification of mRNA: comparison of methods currently used for real-time PCR data analysis. *BMC Mol Biol*, 8, 113.

SUPPLEMENTAL FIGURE LEGENDS:

Supplemental Table 1. Primer and siRNA sequences.

Supplemental Figure 1. Validation of the specificity of the stem-loop RT-PCR method for detection of miRNAs. (**A**) Diagram of the DNA templates used for in vitro transcription to generate synthetic mRNAs and miRNAs. The primer locations that were used for *in vitro* transcription are indicated. (**B**) Stemloop RT-PCR only amplifies RNA with 3' ends that match the stem-loop primer. Internal sequences that match the stemloop primer are not amplified. *In vitro* transcribed RNA was generated using the T7 Forward and Gene Specific Reverse primers for both wildtype (WT) and splicing-deficient (Δss) transcripts. The *in vitro* transcribed RNA was reverse transcribed in the presence (+) or absence (-) of reverse transcriptase using stemloop primers (to detect miRNA specifically) or gene specific primers (to detect and demonstrate the quality of the synthetic RNA). The cDNA was amplified via radiolabelled RT-PCR. (**C**) Synthetic miRNAs are specifically detected using the stem-loop RT-PCR method. RNA with the 3' end corresponding to the stemloop primer only amplifies in the presence of reverse transcriptase. *In vitro* transcribed RNA was generated using the T7 Forward and miRNA specific primers for both wildtype (WT) and splicing-deficient (Δss) transcripts. The *in vitro* transcribed RNA was reverse transcribed in the presence (+) or absence (-) of reverse transcriptase using stemloop primers. The cDNA was amplified via radiolabelled RT-PCR. M indicates a synthetic size marker for the mature miRNA species. • Indicates a non-specific primer dimer.

Supplemental Figure 2. Sequencing of miRNA. DNA sequencing traces of cloned (**A**) pre-miR-1225 and mature miR-1225-5p, (**B**) mature miR-1228 and (**C**) mature miR-877. The location of forward and linker-specific or stemloop primers used to clone and sequence the RNAs are shown. All samples were PCR amplified with indicated primers, ligated into pGEM-T Easy vector (Promega) and sequenced.

Supplemental Figure 3. qPCR analysis of simtrons. RNAs from wild-type (WT), splicing deficient (Δss), and empty vector (-) transfections for miR-1225 and miR-1228 were amplified via Taqman qPCR (Applied Biosystems). snoRNA48 (sno48) was used as a control. miRNA abundance was analyzed via the Liu and Saint method (81) and normalized using the equation: miRNA/sno48, n=5. Bars represent the average expression +/-SEM.

Supplemental Figure 4. Simtron 1228 processing is context independent and associates with Argonaute proteins. (**A**) Control, *Dicer* (Dicer-/-) or *DGCR8* (DGCR8-/-) knockout mouse embryonic stem cells were transiently transfected with the intergenic wild-type minigene (WT), or intergenic splicing-deficient minigene (Δss). Minigene-derived miRNAs and endogenous miR-16 were analyzed by stemloop RT-PCR. sno65 was analyzed as a loading control. (**B**) Graph shows quantitation of miR-1228 abundance using the equation: $(miRNA_{Dicer\text{-}/\text{-} \text{ or } DGCR8\text{-}/\text{-}}/sno65)/(miRNA_{control}/sno65)$. Bars represent the average values +/- SEM, n=4 for Dicer-/- and n=3 for DGCR8-/-. The horizontal dotted line indicates normalized control cell levels. (**C**) Luciferase expression in HEK-293T cells transiently co-transfected with pmiRGLO containing matching miRNA target sequences or mis-match miRNA target sequences and two different concentrations (0.4,

1 or 1.5 μg) of the wild-type (WT) or splicing-deficient (Δss) minigene. Horizontal dotted line represents the normalized mismatch control value. The following equation was used to analyze the data: $[(\text{Luciferase Target / Renilla})_{miR_match}]_{avg}/[(\text{Luciferase Target / Renilla})_{miR_mismatch}]_{avg}$. Each sample was analyzed in triplicate. Bars represent the average +/- SEM, n=3 independent experiments. (**D**) miR-1228 derived from WT and Δss minigenes were co-transfected with pFLAG-GFP, pFLAG-Ago1, pFLAG-Ago2, pFLAG-Ago3, pFLAG-Ago4 or mock (-) into HEK-293T cells. FLAG-tagged proteins were immunoprecipitated from cell lysates and associated miRNAs were analyzed by radiolabelled stemloop RT-PCR. Un indicates the unbound fraction, and IP is the immunoprecipitated miRNA. For wild-type minigenes, unbound is 1/20 of IP. For Δss minigenes, unbound is 1/5 of IP. Sno65 is a control non-coding RNA. miR-16 is a canonical miRNA control. (**E**) The graphs represent the percent of the pre-miRNA in the IP fraction as determined using the equation: (IP/(IP+(Un*20 or 5))*100). Left panel: WT miR-1228, Right panel: Δss miR-1228.

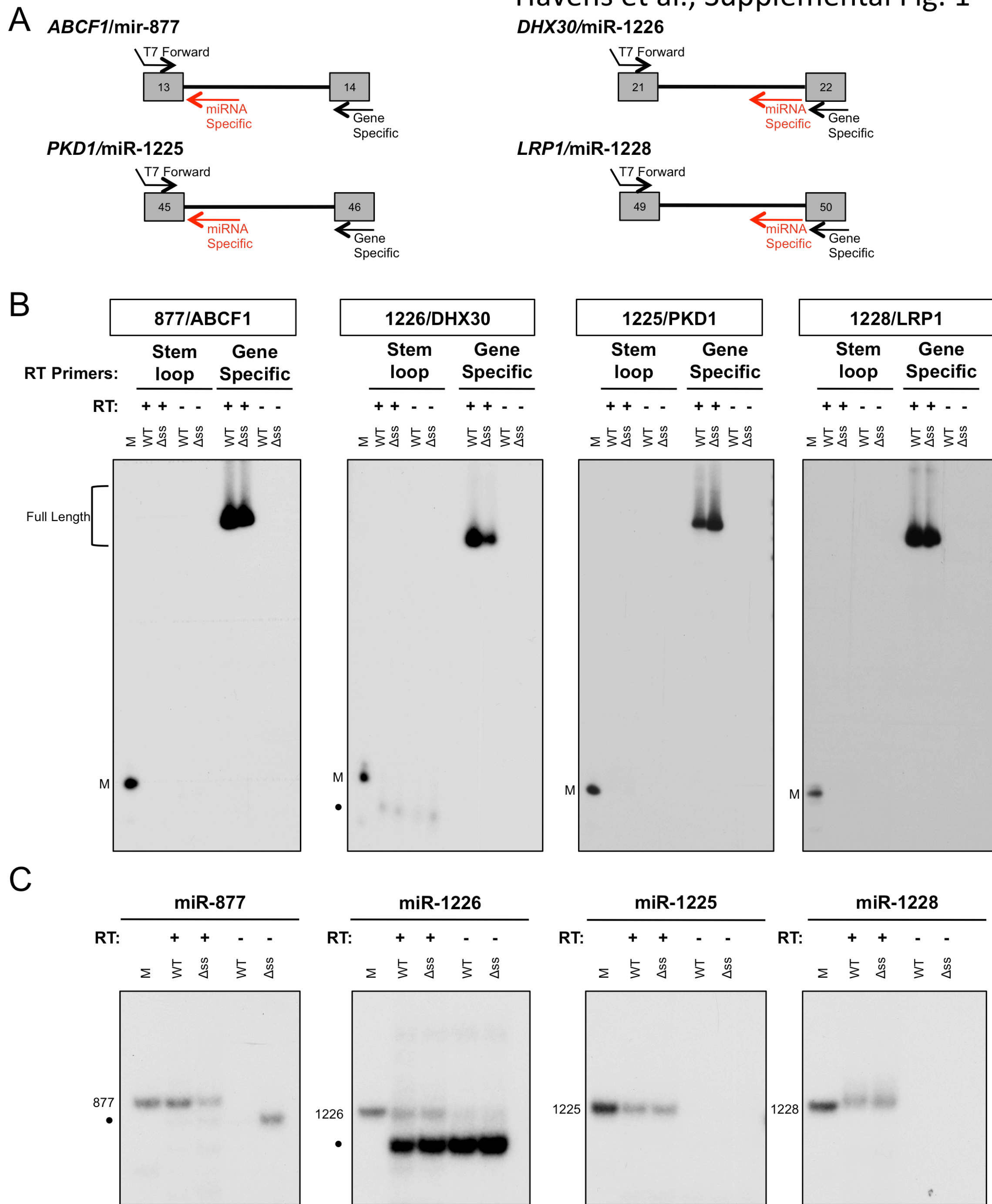
Havens et al. Supplemental Table 1

	Primer Name	Sequence
Stem loops		
	miR-16	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCGCCAA
	miR-1225	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCCCCCC
	miR-1228	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGGGGGG
	miR-1226	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCTAGGG
	miR-877	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCCCTGC
	snoRNA65	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGGCTGT
	MT miR-1226	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACATAGGG
	MT miR-877	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGTGGGA
	miR-877 + 1 NT 3'	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTGTCCC
	milR-1228 + 2NT 3'	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCTGGGG
	miR-1228 +2 NT 3' MT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGTGGGG
	miR-1228 +1 NT 3'	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTGGGGG
RNA linker 1		
	Linker 1 (IDT)	rAppCTGTAGGCACCATCAAT/3ddC/
miRNA Specific		
	miR-16 Forward	TAGCAGCACGTAAA
	miR-1225-Forward	GTGGGTACGGCCCA
	miR-1228 Forward	TCACACCTGCCTC
	miR-1226 Forward	TCACCAGCCCTGTG
	miR-877 Forward	GTAGAGGAGATGG
	snoRNA65 Forward	TGACTCTGTCCCGAAAGCAT
	Stemloop Reverse	CCAGTGCAGGGTCCGAG
miRNA Specific Reverse		
	mir-1225 Reverse	CCCCCCACAGGGCCGTACCC
	miR-1228 Reverse	GGGGGGCGAGGCAGGTGTGA
	miR-1226 Reverse	AGGGAACACAGGGCTGGTGA
	miR-877 Reverse	CCCTGCGCCATCTCCTCT
Gene Specific		
	PKD1 Forward	CGAAGCTTGCCTCGCTGCTCTTCCTG
	DHX30 Forward	CGAAGCTTACTCATCAAGCAGTTCTC
	LRP1 Forward	CGAAGCTTAGACAGCCTGTGGTGTGGGC
	DBR1EXON 5 Forward	TTGCCGCCTTGATGCAG
	DBR1 EXON 8 Reverse	CTGTTGCACTATAATCCC
	Drosha Forward	CATGCACCAGATTCTCCTGTA
	Drosha Reverse	GTCTCCTGCATAACTCAACTG
	DGCR8 Forward	TATCAGATCCTCCACGAGTG
	DGCR8 Reverse	TCTTGGAGCTTGCTGAGGAT
	GAPDH Forward	GAAGGTGAAGGTCGGAGTC
	GAPDH Reverse	GAAGATGGTGATGGGATTTC
	XPO5 Forward	CCTCCCGAGCACAACAAGG
	XPO5 Reverse	CTTAGAGAAGACCTGGGGC
	3' linker	ATTGATGGTGCCTAC
Vector specific		
	pTT3 Reverse	GGTCGAGGTCGGGGGATCC
	pCI Forward	GACTCACTATAGGCTAGCCTCG
	pCI Reverse	GTATCTTATCATGTCTGCTCG
siRNA		
	siDGCR8	Sense: CAUCGGACAAGAGUGUGAUTT Antisense: AUCACACUCUUGUCCGAUGTT
	siXPO5	Sense: AGAUGUUUCGAACACUAAATT Antisense: UUUAGUGUUCGAAACAUCUGG
Cloning		
	HindIII PKD1 ex 43 For	CGAAGCTTGCCTCGCTGCTCTTCCTG

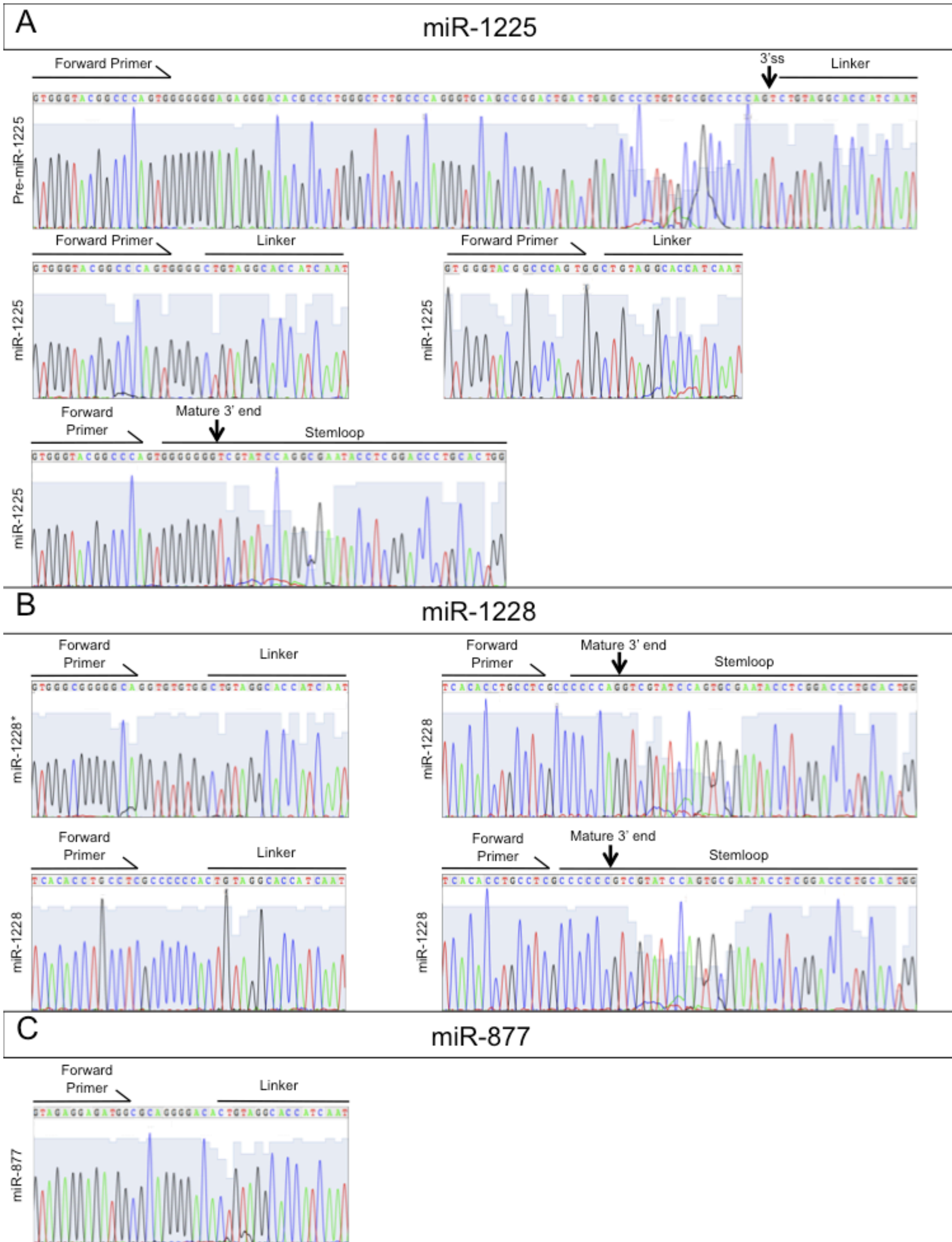
	BamHI PKD1 ex 46 Rev	CGGGATCCCTGGAGGAGGTGGAGGGG
	Xho ABCF1 ex 12 For	GGCTCGAGCAAGGGCAAGACCACAC
	NotI ABCF1 EX 15 REV	GGGCGGCCGCTTATTAAGCCAGATGACAGC
	HindIII LRP1 Ex 48 For	CGAAGCTTAGACAGCCTGTGGTGTGGGC
	BamHI LRP1 ex 50 Rev	CGGATCCGTCACTCACTGCAGTGGG
	HindIII DHX30 ex For	CGAAGCTTACTCATCAAGCAGTTCTC
	BamHI DHX30 ex Rev	CGGGATCCCACGGATGTGCACGTCCC
	miR-1225-5p Target 3' UTR	TCGAGGGCCCCCCACTGGGCCGTACCCACGTTTTGCCCCCCACTGGGCCG TACCCACAATCGATGATCT
	miR-1225-5p Target 3' UTR R	CTAGAGATCATCGATTGTGGGTACGGCCCAGTGGGGGCAAAACGTGGGT ACGGCCCAGTGGGGGGCC
	miR-877 Target 3' UTR	TCGAGGGCCCTGCGCCATCTCCTCTACGTTTTGCCCTGCGCCATCTCCTCT ACAATCGATGATCT
	miR-877 Target 3' UTR R	CTAGAGATCATCGATTGTAGAGGAGATGGCGCAGGGCAAAACGTAGAGGA GATGGCGCAGGGCC
	miR-1228 Target 3’ UTR	TCGAGGGGGGGGCGAGGCAGGTGTGAGTTTTGGGGGGGCGAGGCAGG TGTGAAATCGATGATCT
	miR-1228 Target 3’ UTR R	CTAGAGATCATCGATTTCACACCTGCCTCGCCCCCCCAAAACTCACACCTG CCTCGCCCCCCCC
	Mirtron intron 1225 BamHI	CGGGGATCCGTGGGTACGGCCCAGTGGGG
	Mirton intron 1225 EcoRI	CGGGAATTCCTGGGGGCGGCACAGGGGC
	MT Mirtron intron 1225 BamHI	CGGGGATCCGAGGGTACGGCCCAGTGGGG
	MT Mirton intron 1225 EcoRI	CGGGAATTCCAGGGGGCGGCACAGGGGC
	Mirtron intron 1228 BamHI	CGGGGATCCGTGGGCGGGGGCAGGTGTG
	Mirtron intron 1228 EcoRI	CGGGAATTCCTGGGGGGCGAGGCAGGTG
	MT Mirtron intron 1228 BamHI	CGGGGATCCCTGGGCGGGGGCAGGTGTG
	MT Mirtron intron 1228 EcoRI	CGGGAATTCGTGGGGGGCGAGGCAGGTG
	Mirtron intron 877 BamHI	CGGGGATCCGTAGAGGAGATGGCGCAGGG
	Mirtron intron 877 EcoRI	CGGGAATTCCTGGGAGGAGGGAGAAGAG
	MT Mirtron intron 877 BamHI	CGGGGATCCCTAGAGGAGATGGCGCAGGG
	MT Mirtron intron 877 EcoRI	CGGGAATTCGTGGGAGGAGGGAGAAGAG
	Mirtron intron 1228 BamHI	CGGGGATCCGTGGGCGGGGGCAGGTGTG
	Mirtron intron 1228 EcoRI	CGGGAATTCCTGGGGGGCGAGGCAGGTG
	MT Mirtron intron 1228 BamHI	CGGGGATCCCTGGGCGGGGGCAGGTGTG
	MT Mirtron intron 1228 EcoRI	CGGGAATTCGTGGGGGGCGAGGCAGGTG
Quik change		
	PKD1 5' SS For	TCAAGGAGGAGGGTACGGCC
	PKD1 5' SS Rev	GGCCGTACCCTCCTCCTTGA
	PKD1 3' SS For	TGCCGCCCCCTGTTCCGCCAC
	PKD1 3' SS Rev	GTGGCGGAACAGGGGGCGGCA
	ABCF1 5' SS For	CTAGAGAAGCTAGAGGAGATG
	ABCF1 5' SS Rev	CATCTCCTCTAGCTTCTCTAG
	ABCF1 3' SS For	CTCCTCCCACGTGTATGAG
	ABCF1 3' SS Rev	CTCATACACGTGGGAGGAG
	DHX30 5' SS For	CCATTAACAGTTGAGGGCATGC
	DHX30 5' SS Rev	GCATGCCCTCAACTGTTAATGG
	DHX30 3' SS For	GTGTTCCCTATGGAGGCCACACGG
	DHX30 3' SS Rev	CCGTGTGGCCTCCATAGGGAACAC
	LRP1 5' SS For	GACTGCCCAGCTGGGCGGGGG
	LRP1 5' SS Rev	CCCCCGCCCAGCTGGGCAGTC
	LRP1 3' SS For	CGCCCCCCACGTGTGAAACG
	LRP1 3' SS Rev	CGTTTCACACGTGGGGGGCG
	Drosha glu1222 Forward	GGCGGACCTTTTGGAATCATTTATTGCAGCG
	Drosha glu 1222 Reverse	CGCTGCAATAAATGATTCCAAAAGGTCCGCC
	Drosha glu 1045 Forward	GGCCAATTGTTTTGAAGCGTTAATAGG
	Drosha glu 1045 Reverse	CCTATTAACGCTTCAAAACAATTGGCC
***In vitro* Transcript.**		

	T7 PKD1 Forward	TAATACGACTCACTATAGGCTCGTGTCTTCCTGTGTGGAC
	PKD1 ex46 Reverse	AAGACTCACCTGGAGGAGG
	T7 pri-miR-16-1	TAATACGACTCACTATAGGCAGGCCATATTGTGCTGCCTC
	pri-miR-16-1 Reverse	CGGGATCCCAGAATCATACTAAAAATAAC
	T7 ABCF1 Forward	TAATACGACTCACTATAGGAGGTGGTAGCAGATGAG
	T7 ABCF1 ex14 Reverse	ATGGCCCACCTGGCCAGGG

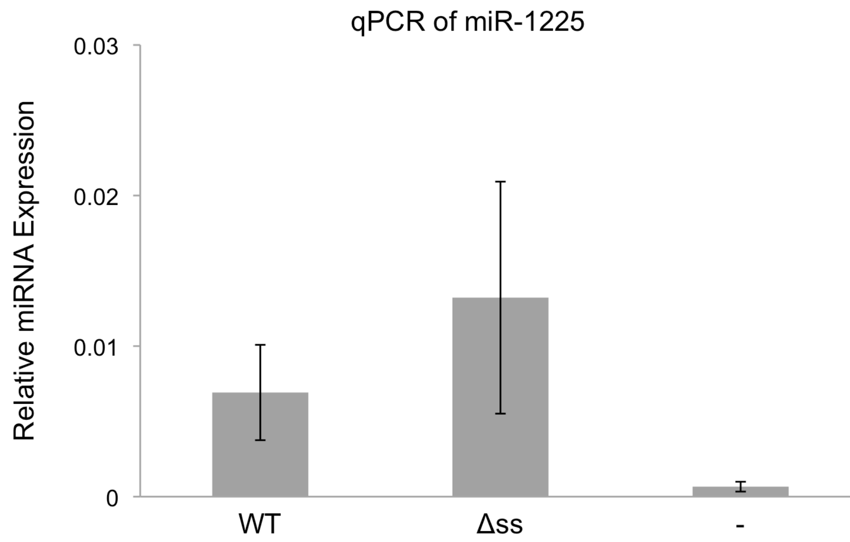
Havens et al., Supplemental Fig. 1

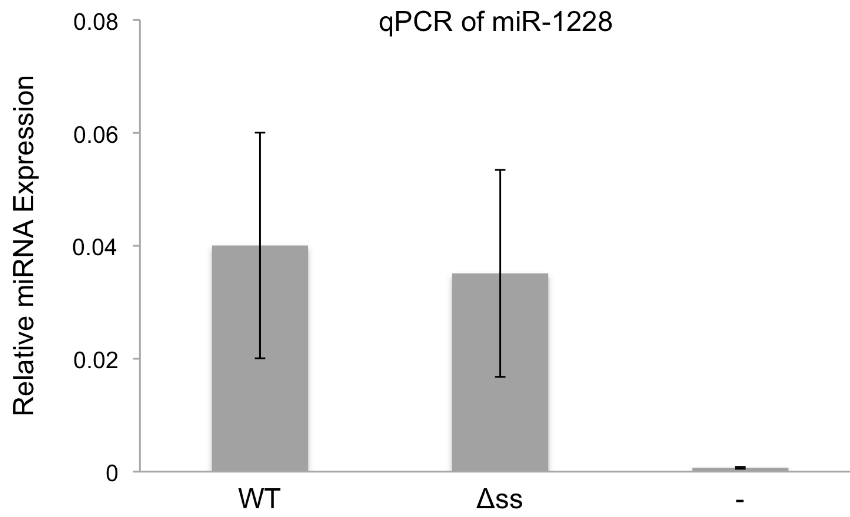


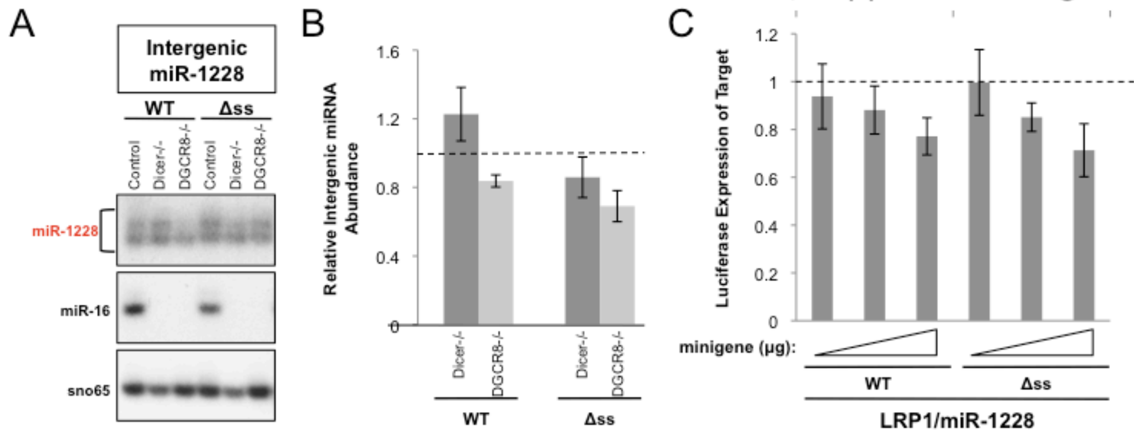
Havens et al., Supplemental Fig. 2

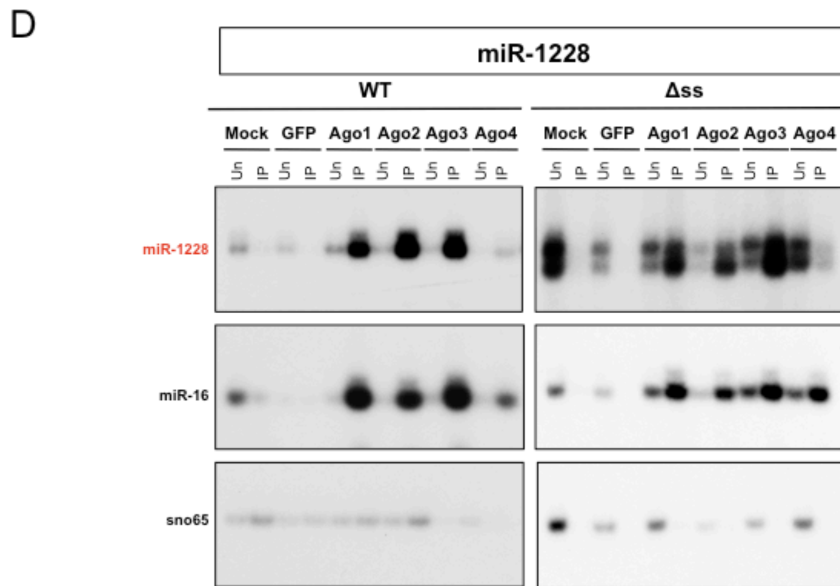


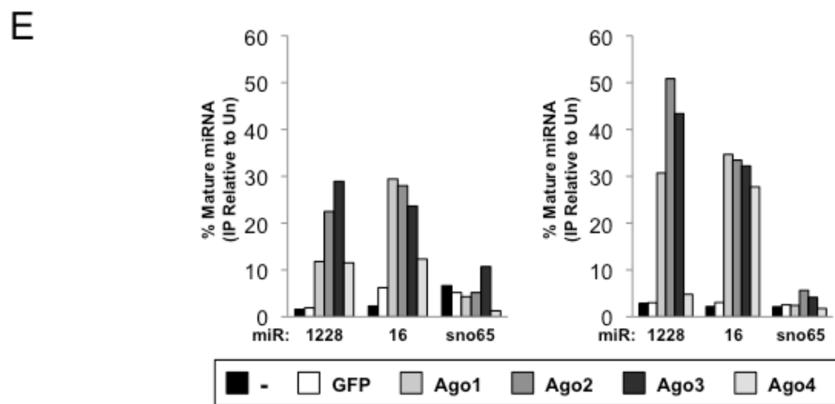
Havens et al., Supplemental Fig. 3






A
Intergenic miR-1228
WT
Δss
Control
Dicer-/-
DGCR8-/-
miR-1228
miR-16
sno65
B
Relative Intergenic miRNA Abundance
1.6
1.2
0.8
0.4
0
Dicer-/-
DGCR8-/-
WT
Δss
C
Luciferase Expression of Target
1.2
1
0.8
0.6
0.4
0.2
0
minigene (μg):
WT
Δss
LRP1/miR-1228


D
miR-1228
WT
Δss
Mock
GFP
Ago1
Ago2
Ago3
Ago4
Un
IP
miR-1228
miR-16
sno65


E
% Mature miRNA (IP Relative to Un)
60
50
40
30
20
10
0
miR:
1228
16
sno65
-
GFP
Ago1
Ago2
Ago3
Ago4