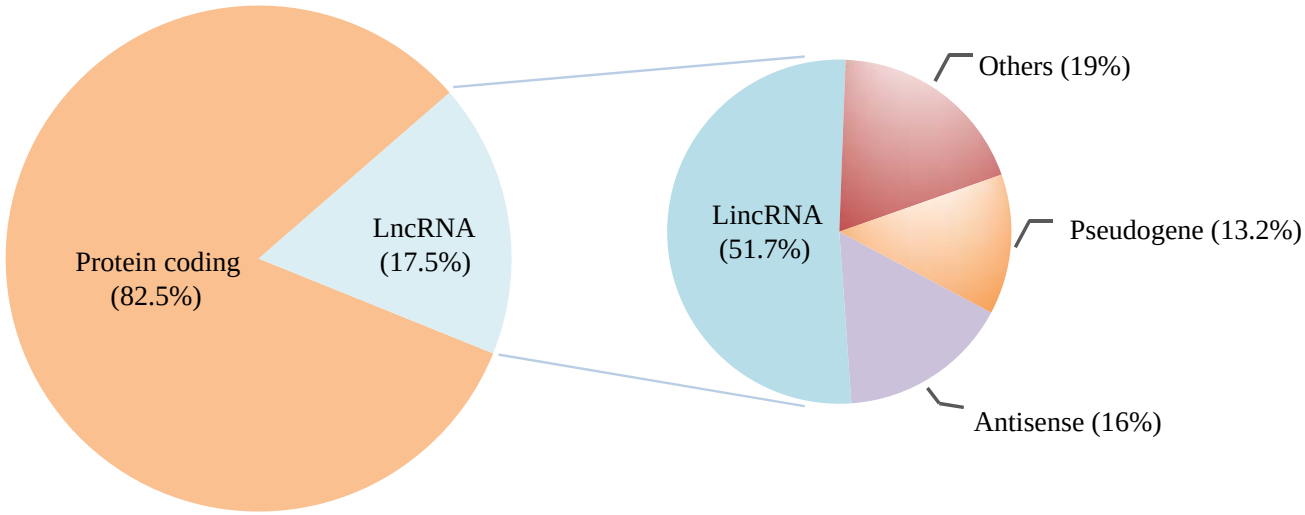
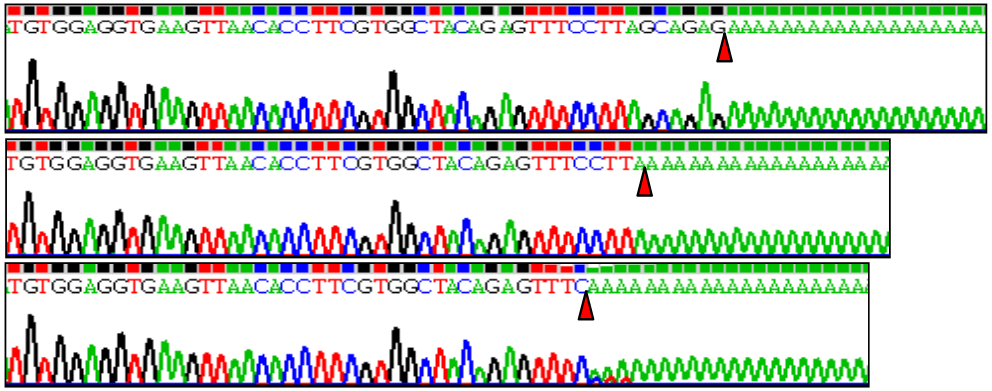
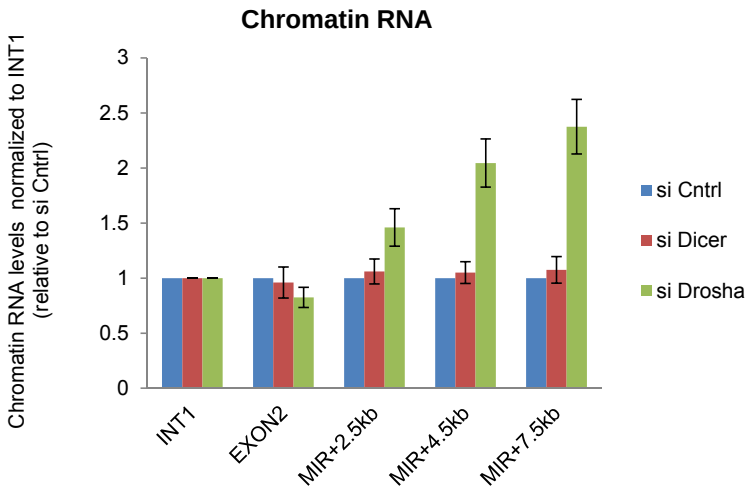
Supplementary Figure 1



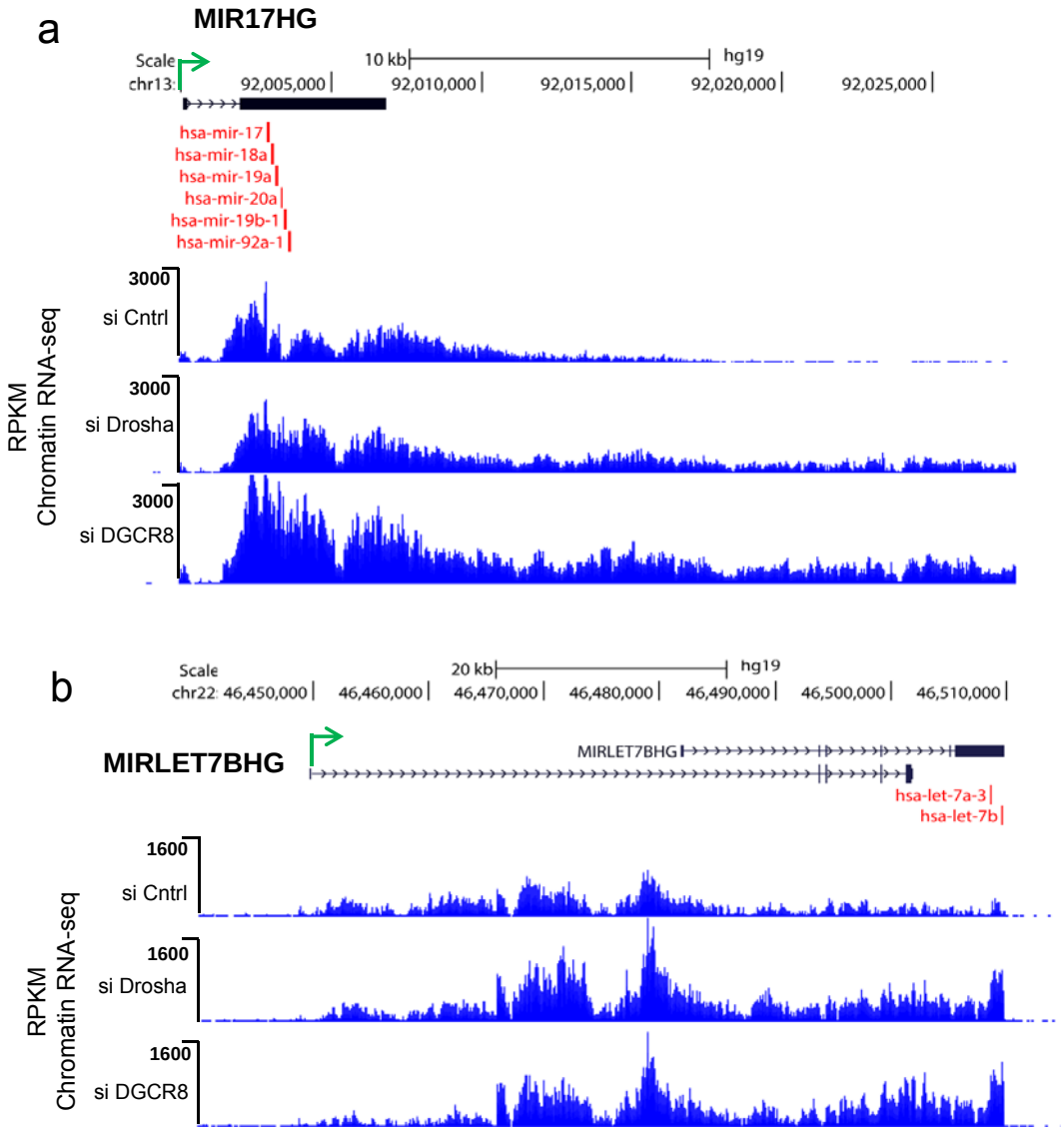
Supplementary Figure 2



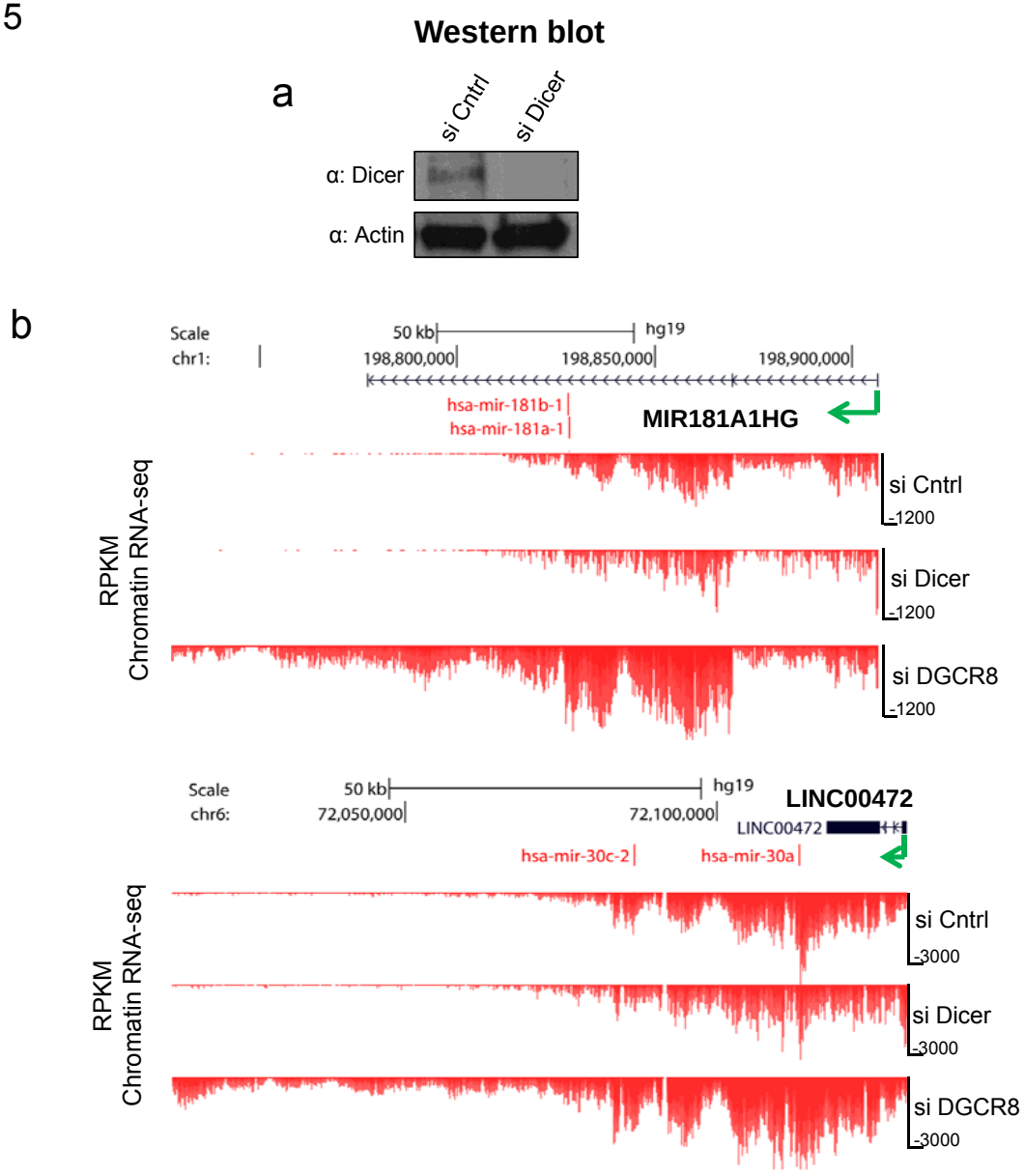
Supplementary Figure 3



Supplementary Figure 4

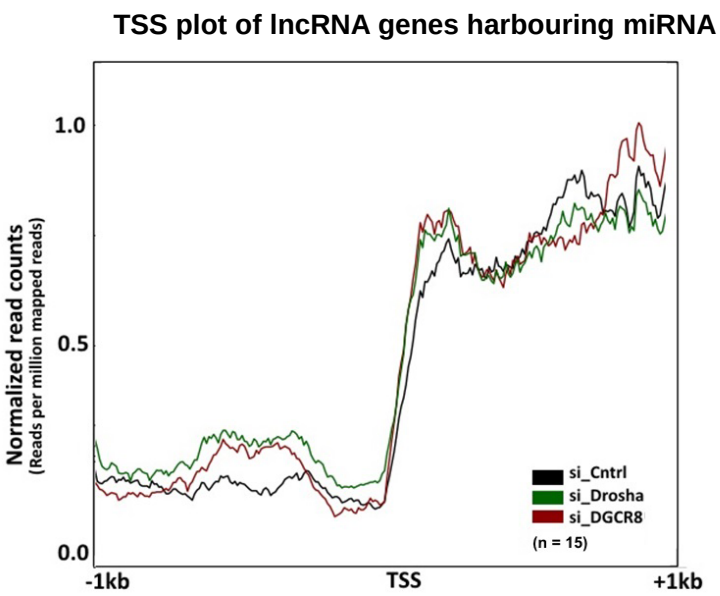


Supplementary Figure 5

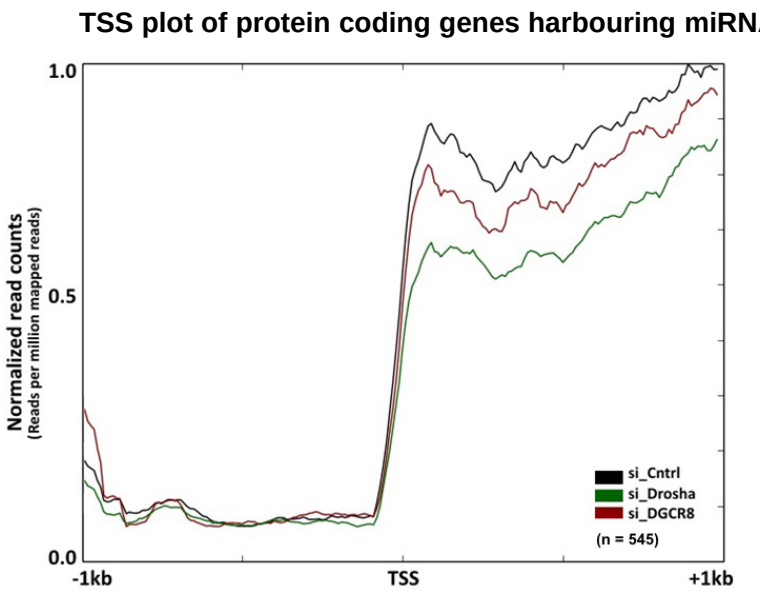


Supplementary Figure 6

a

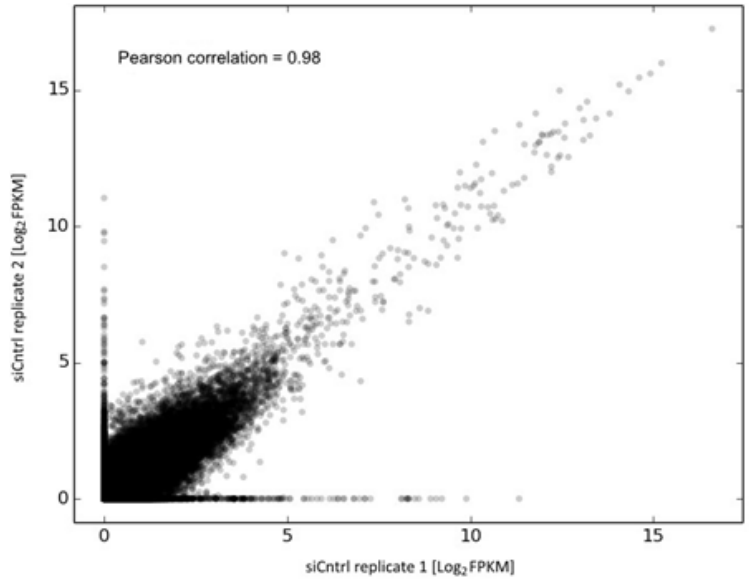


b

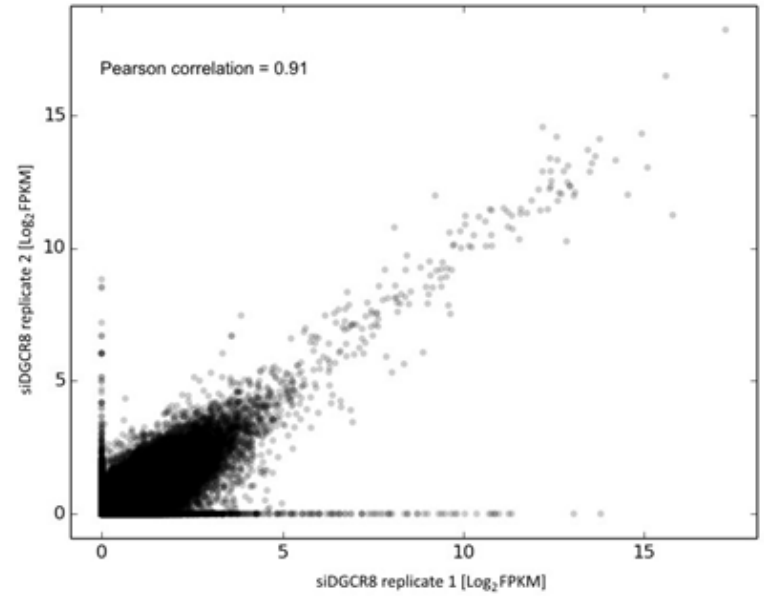


Supplementary Figure 7

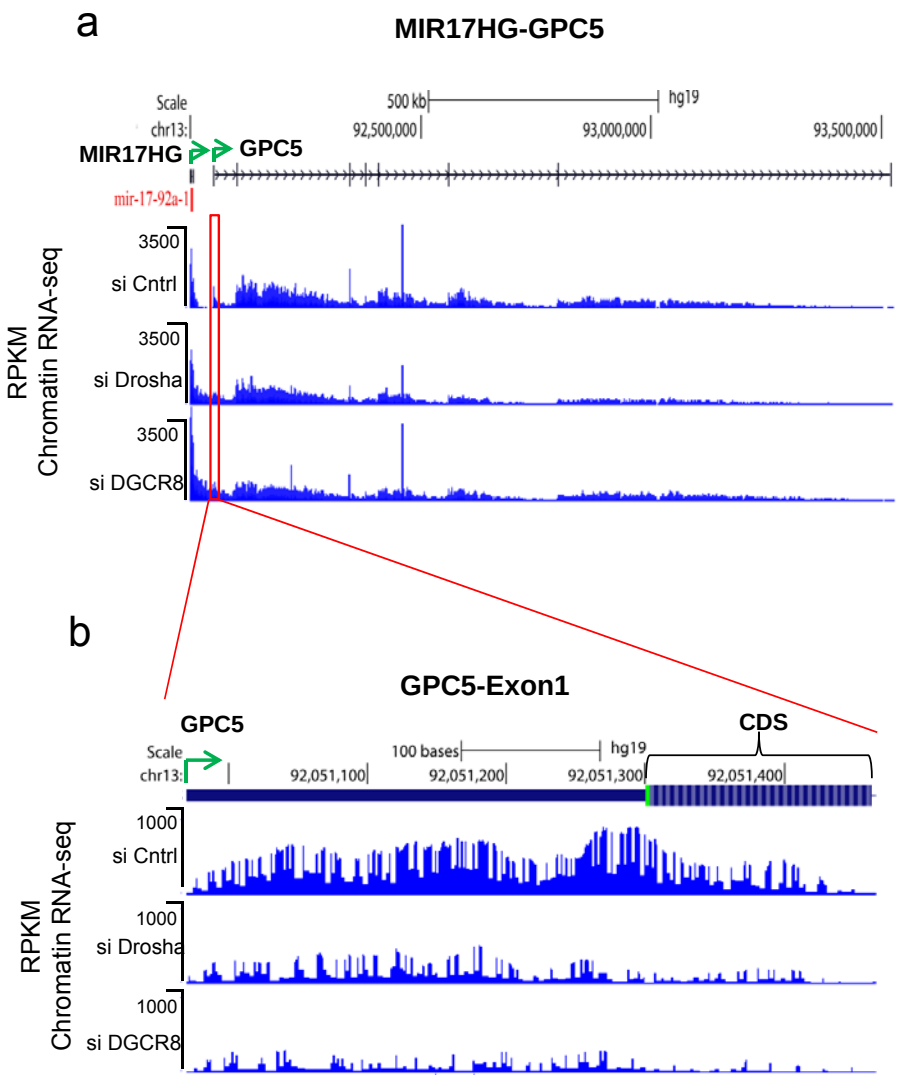
a



b



Supplementary Figure 8



Online Methods

PCR primers and siRNA sequences

See Supplementary Table 6.

Antibodies

anti-Glypican5 (ab124886; Abcam), anti-OGFRL1 (SC-137654; Santa Cruz), anti-Actin (A2103; Sigma), anti-Drosha (ab12286; Abcam), anti-DGCR8 (10996-1-AP; Proteintech), anti-CPSF73 (A301-090A; Bethyl), anti-α-Tubulin (T5168; Sigma), anti-H3 (ab1791; Abcam), anti-β-tubulin (ab6046; Abcam), anti-Dicer (13D6; Abcam).

Plasmid Constructs

Construction of βwt (formerly labeled HIVβ) has been described previously (Dye and Proudfoot, 1999). The pri-miR-122wt construct was made by insertion of a genomic PCR fragment generated using the primers Pri122qF/Pri122qR on Huh7 genomic DNA. The resulting ~8kb PCR fragment was ligated into a cloning vector prepared by long range PCR amplification of βwt with primers Open_BetaqR/B10 qF using Prime STAR HS DNA polymerase (Takara). The Quick change II XL site directed mutagenesis kit (Stratagene) was used to generate the various mutants using the following primer sets: primiR-122Δ (DELTA mirqF/DELTA mirqR); pA1mt (pA1MTqF/pA1MTqR).

Cell culture and transfection of siRNA and plasmids

HeLa, Huh7 and HepG2 cells were maintained in DMEM supplemented with 10% fetal bovine serum. For Huh7 and HepG2 culture, 1% non-essential amino acids (Invitrogen) were also included in the culture media. RNAi was performed using lipofectamine RNAiMax (Invitrogen), with siRNA delivered at 30nM final concentration. A second siRNA treatment was performed at 48h, and cells harvested at 72h after the first hit. Lipofectamine 2000 (Invitrogen) was used to deliver 0.1μg pri-miR-122 plasmid and 0.025μg pTAT per well of a 6 well plate.

RNA isolation and northern blot

RNA was isolated using TRIzol reagent (Ambion) according to the manufacturer's instructions. Total RNA from human liver was purchased from Agilent Technologies (Cat. No. 540017). Northern blotting was carried out using standard procedures on equal molar quantities of RNA. Membranes were probed with a random-primed ^{32}P-labelled DNA fragment corresponding to nt 3077-3707 (exon probe) and nt 1563-2005 (intron probe) of pri-miR-122 in Ultrahyb (Ambion). A fragment corresponding to nt 685-1171 of γ-actin was used as a loading control. For miRNA northern blots, the small RNA fraction was isolated by dissolving total RNA in 300µL of TE buffer with addition of equal amounts of PEG solution (20% PEG 8000, 2M NaCl). Samples were mixed and incubated for at least 30 min on ice, followed by centrifugation at 14,000*g* for 15min and isopropanol precipitation of the supernatant. Small RNA were run on a 18% polyacrylamide (19:1) urea gel and analyzed as described before[54] using a ^{32}P-end-labeled oligonucleotide complementary to miR-122.

Reverse transcription and real-time qPCR analysis

Total RNA was treated with DNase I (Roche) and reverse-transcribed using SuperScript Reverse Transcriptase III (Invitrogen) and random primers (Invitrogen). Real-time quantitative PCR (qPCR) was performed with 2x Sensimix SYBR mastermix (Bioline) and analysed on a Corbett Research Rotor-Gene GG-3000 machine.

Nuclear-cytoplasmic fractionation

The procedure for isolating nuclear and cytoplasmic RNA has been described elsewhere[55].

m^{7}G cap selection

Capped nuclear RNA was immunoprecipitated using a mouse monoclonal antibody against the 5´-terminal m^{7}G cap (cat.No.201001; SYSY Synaptic Systems) according to the manufacturer's protocol. RNA was analyzed by qPCR as described above and compared to 20% input.

PolyA+ and polyA- RNA separation

DNase I-treated nuclear RNA was incubated with oligodT magnetic beads (Dynabeads mRNA purification kit, Invitrogen) to isolate either polyA+ RNA, which was bound to beads, or polyA- RNA, which was present in the flowthrough after incubation. OligodT magnetic bead selection was performed twice to ensure pure polyA+ or polyA- populations. The polyA- RNA population was further processed with the Ribo-Zero Magnetic Kit (Human/Mouse/Rat, Epicentre) to deplete most of the abundant ribosomal RNA.

***In vitro* polyA tailing and 3' RACE**

To detect the 3' end of pri-miR-122, *in vitro* polyA tailing of nuclear RNA from Huh7 was carried out first using the polyA tailing kit (AM1350; Ambion) according to the manufacturer's protocol. Purified RNA was reverse-transcribed using an oligodT_{24}V anchor using SuperScript III (Invitrogen). Compatible linker and sense primer (fwd1) were used for PCR. PCR fragments were gel purified and cloned into TA cloning vector (StrataClone PCR cloning kit) for sequence verification.

RNA stability

Transcript half-life was estimated after actinomycin-D (5mg/ml) treatment of Huh7 cells. Transcript levels at the indicated time points were analyzed by qPCR. The relative levels of expression of each transcript at the different time points were normalized according to the levels at time 0.

Western blot

Cells extracts were prepared in 15mM HEPES (pH 7.5), 0.25M NaCl, 0.5% NP-40, 10% glycerol, 1x protease inhibitor (Roche) and 1mM phenyl methyl sulphonyl fluoride. Proteins were separated by 4-12% tris-glycine SDS–PAGE and transferred to nitrocellulose (0.45 µM, Amersham Biosciences), and protein detection was carried out with standard western blotting techniques. Secondary antibodies were anti-mouse (Sigma) and anti-rabbit (Sigma). Signals were detected with an ECL kit (GE Healthcare) and quantified using ImageJ software.

Br-UTP nuclear run-on analysis

The Br-UTP NRO was carried out largely as described[30] followed by qPCR analysis as described above. The primers used are listed in Supplementary Table 6.

Chromatin RNA isolation

The procedures for separating nuclear RNA into chromatin-associated and released fractions have been described before[56]. Chromatin-associated RNA from Huh7 cells was analyzed using qPCR as detailed above. The primers used are listed in Supplementary Table 6.

RNA-sequencing

Chromatin-associated RNA was isolated as described above with the omission of tRNA in solution preparation. Nuclear poly(A)- or poly(A)+ RNA were prepared as described above. RNA-seq was performed by the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics, University of Oxford. RNA Samples were ribodepleted using Ribo-Zero rRNA removal kit (Human/Mouse/Rat, EpiCentre RZH110424). Libraries were prepared using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina, v1.0 (cat # E7420) using manufacturer's guidelines with an exception of using our own 8bp tags for indexing according to[57]. Libraries were sequenced on an Illumina HiSeq-2000 using 100bp and 50bp paired end reads, v3 chemistry.

Bioinformatic analysis

Datasets

All human miRNAs and their genomic coordinates were obtained from miRBase release 20[58]. Annotation for protein coding genes (GRCh37) was obtained from Ensembl, release 74[59]. Annotation for long noncoding RNA (lncRNA) was obtained from GENCODE v7 catalog of human long noncoding RNA[50].

We separated the miRNA into two groups: lnc-pri-miRNA (if the miRNA overlapped with an lncRNA) and protein coding miRNA (if the miRNA overlapped with an Ensembl protein coding gene), based on the genomic coordinates and taking the strand orientation into account. miRNA harboring genes having length ≥200 bp were

considered for this study. This resulted in 112 lnc-pri-miRNA and 967 protein coding pri-miRNA. An additional 18 lnc-pri-miRNA were added to this list which mapped to genes having 'lincRNA' as Ensembl gene biotype and gene length ≥200 bp.

Genomic distribution of miRNA as belonging to protein coding genes and long non-coding genes was based on Ensembl gene biotype annotation (Supplementary Fig.1).

Mapping of sequencing reads

Paired-end reads for each sample were mapped to the human genome reference assembly GRh37/hg19 (build 37.2, Feb 2009) using the Bowtie2 alignment software[60]. Prior to alignment, the first 12 nucleotides were trimmed from all the reads owing to the low quality of the bases. Uniquely mapped reads with no more than two mismatches were retained for further analysis. For nuclear polyA+ data, we filtered out reads that had 8 or more genomically encoded A-stretch at their 3' ends. A statistical summary of read alignments can be found in Supplementary Table 5 for HeLa and Huh7 chromatin RNA-seq and HeLa nuclear polyA+ and polyA- RNA-seq.

Calculation of metagene profiles

We used the Ensembl gene annotation to define transcription start and end sites. In-house Perl and Python scripts were used to compute metagene profiles. To get a list of miRNA expressed in HeLa cells, miRNA FPKM was calculated using Cufflinks on small RNA-seq data for HeLa cells downloaded from the ENCODE Experiment Matrix available as ENCODE Project at UCSC[61]. This resulted in 15 lncRNA expressing 34 miRNA in HeLa cells. For the protein coding dataset, miRNA harboring genes with length ≥2 kb and the gene body RPKM ≥ 1 were considered. 545 genes were identified in this category.

Metagene profile (Figure 6) for si Cntrl, si Drosha and si DGCR8 in HeLa cells was calculated for a region from the start of the miRNA host gene to 1 kb upstream of the start of the next downstream gene. For this, read counts were normalized to total sequencing depth. The region extending from TSS to TES was scaled to 4 kb and the region from downstream of TES to 1 kb upstream of the start of the next gene was scaled to 10 kb. Normalized read counts were plotted for each 20 bp bin.

To investigate the profile surrounding TSS upon Microprocessor depletion, we plotted normalized read count across a region of 1kb upstream and downstream of annotated TSS for si Cntrl, si Drosha and si DGCR8 in HeLa cells.

For supplementary Table 1 and 2, normalized read count (RPKM) was calculated over a region from the TES (3'end) of the lnc-pri-miRNA to 1 kb upstream of the TSS of the next downstream gene for miRNA harboring lncRNA genes that are expressed (RPKM≥1) in HeLa and Huh7 cell lines.

Classification of polyA+, polyA- and bimorphic transcripts

Classification of transcripts into polyA+, polyA- and bimorphic was done as described[62]. Briefly, all expressed transcripts were classified as polyA+, polyA- and bimorphic predominant transcripts based on their relative abundance, calculated using BPKM (bases per kilobase of gene model per million mapped bases) in the polyA+ and polyA- sample for each condition (si Cntrl and si DGCR8). See Supplementary Table 3.

References

54. Pall, G.S. & Hamilton, A.J. Improved northern blot method for enhanced detection of small RNA. *Nat. Protoc.* **3**, 1077-84 (2008).
55. Dye, M.J., Gromak, N. & Proudfoot, N.J. Exon tethering in transcription by RNA polymerase II. *Mol. Cell* **21**, 849-59 (2006).
56. West, S., Proudfoot, N.J. & Dye, M.J. Molecular dissection of mammalian RNA polymerase II transcriptional termination. *Mol. Cell* **29**, 600-10 (2008).
57. Lamble, S. et al. Improved workflows for high throughput library preparation using the transposome-based nextera system. *BMC Biotechnol.* **13**, 104 (2013).
58. Kozomara, A. & Griffiths-Jones, S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **39**, D152-7 (2014).
59. Flicek, P. et al. Ensembl 2014. *Nucleic Acids Res.* **42**, D749-55 (2014).
60. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357-9 (2012).

61. Kent, W.J. et al. The human genome browser at UCSC. *Genome Res.* **12**, 996-1006 (2002).
62. Yang, L., Duff, M.O., Graveley, B.R., Carmichael, G.G. & Chen, L.L. Genomewide characterization of non-polyadenylated RNAs. *Genome Biol.* **12**, R16 (2011).

Supplementary figure legends

Figure 1: Distribution of miRNA between lncRNA and protein coding genes.

17.5% of all miRNA derive from lncRNA. These can be further subdivided into intergenic (lincRNA) or other less well-characterized subdivisions such as pseudogene or antisense.

Figure 2: Transcript properties of lnc-pri-miR-122.

Sequencing results of lnc-pri-miR-122 3'RACE products amplified by poly(A) polymerase dependent method.

Figure 3: Drosha but not Dicer is required for lnc-pri-miRNA-122 transcription termination.

Chromatin-associated RNA was analyzed by qPCR as in Fig. 3a. The efficiency of siRNA mediated knockdown of Drosha and Dicer is shown by western blots in Fig 6e and Supplementary Fig 5a respectively.

Figure 4: Chromatin RNA-seq profiles of two lnc-pri-miRNA in HeLa cells.

a. Chromatin RNA-seq profile showing that Drosha or DGCR8 depletion leads to transcriptional readthrough at the MIR17HG locus. **b.** Chromatin RNA-seq profile showing no transcriptional readthrough at the MIRLET7BHG locus following Drosha or DGCR8 depletion.

Figure 5: Dicer depletion does not lead to transcriptional readthrough on lnc-pri-miRNA.

a. Western blot showing effective Dicer depletion by siRNA transfection in HeLa cells. **b.** Chromatin RNA-seq profiles for MIR181A1HG and LINC00472 showing transcriptional readthrough on DGCR8 but not Dicer depletion.

Figure 6: Effect of Microprocessor knockdown on levels of TSS transcripts harboring miRNA in HeLa cells.

a. and **b.** TSS metagene plot of chromatin RNA-seq of lnc-pri-miRNA versus protein coding genes harboring intronic miRNA showing region 1 kb before and after TSS.

Figure 7: Scatter plots showing reproducibility of replicate Chromatin RNA-seq.

a. si Cntrl (HeLa) **b**. si DGCR8 (HeLa).

Figure 8: Additional views of GPC5 Chromatin RNA-seq profiles.

a. Compressed view showing full extent of GPC5 transcription unit. **b.** Magnified view of GPC5 exon 1 showing loss of reads following Microprocessor knockdown. Coding sequence (CDS) that translates first 55 amino acids of the GPC5 protein are located in the exon 1 and are shown with bracket. HeLa cell RNA employed.

Supplementary Table 1: List of miRNA harbouring lncRNA host genes showing termination defect following Drosha/DGCR8 depletion in HeLa cell line.

lncRNA host gene	miRNAs	strand	Normalized Read counts (RPKM)			Fold increase		Read-through	
			si Cntrl	si Drosha	si DGCR8	si Drosha	si DGCR8	si Drosha	si DGCR8
MIR17HG	hsa-miR-(17~18a~19a~19b-1~20a~92a-1)	(+)	57424.66	193251.73	339808.14	3.37	5.92	YES	YES
MIR181A2HG	hsa-miR-(181a-2~181b-2)	(+)	450.27	2687.66	4790.68	5.97	10.64	YES	YES
RP11-2B6.3	hsa-let-(7a-1~7f-1~7d)	(+)	60.04	433.49	9390.42	7.22	156.41	YES	YES
RP11-631N16.2	hsa-let-7i	(+)	6211.77	13798.26	70840.17	2.22	11.40	YES	YES
LINC00478	hsa-miR- (99a~let-7c~125b-2)	(+)	1891.14	4306.04	6127.20	2.28	3.24	YES	YES
MIR181A1HG	hsa-miR-(181a-1~181b-1)	(-)	630.38	82826.30	112737.08	131.39	178.84	YES	YES
LOC646329	hsa-miR-(29a~29b-1)	(-)	6393.86	148211.68	223252.40	23.18	34.92	YES	YES
MIR31HG	hsa-miR-31	(-)	2401.45	25821.81	26817.37	10.75	11.17	YES	YES
LINC00472	hsa-miR-(30a~30c-2)	(-)	14648.84	74459.86	114837.34	5.08	7.84	YES	YES
MIR100HG	hsa-miR-(100~125b-1~let-7a-2)	(-)	4007.42	22772.90	31538.61	5.68	7.87	YES	YES
DLEU2	hsa-miR-(15a~16-1)	(-)	26145.78	93620.31	62868.93	3.58	2.40	YES	YES
MIRLET7BHG	hsa-let-(7a-3~7b)	(+)	410.31	609.19	494.30	1.48	1.20	NO	NO
LOC284454	hsa-miR-23a~24-2~27a	(-)	2716.64	4739.54	5414.20	1.74	1.99	NO	NO
MIR22HG	hsa-miR-22	(-)	120.07	187.85	69.43	1.56	0.58	NO	NO
MIR600HG	hsa-miR-600	(-)	1921.16	2239.72	1701.04	1.17	0.89	NO	NO

RPKM fold increase >= 2 is considered as genes showing read-through.

Shaded region in the table highlights lncRNA host genes that do not show termination defect following Drosha/DGCR8 depletion.

Supplementary Table 2: List of miRNA harbouring lncRNA host genes showing termination defect following DGCR8 depletion in Huh7 cell line.

lncRNA host gene	miRNAs	strand	Normalized Read counts (RPKM)		Fold increase	Read-through
			si Cntrl	si DGCR8	si DGCR8	si DGCR8
MIR17HG	hsa-miR-(17~18a~19a~19b-1~20a~92a-1)	(+)	16404.90	72189.92	4.40	YES
MIR181A2HG	hsa-miR-(181a-2~181b-2)	(+)	810.49	8748.19	10.79	YES
RP11-2B6.3	hsa-let-(7a-1~7f-1~7d)	(+)	435.26	5901.55	13.56	YES
MIR146A	hsa-miR-(146a~3142)	(+)	1335.80	19336.27	14.48	YES
RP11-65J21.3	hsa-miR-(193b~365a)	(+)	1650.99	7862.95	4.76	YES
RP11-631N16.2	hsa-let-7i	(+)	3347.02	12445.34	3.72	YES
AP000662.4	hsa-miR-130a	(+)	4012.84	9425.13	2.35	YES
lnc-pri-miR-122*	hsa-miR-122	(+)	18556.80	41382.39	2.23	YES
MIR940	hsa-miR-(940~3677)	(+)	60.04	121.50	2.02	YES
MIR181A1HG	hsa-miR-(181a-1~181b-1)	(-)	120.07	20915.81	174.19	YES
LOC646329	hsa-miR-(29a~29b-1)	(-)	2671.61	63840.95	23.89	YES
AB429224	hsa-miR-194-2	(-)	285.17	2395.38	8.40	YES
MIR137HG	hsa-miR-(137~2682)	(-)	450.27	2655.70	5.89	YES
AC034220.3	hsa-miR-3936	(-)	120.07	624.87	5.20	YES
MIR210HG	hsa-miR-210	(-)	300.18	1128.24	3.76	YES
DLEU2	hsa-miR-(15a~16-1)	(-)	24629.87	83472.31	3.39	YES
LINC00472	hsa-miR-(30a~30c-2)	(-)	645.39	1614.25	2.50	YES
MIRLET7BHG	hsa-let-(7a-3~7b)	(+)	14.65	22.43	1.53	NO
LINC00478	hsa-miR- (99a~let-7c~125b-2)	(+)	420.25	833.16	1.98	NO
MIR22HG	hsa-miR-22	(-)	255.15	433.93	1.70	NO
LOC284454	hsa-miR-23a~24-2~27a	(-)	615.37	1041.45	1.69	NO
MIR600HG	hsa-miR-600	(-)	2341.41	2655.70	1.13	NO
MIR100HG	hsa-miR-(100~125b-1~let-7a-2)	(-)	2401.45	2082.90	0.87	NO

RPKM fold increase >= 2 is considered as genes showing read-through.

Shaded region in the table highlights lncRNA host genes that do not show termination defect following DGCR8 depletion

Genes expressed in both HeLa and HUH7 cell lines are highlighted in bold.

*Based on our data: Read through calculated over chromosomal region (chr18:56120306-56130306) due to lack of available gene annotation.

Supplementary Table 3: Polyadenylation status of miRNA-expressing lncRNA host genes in control vs DGCR8 depleted HeLa cells.

	si Cntrl				si DGCR8			
lncRNA host gene	BPKM pA(-)	BPKM pA(+)	Fold change	polyA status	BPKM pA(-)	BPKM pA(+)	Fold change	polyA status
MIR17HG	21.37	1.89	11.31	pA(-)	43.69	5.8	7.53	pA(-)
MIR181A2HG	3.24	0.49	6.59	pA(-)	4.25	1.03	4.12	pA(-)
RP11-2B6.3	3.37	0.37	9.03	pA(-)	7.3	0.95	7.7	pA(-)
RP11-631N16.2	6.12	1.03	5.91	pA(-)	13.35	2.66	5.02	pA(-)
LINC00478	1.46	0.17	8.45	pA(-)	2.17	0.32	6.78	pA(-)
MIR181A1HG	1.07	0.07	10.31	pA(-)	3.2	0.45	7.04	pA(-)
LOC646329	2.02	0.5	4	pA(-)	10.6	2.96	3.58	pA(-)
MIR31HG	4.18	0.51	8.12	pA(-)	6.95	1.22	5.68	pA(-)
LINC00472	7.65	1.02	7.62	pA(-)	22.9	3.81	6.01	pA(-)
MIR100HG	1.59	0.27	5.93	pA(-)	7.58	1.26	6.04	pA(-)
DLEU2	2.27	0.58	3.93	pA(-)	3.2	0.97	3.28	pA(-)
MIRLET7BHG	9.36	3.67	2.55	pA(-)	20.07	57.29	0.35	pA(+)
LOC284454	15.49	46.46	0.33	pA(+)	19.96	69.89	0.29	pA(+)
MIR22HG	6.05	1.58	3.82	pA(-)	9.82	7.27	1.35	bimorphic
MIR600HG	1.41	1.12	1.26	bimorphic	1.73	1.38	1.26	bimorphic

Shaded region in the table highlights lncRNA host genes that do not show termination defect following Drosha/DGCR8 depletion.

Criteria for polyA status:
pA(-): Fold change values ≥2
pA(+): Fold change values ≤0.5
Bimorphic: Fold change values between 0.5 and 2

Supplementary Table 4: Gene length and distance of pre-miRNA from TSS in lnc-pri-miRNA showing Microprocessor-mediated transcription termination in (a) HeLa and (b) Huh7 cells.

a

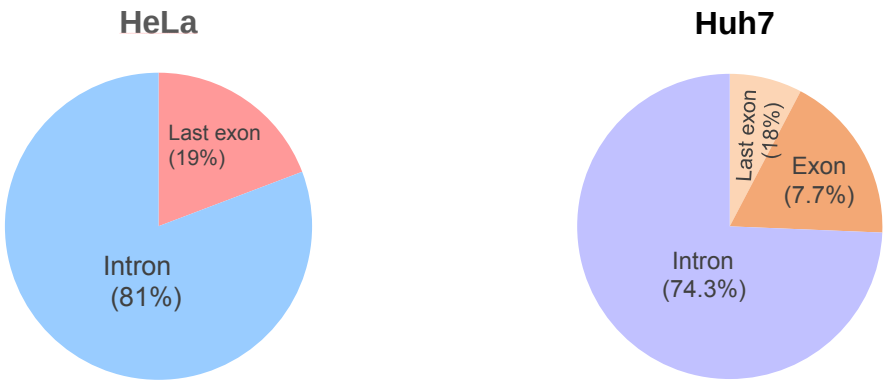
lncRNA host gene	Gene length	Distance from TSS ǂ
MIR17HG	6759	2785
MIR181A2HG	40164	33975
LINC00478	556874	468567
RP11-2B6.3	34338	9668
RP11-631N16.2	11323	934
LINC00472	76425	32616
LOC646329	39660	4255
MIR181A1HG	129936	51380
DLEU2	98587	21840
MIR31HG	104027	56473
MIR100HG	113959	10654

b

lncRNA host gene	Gene length	Distance from TSS ǂ
MIR210HG	2797	2429
AB429224	4987	2893
MIR940	5263	2558
MIR17HG	6759	2785
RP11-631N16.2	11323	934
AP000662.4	14766	3174
MIR146A	19158	6134
RP11-65J21.3	24065	1679
RP11-2B6.3	34338	9668
LOC646329	39660	4255
MIR181A2HG	40164	33975
AC034220.3	58630	54204
MIR137HG	61863	57242
LINC00472	76425	32616
DLEU2	98587	21840
MIR181A1HG	129936	51380

ǂ For miRNAs in a cluster, distance is calculated from the TSS to the first miRNA

c



SupplementaryTable 5: RNA-seq mapping statistics.

Sample ID	RNA-seq	Type	Read size (nt)	Number of PE reads	Mapped PE reads	% of mapped PE reads
si_Cntrl	Chromatin	Hela	51	38343118	33313212	86.88%
si_DGCR8	Chromatin	Hela	51	33569636	28805961	85.81%
si_Drosha	Chromatin	Hela	51	38989688	34602536	88.75%
si_Dicer	Chromatin	Hela	51	17923502	12819860	71.53%
si_Cntrl	Chromatin	Huh7	100	20678124	18305816	88.53%
si_DGCR8	Chromatin	Huh7	100	25042908	21524505	85.95%
si_Cntrl_PolyAplus	Nuclear	Hela	100	22919390	15798545	68.93%
si_Cntrl_PolyAminus	Nuclear	Hela	100	21002155	14045951	66.88%
si_DGCR8_PolyAplus	Nuclear	Hela	100	24108556	16607260	68.89%
si_DGCR8_PolyAminus	Nuclear	Hela	100	25086373	17768424	70.83%

PE denotes paired end

Supplementary Table 6: Oligonucleotide primers and siRNAs used in study

Primer (qRT-PCR)	**Gene**	**Sequence (5'-3')**
EX1 forward	lnc-pri-miR-122	CAGAAGCTGTGGAAGGCGC
EX2 reverse	lnc-pri-miR-122	GAAACAGTGAGAGGTGAACAATTCA
INT1 forward	lnc-pri-miR-122	TGCCACTCAGCACAGCACTA
INT1 reverse	lnc-pri-miR-122	TGAACCCATCCTGCTCATA
GAPDH mRNA forward	GAPDH	AAGGTGAAGGTCGGAGTCAA
GAPDH mRNA reverse	GAPDH	AATGAAGGGGTCATTGATGG
U6 snRNA forward	U6 snRNA	CTCGCTTCGGCAGCACA
U6 snRNA reverse	U6 snRNA	AACGCTTCACGAATTTGCGT
EXON2 forward	lnc-pri-miR-122	ACATGCCTGTGTCCACTGCT
EXON2 reverse	lnc-pri-miR-122	CTCTAGGTGGCCCCAGTCAC
MIR+2.5kb forward	lnc-pri-miR-122	GGACTGGGTCTTCCATGCTC
MIR+2.5kb reverse	lnc-pri-miR-122	AAACGAATCCAGGCAGCAAT
MIR+4.5kb forward	lnc-pri-miR-122	CATTCACTCCTGCACGCATT
MIR+4.5kb reverse	lnc-pri-miR-122	GGTTCTGGGGAGAGCATCAC
MIR+7.5kb forward	lnc-pri-miR-122	GCAACTGAAGCAGCATCGTT
MIR+7.5kb reverse	lnc-pri-miR-122	CACGGGGGAGCTTTTCTCTT
INT1 forward	GAPDH	CCCCTTCATACCCTCACGTA
INT1 reverse	GAPDH	GACAAGCTTCCCGTTCTCAG
PAS forward	GAPDH	CTGAATCTCCCCTCCTCACA
PAS reverse	GAPDH	TGCCCCAGACCCTAGAATAA
PAS+1.1kb forward	GAPDH	TCCAGCCTAGGCAACAGAGT
PAS+1.1kb reverse	GAPDH	TGTGCACTTTGGTGTCACTG
Exon1 forward	GPC5	TCGGCTAGGGAAGAAGACCA
Exon1 reverse	GPC5	GCAGTTAGCTCTTCCCGGAG
-2kb GPC5 forward	GPC5	TGACCTCCCACCTGTTTTGT
-2kb GPC5 reverse	GPC5	AACGGTGTGTCTTGCAAGGA
exon1-2 forward	GPC5	GAAACTTTTCCAGTGGCGGC
exon1-2 reverse	GPC5	CTCCTCCATCTTCCTGGTGC
exon3-4 forward	OGFRL1	TCCATTCAAGCCAGATGGTGT
exon3-4 reverse	OGFRL1	AGTGTGGTTGTGCTCCAGTT
Primer (PCR-Gel)	**Gene**	**Sequence (5'-3')**
exon2 (Total) forward	lnc-pri-miR-122	GGTGAAGAGGTGAGAGTTGGA
exon2 (Total) reverse	lnc-pri-miR-122	CAGCAAACCCCTTCTGCAAA
EX1 (spliced) forward	lnc-pri-miR-122	CAGAAGCTGTGGAAGGCGC
EX2 (spliced) reverse	lnc-pri-miR-122	GAAACAGTGAGAGGTGAACAATTCA
INT1 (unspliced) forward	lnc-pri-miR-122	ACAACTGGGCTCAAGGGATC
INT1 (unspliced) reverse	lnc-pri-miR-122	TCTGTTTAGCTTCCCTGGGC
GAPDH mRNA forward	GAPDH	AAGGTGAAGGTCGGAGTCAA
GAPDH mRNA reverse	GAPDH	AATGAAGGGGTCATTGATGG
fwd1 (3'end mapping)	lnc-pri-miR-122	CCCGTGATGCTTCTTTTCTC
Linker (3'end mapping)		GCTGTCAACGATACGCTACGTAACG
3'RACE-$(dT)_{24}$V		GCTGTCAACGATACGCTACGTAACG$(T)_{24}$V
Anti-miR-122 **(Northern blot)**		AACACCAUUGUCACACUCCAUA (RNA oligo)
siRNA	**Gene**	**siRNA Forward Strand**
si Cntrl	Luciferase	GAUUAUGUCCGGUUAUGUAUU
si DGCR8	DGCR8	CAUCGGACAAGAGUGUGAU.dTdT
si Drosha	Drosha	CGAGUAGGCUUCGUGACUU.dTdT
si CPSF-73	CPSF3	ON-TARGETPLUS SMARTPOOL CAT NO. L-006365-01
si Dicer	Dicer	UGCUUGAAGCAGCUCUGG.dTdT
Primers for pri-miR-122 cloning and mutagenesis		**Sequence (5'-3')**
Open_Beta qR		AGCTTTATTGAGGCTTAAGC
B10 qF		CAGGAAACTATTACTCAAAGGG
Pri122 qF		CAGAAGCTGTGGAAGGCGC
Pri122 qR		ACTGTGTTCTCTGCTGCCTC
DELTA mir qF		GTTTCCTTAGCAGAGCTGCAATCCTTCCCTCGATAA
DELTA mir qR		TTATCGAGGGAAGGATTGCAGCTCTGCTAAGGAAAC
pA1MTqF		CAATCTTGTGTACTTACTGAAGAAAGTCTGGCTCTTTTGCACT
pA1MTqR		AGTGCAAAAGAGCCAGACTTTCTTCAGTAAGTACACAAGATTG