Table S1. Conserved miRNAs encoded by newly annotated pri-miRNAs.

	Illumina BodyMap 2.0	**Human cell lines**	**Mouse cell lines**
Class I Independent noncoding transcription units	miR-23a/24-2/27a[b] miR-101-1/3671 miR-141/200c[b] miR-142[a] miR-193b/365a[b] miR-219-2[a] miR-223[c]	let-7a-1/7f-1/7d let-7i miR-10b miR-23a/24-2/27a[b] miR-29c/29b-2 miR-30a/30c-2[b] miR-30b/30d[b] miR-34a miR-92b[b] miR-101-1/3671 miR-129-2 miR-130a miR-130b/301b miR-132/212[b] miR-138-1 miR-141/200c[b] miR-144/451a/4732[b] miR-146a/3142[b] miR-148a[b] miR-187[b] miR-192/194-2[b] miR-193b/365a[b] miR-194-1/215 miR-200a/200b/429 miR-221/222 miR-302a/302b/302c/302d/367	let-7a-1/7f-1/7d let-7i miR-7a-2 miR-17/18a/19a/20a/19b-1/92a-1 miR-31 miR-129-1[a] miR-129-2 miR-130a miR-133b/206[a] miR-137 miR-138-1 miR-138-2[a] miR-142[a] miR-150[a] miR-155 miR-191/425 miR-194-1/215 miR-199a-1[a] miR-219-2[a] miR-221/222 miR-302a/302b/302c/302d/367 miR-384 miR-670[a] miR-3074-1[a]
Class II Extension of existing protein-coding transcripts	miR-21 miR-505	miR-7-2 miR-21 miR-34b/34c miR-181c/181d[b] miR-196a-1 miR-219-1 miR-324 miR-505	miR-10a miR-34b/34c miR-196a-1 miR-196a-2[a] miR-196b miR-200a/200b/429 miR-219-1 miR-320[a] miR-324 miR-331[a] miR-345[a] miR-505
Class III Extension of existing non-coding transcripts	miR-29a/29b-1 miR-370	let-7e/miR-99b/miR-125a miR-9-3 miR-29a/29b-1 miR-296/298 miR-370	let-7e/miR-99b/miR-125a miR-9-2 miR-18b/19b-2/20b/92a-2/106a/363[a] miR-29a/29b-1 miR-296/298

[a] not mapped in human cell lines

[b] not mapped in mouse cell lines

[c] not mapped in human and mouse cell lines

Table S2. Evaluation of the performance of four transcriptome assembly programs on pri-miRNAs that are annotated in Refseq.

Program	**Number of predicted pri-miRNA transcripts matching the RefSeq annotation**	**Number of RefSeq pri-miRNAs for which at least one transcript was assembled correctly by the program**	**Running Time (hours:minutes :seconds)**
StringTie	561	467	1:13:23
Cufflinks	378	337	21:01:08
IsoLasso	90	82	14:36:04
Scripture	293	200	65:57:32

Note: There are 788 Refseq genes (1836 transcripts) that overlap 876 miRNAs annotated in miRBase release 20 (out of 1871 total miRNAs).

Table S3. RNAseq mapping statistics.

	Cell type	**Read count**	**Mapping frequency**
Human	A172	184,705,740	92.50%
	A673	174,578,382	93.40%
	Fibroblast	142,718,780	92.20%
	HCT116	150,638,560	90.20%
	HEK293	193,346,087	86.70%
	HepG2	221,060,288	91.10%
	MCF7	160,067,256	91.00%
	NCCIT	165,209,310	93.30%
Mouse	C2C12	163,248,130	92.50%
	CT-26	215,970,827	90.90%
	E14TG2a	211,111,824	91.00%
	Hepa1-6	150,418,313	93.40%
	MEF	200,927,640	90.20%
	Neuro-2a	193,572,149	89.50%

Table S4. Novel potential regulatory mechanisms for conserved human non-protein coding pri-miRNAs.

Encoded human miRNA(s)	Multiple promoters	Partial production of cluster	miRNA spans splice site
let-7a-1/let-7f-1/let-7d	Yes		
let-7a-3/let-7b	Yes		
let-7c/miR-99a/miR-125b-2	Yes		
miR-9-2	Yes		
miR-9-3	Yes		
miR-15a/miR-16-1	Yes		
miR-17/miR-18a/miR-19a/miR-20a/miR-19b-1/miR-92a-1	Yes		
miR-22	Yes		
miR-23a/miR-24-2/miR-27a	Yes		
miR-29a/miR-29b-1	Yes		
miR-30b/miR-30d	Yes		
miR-31	Yes		
miR-101-1/miR-3671	Yes		
miR-130a	Yes		
miR-135a-2/miR-1251	Yes		
miR-135b	Yes		
miR-137/miR-2682	Yes		
miR-181c/miR-181d	Yes		
miR-193b/miR-365a	Yes		
miR-195/miR-497	Yes		
miR-221/miR-222	Yes		
miR-675	Yes		
let-7a-2/miR-100/miR-125b-1	Yes	Yes	
miR-30a/miR-30c-2	Yes	Yes	
miR-374a/miR-374b/miR-421/miR-545	Yes	Yes	
miR-132/miR-212		Yes	
miR-130b/miR-301b			Yes (miR-130b)
miR-199a-2/miR-214			Yes (miR-199a-2)
miR-202			Yes
miR-205			Yes

Table S5. Novel potential regulatory mechanisms for conserved mouse non-protein coding pri-miRNAs.

Encoded mouse miRNA(s)	Multiple promoters	Partial production of cluster	miRNA spans splice site
let-7b/let-7c-2	Yes		
miR-15a/miR-16-1	Yes		
miR-17/miR-18a/miR-19a/miR-20a/miR-19b-1/miR-92a-1	Yes		
miR-29a/miR-29b-1	Yes		
miR-31	Yes		
miR-101a	Yes		
miR-196b	Yes		
miR-221/miR-222	Yes		
miR-345	Yes		
miR-374/miR-421	Yes		
let-7a-2/miR-100/miR-125b-1	Yes	Yes	
miR-670			Yes

Table S6. Sequences of oligonucleotides used in this study

Primer sequences for mutagenesis

Mutation	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')
E1045Q	AAGCGTTAATAGGAGCTGTTTACTTGGAGGGAAG	GAAAACAATTGGCCATTGCATGTCGAAGGTCCG
E1222Q	AATCATTTATTGCAGCGCTGTACATTGATAAGGATTTGGAATATG	GCAAAAGGTCCGCCAAGGTCTTGGTGCGAAG
Synonymous mutation T438-L444	GAGAGATCTGTATGACAAATTTGAGGAGGAGTTGGGGAGC	AATCGTGATGTTCCAACCACTGTAGAATCTCCCACCTG

Primer sequences for real-time RT-PCR

Gene/Amplicon	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')
Human DLEU2	TGCATTGGAACATGACATGAG	AAGAATTGCTGAGCTAAGTAGAGGTC
Human C13orf25	GGCCTCCGGTCGTAGTAAAG	GCAGTTAGGTCCACGTGTATGA
Human pri-miR-15a	TAGGCGCGAATGTGTGTTTA	TGCTATCATAAGAGCTATGAATAAAAAG
Human pri-miR-16-1	CTTTTTATTCATAGCTCTTATGATAGC	TCAATAAAACTGAAAACACATTAGTAACA
Human pri-miR-17	CACCTTGTAAAACTGAAGATTGTGA	CCTGCACTTTAAAGCCCAACT
Human pri-miR-18a	AGGGCCTGCTGATGTTGAGT	AACACCTATATACTTGCTTGGCTTG
18S rRNA	GTAACCCGTTGAACCCCATT	CCATCCAATCGGTAGTAGCG

Primer sequences for RACE in Fig S5

Amplicon	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')
5'RACE from exon A	CGACTGGAGCACGAGGACACTGA	CGCTCGCCTGACAGCTGATG
Nested 5'RACE from exon A	GGACACTGACATGGACTGAAGGAGTA	GCAGGAGGAGGAGGGGAGAA
3'RACE from exon B	ATCCCTCCCTGTCACACACG	GCTGTCAACGATACGCTACGTAACG
Nested 3'RACE from exon B	GATGGGTGGTCGCTTACCTGTG	CGCTACGTAACGGCATGACAGTG
5'RACE from exon C	CGACTGGAGCACGAGGACACTGA	TGCTCTAAAGTCTGCTCCCAGAGAGG
Nested 5'RACE from exon C	GGACACTGACATGGACTGAAGGAGTA	CTGCTCCCAGAGAGGACTTGT
3'RACE from exon D	TGGCGCCACTTTCCTGAGAT	GCTGTCAACGATACGCTACGTAACG
Nested 3'RACE from exon D	ACTTCCAGCCAGTTTGGGTCA	CGCTACGTAACGGCATGACAGTG

Primer sequences for RACE in Fig S10

Amplicon	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')
5'RACE from exon A	CGACTGGAGCACGAGGACACTGA	CCACACGCACCTCCTGGTTG
Nested 5'RACE from exon A	GGACACTGACATGGACTGAAGGAGTA	TGTCTTGGTTCTGTCTGTCTGATG
5'RACE from exon D	CGACTGGAGCACGAGGACACTGA	AAACCTGCTTCCATCTTGTTAGGC
Nested 5'RACE from exon D	GGACACTGACATGGACTGAAGGAGTA	GGCTAATATCTTCAAATCATCCACACG

5'RACE from exon E	CGACTGGAGCACGAGGACACTGA	GTGGCACCATCCCGAGCAAG
Nested 5'RACE from exon E	GGACACTGACATGGACTGAAGGAGTA	AGAGCTCTCAGTGCGCTAGG

Primer sequences for RACE in Fig S12

Amplicon	**Forward primer sequence (5'-3')**	**Reverse primer sequence (5'-3')**
5'RACE from exon A	CGACTGGAGCACGAGGACACTGA	AGGCCCTCAGCTAGCGGTCTG
Nested 5'RACE from exon A	GGACACTGACATGGACTGAAGGAGTA	GGTCTGAGTCCTGGGTTCCAAA
5'RACE from exon B	CGACTGGAGCACGAGGACACTGA	CGGAGGATGGAGGCGTCTTCT
Nested 5'RACE from exon B	GGACACTGACATGGACTGAAGGAGTA	CCAAAGCCAGGAAGTGAAAATGA
5'RACE from exon C	CGACTGGAGCACGAGGACACTGA	AAATGCGGCCACACGGACTTT
Nested 5'RACE from exon C	GGACACTGACATGGACTGAAGGAGTA	GGCCACACGGACTTTGAAGG

Primer sequences for RT-PCR

Primer name	**Primer sequence (5'-3')**
507	GAGTAGGCGCGTGGAGTC
508	TCTTGCACGATCAAAATAGGG
509	GCCACATGTGATAGATGACCA
510	GGGTGATCCTTTGCCTTCT
511	CAGGCAGGACGAGAGAAAGA
513	TGCAATGTAAGCTTCTGTTTCC
514	GGGGAGAGGATGGAGAGC
515	TCATTTTCTCCGCAGCATC
517	CGAGCTCAGTTATGGCACAC
518	GGGAGTCTAAGGGCAGCAG
519	TGCTGCTGCTGCTGCTAC
520	TAGCGGGAAGAACAAAGGAA
521	GGGACGCTGGAGTCTGG
522	TTCTGGTGGCTGCATTACTCT
523	GGAGAGAGGAAGAGCGGAGT
524	AAAGGCGCTTCTTTTCACCT
525	CCTGTCAGTCACCGTGTCC
552	AAGAGGGTGAGCGTTTGGA
553	CCAGGGACGTCATTTTCACT
554	CCCTTCAAAGTCCGTGTGG
555	GGTGGCTAGGTGACAGGAGA
556	GGGTGACTTTCTCGACTCGT
557	CTGGCCCATGTCTCTCTGTT
559	CAAGACATCTGAGGGGCAAC
560	GCAGAGGAGGTGTCTTCAGG
561	CACTAGTGTCTCCCCTGCTTC
563	CAGCCTAGCGCACTGAGAG
565	GTCCTCTCTGGGAGCAGACTT
566	TTTGAACCATGAATTCCACCT
575	TCTTTGGACAAAATTGAGAAGAACT

Table S7. Transfection methods.

Cell lines	**Transfection reagent**	**Molar Ratio of plasmids transfected***
A-172	Cell Line Kit V; Program U-029; Nucleofector 2b (Amaxa)	3:1
A-673	Cell Line Kit V; Program X-001; Nucleofector 2b (Amaxa)	3:1
C2C12	Cell Line Kit V; Program B-032; Nucleofector 2b (Amaxa)	3:1
CT-26	Cell Line Kit SE; Program CM-137; 4D-Nucleofector (Amaxa)	3:1
E14TG2a embryonic stem cells	Xfect (Clontech)	3:1
HCT116	FuGENE HD (Promega)	2:1
HEK293T	FuGENE HD (Promega)	1:1
Hepa1-6	Cell Line Kit SF; Program EH-100; 4D-Nucleofector (Amaxa)	3:1
HepG2	FuGENE HD (Promega)	3:1
Human primary fibroblasts	FuGENE HD (Promega)	3:1
MCF-7	Lipofectamine LTX (Life Technologies)	4:1
MEF	MEF Kit 2; Program T-020; Nucleofector 2b (Amaxa)	3:1
Neuro-2a	Cell Line Kit V; Program T-024; Nucleofector 2b (Amaxa)	3:1
NCCIT	FuGENE HD (Promega)	4:1

*Molar ratio of pcDNA3.1/V5-His-TN-DROSHA to pcDNA5/FLAG-HA-DGCR8