Supporting Information

Londin et al. 10.1073/pnas.1420955112

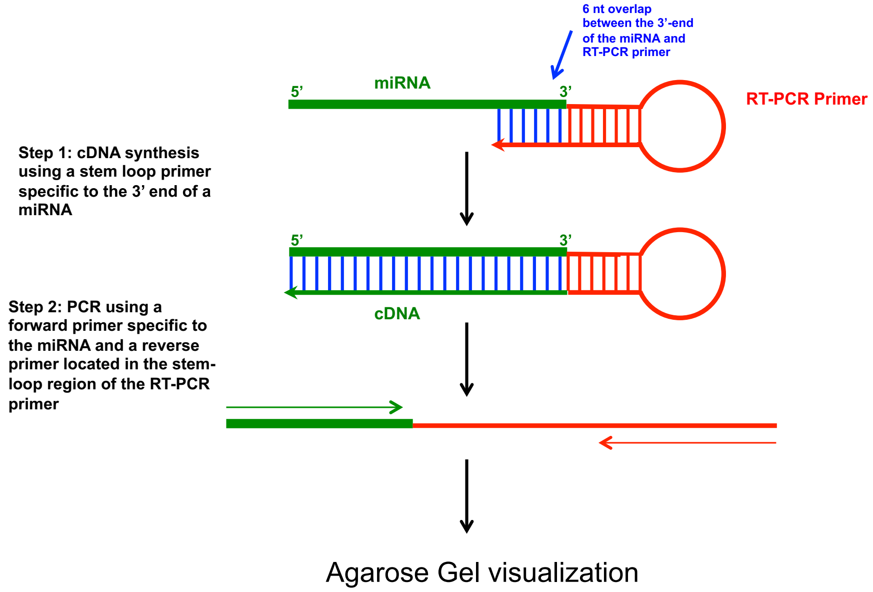

Fig. S1. PCR schematic for amplification of novel miRNAs. cDNA synthesis occurs using a stem-loop primer specific to the 3′ end of a miRNA with a 6 nt overlap between the between miRNAs and RT-PCR primer. The PCR is carried out using a forward primer specific to the miRNA and a reverse primer specific to the stem-loop region of the RT-PCR primer. The resulting output is visualized on an agarose gel.

Table S1. Among the novel miRNAs are seed-paralogues of “popular” miRNAs

	Known miRNAs sharing seed		Novel miRNAs that are seed-paralogues	
Seed	How many?	Name	How many?	Name
AAAGTG	3	miR-17–5p miR-20b-5p miR-20a-5p	1	TJU_CMC_MD2.ID02276.3p-miR
AACACT	1	miR-200a-3p	1	TJU_CMC_MD2.ID00971.5p-miR
AATCAC	1	miR-34b-3p	1	TJU_CMC_MD2.ID03005.3p-miR
ACCCTG	2	miR-10b-5p miR-10a-5p	3	TJU_CMC_MD2.ID00662.3p-miR TJU_CMC_MD2.ID00574.5p-miR TJU_CMC_MD2.ID01550.3p-miR
AGCACC	1	miR-29a-3p	2	TJU_CMC_MD2.ID02353.3p-miR TJU_CMC_MD2.ID00524.3p-miR
AGCAGC	3	miR-16–5p miR-15b-5p miR-15a-5p	9	TJU_CMC_MD2.ID03212.5p-miR TJU_CMC_MD2.ID01415.3p-miR TJU_CMC_MD2.ID03136.3p-miR TJU_CMC_MD2.ID00069.3p-miR TJU_CMC_MD2.ID00807.5p-miR TJU_CMC_MD2.ID01398.3p-miR TJU_CMC_MD2.ID01521.3p-miR TJU_CMC_MD2.ID00930.3p-miR TJU_CMC_MD2.ID01325.3p-miR
AGCTTA	2	miR-21–5p miR-590–5p	3	TJU_CMC_MD2.ID02864.5p-miR TJU_CMC_MD2.ID03495.5p-miR TJU_CMC_MD2.ID01413.3p-miR
AGGCAG	1	miR-34b-5p	5	TJU_CMC_MD2.ID02203.5p-miR TJU_CMC_MD2.ID02387.5p-miR TJU_CMC_MD2.ID01241.3p-miR TJU_CMC_MD2.ID01229.3p-miR TJU_CMC_MD2.ID03410.5p-miR
AGGCCA	1	miR-15a-3p	2	TJU_CMC_MD2.ID01813.5p-miR TJU_CMC_MD2.ID01250.3p-miR
ATCTTA	1	miR-200a-5p	1	TJU_CMC_MD2.ID02273.3p-miR
ATTGCA	2	miR-92a-3p miR-92b-3p	1	TJU_CMC_MD2.ID03303.3p-miR
CAATAT	1	miR-16–2-3p	1	TJU_CMC_MD2.ID01975.3p-miR
CCCTGA	5	miR-125a-5p miR-125b-5p miR-345–3p miR-4319 miR-4732–3p	3	TJU_CMC_MD2.ID01085.3p-miR TJU_CMC_MD2.ID01576.3p-miR TJU_CMC_MD2.ID02482.3p-miR
CTATTC	2	miR-26a-1–3p miR-26a-2–3p	1	TJU_CMC_MD2.ID02575.3p-miR
CTGCAG	1	miR-17–3p	6	TJU_CMC_MD2.ID02050.3p-miR TJU_CMC_MD2.ID01376.5p-miR TJU_CMC_MD2.ID03246.3p-miR TJU_CMC_MD2.ID00085.3p-miR TJU_CMC_MD2.ID02871.5p-miR TJU_CMC_MD2.ID01424.5p-miR
CTGCCC	1	miR-18a-3p	6	TJU_CMC_MD2.ID03302.3p-miR TJU_CMC_MD2.ID00831.5p-miR TJU_CMC_MD2.ID01696.3p-miR TJU_CMC_MD2.ID02348.3p-miR TJU_CMC_MD2.ID02192.5p-miR TJU_CMC_MD2.ID03001.3p-miR
GAGGTA	6	let-7a-5p let-7f-5p let-7g-5p let-7b-5p let-7e-5p let-7d-5p	1	TJU_CMC_MD2.ID00832.5p-miR

Table S1. Cont.

Seed	Known miRNAs sharing seed: How many?	Known miRNAs sharing seed: Name	Novel miRNAs that are seed-paralogues: How many?	Novel miRNAs that are seed-paralogues: Name
GCAGCA	2	miR-103a-3p miR-107	5	TJU_CMC_MD2.ID02151.5p-miR TJU_CMC_MD2.ID00091.5p-miR TJU_CMC_MD2.ID00676.3p-miR TJU_CMC_MD2.ID01372.3p-miR TJU_CMC_MD2.ID02652.5p-miR
GCTTCT	1	miR-103a-2–5p	2	TJU_CMC_MD2.ID00400.5p-miR TJU_CMC_MD2.ID02736.5p-miR
GGCAGT	2	miR-34c-5p miR-34a-5p	4	TJU_CMC_MD2.ID00706.5p-miR TJU_CMC_MD2.ID02205.5p-miR TJU_CMC_MD2.ID01026.3p-miR TJU_CMC_MD2.ID01982.5p-miR
GGTTGG	1	miR-92a-1–5p	4	TJU_CMC_MD2.ID01723.5p-miR TJU_CMC_MD2.ID00051.3p-miR TJU_CMC_MD2.ID00293.3p-miR TJU_CMC_MD2.ID03385.3p-miR
GTGCAA	1	miR-19b-3p	2	TJU_CMC_MD2.ID03086.5p-miR TJU_CMC_MD2.ID00745.5p-miR
TCCTAC	1	miR-155–3p	2	TJU_CMC_MD2.ID02078.5p-miR TJU_CMC_MD2.ID01758.3p-miR

The table lists alphabetically the seeds contained in the mature miRNAs of several precursors that are frequently cited in the literature. For each seed, we show the known mature miRNA(s) containing the seed and the identifiers of the seed-paralogues among the newly discovered miRNAs. It is worth noting that a few of the shown miRNAs each have several seed-paralogues among the newly discovered miRNAs.

Table S2. Precursor/mature conservation for a less stringent threshold

Genome where present	No. of precursors:mature combinations	No. of precursors:mature combinations whose novel miRNAs have support in Ago CLIP-seq	No. of miRBase precursors:mature combinations
Human	3,707	1,657	2,772
Chimpanzee	1,310 (35.3%)	558 (33.7%)	2,163 (78.0%)
Gorilla	1,371 (37.0%)	603 (36.4%)	2,334 (84.2%)
Orangutan	1,136 (30.6%)	468 (28.2%)	2,005 (72.3%)
Macaque	807 (21.8%)	351 (21.2%)	1,881 (67.9%)
Mouse	262 (7.1%)	136 (8.2%)	778 (28.1%)
Drosophila	103 (2.8%)	44 (2.7%)	76 (2.7%)
Worm	81 (2.2%)	36 (2.2%)	50 (1.8%)

In Table 2, we showed conservation results for searches that required the presence of at least 85% of the mature miRNA sequence (with 100% of the seed matching) and the presence of at least 85% of the precursor sequence. This table shows conservation results for searches that required the presence of at least 85% of the mature miRNA sequence (with 100% of the seed matching) and the presence of at least 50% of the precursor sequence: i.e., more tolerant than the setting in Table 2. The results are largely unchanged, with the great majority of the newly identified sequences remaining primate-specific.

Table S3. GO terms for RNA22 predicted targets for four miRNAs

mir-19a-3p	mir-19b-3p	TJU_CMC_MD2.ID00745-5p	TJU_CMC_MD2.ID03086-5p	Term
—	—	—	2.405349619	Negative regulation of protein-complex disassembly
—	—	—	2.354084708	Regulation of protein-complex disassembly
—	—	—	1.878832007	Microtubule-associated complex
—	—	1.739631004	1.421497967	Cell-projection organization
—	—	1.798758987	—	Dendrite
—	—	1.78423613	—	Plasma-membrane part
—	—	1.570531031	—	Protein-kinase activity
1.603993145	1.504485823	—	—	Nucleoside-triphosphatase regulator activity
1.614881337	1.507646301	—	—	GTPase regulator activity
1.670034031	1.622729454	—	—	Regulation of small GTPase-mediated signal transduction
1.690154923	1.649222996	—	—	Regulation of Ras protein signal transduction
1.740009706	1.626187051	—	—	Small GTPase regulator activity
1.947991476	1.963241996	—	—	Guanyl-nucleotide exchange factor activity
—	2.043923174	—	—	Ras guanyl-nucleotide exchange factor activity
—	2.099164341	—	—	Rho guanyl-nucleotide exchange factor activity

All four of these miRNAs share the same seed sequence: GTGCAA. Cells marked by ”−“ indicate that the associated GO terms were not identified as being significant.

Table S4. Primer sequences used for the miRNA PCR validations performed in Fig. 5

miRNA	miRNA sequence (5′→ 3′)	Length of miRNA	Primers for cDNA synthesis	Primers for RT PCR
TJU_CMC_MD2.ID00126.5p-miR	CTCAGGACTGGGTCATTCTGTT	22	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAACAGA	CAGGACTGGGTCATTCTGTT
TJU_CMC_MD2.ID00259.5p-miR	AGCTACAGGTGAGATCAGGTTT	22	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAAACCT	CTACAGGTGAGATCAGGTTT
TJU_CMC_MD2.ID00400.5p-miR	AGCTTCTTTACAGTGTTGCCTTG	23	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCAAGGC	TTCTTTACAGTGTTGCCTTG
TJU_CMC_MD2.ID00480.5p-miR	GGCACTGACCACGTCACTGGAT	22	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACATCCAG	CACTGACCACGTCACTGGAT
TJU_CMC_MD2.ID00841.3p-miR	GCTAAAGGATATTTTGGGTTCAGA	24	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTCTGAA	AAGGATATTTTGGGTTCAGA
TJU_CMC_MD2.ID00851.5p-miR	CCCGATGCGGAGGCGGGTCC	20	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGGACCC	CCCGATGCGGAGGCGGGTCC
TJU_CMC_MD2.ID01125.3p-miR	CCATGCTGAGAATCGCTGTTC	21	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGAACAG	CATGCTGAGAATCGCTGTTC
TJU_CMC_MD2.ID01196.3p-miR	GCACACACGTCCATTCCCTACT	22	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAGTAGG	ACACACGTCCATTCCCTACT
TJU_CMC_MD2.ID01548.5p-miR	TGGCCCAGGTAGAGGGGTCCGCG	23	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCGCGGA	CCCAGGTAGAGGGGTCCGCG
TJU_CMC_MD2.ID01613.5p-miR	TCAAGTGATGGAGAGCAATA	20	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTATTGC	TCAAGTGATGGAGAGCAATA
TJU_CMC_MD2.ID01963.3p-miR	CTAGAGAACCATCTGAAAGACT	22	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAGTCTT	AGAGAACCATCTGAAAGACT
TJU_CMC_MD2.ID02236.3p-miR	GATGCCTGGGAGTTGCGATCTGC	23	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGCAGAT	GCCTGGGAGTTGCGATCTGC
TJU_CMC_MD2.ID02236.5p-miR	GACGAATTTTTGAGCGGGTAAA	22	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTTTACC	CGAATTTTTGAGCGGGTAAA
TJU_CMC_MD2.ID02704.5p-miR	CGGGCTGGCTTTAGCTCAGCGGT	23	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACACCGCT	GCTGGCTTTAGCTCAGCGGT
TJU_CMC_MD2.ID03061.3p-miR	TACTCTCTCGGACAAGCTGTAG	22	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCTACAG	CTCTCTCGGACAAGCTGTAG
TJU_CMC_MD2.ID03061.5p-miR	AATTACAGATTGTCTCAGAGA	21	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTCTCTG	ATTACAGATTGTCTCAGAGA
TJU_CMC_MD2.ID03327.3p-miR	CTCACTGATCAATGAATGCAA	21	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTTGCAT	TCACTGATCAATGAATGCAA
TJU_CMC_MD2.ID03399.5p-miR	TACGGTCAATGCTCTGAAATC	21	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGATTTC	ACGGTCAATGCTCTGAAATC
TJU_CMC_MD2.ID03403.5p-miR	CGCCGCGCTCTCGGGGCAGGGCC	23	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACGGCCCT	CGCGCTCTCGGGGCAGGGCC
TJU_CMC_MD2.ID03495.5p-miR	TAGCTTATCTCTGCAGCCTTGACA	24	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTGTCAA	TTATCTCTGCAGCCTTGACA
Reverse Primer in stem loop region	GTGCAGGGTCCGAGGT			

Red, the miRNA-specific region; blue, the stem-loop sequence.

Other Supporting Information Files

Dataset S1 (XLSX)
Dataset S2 (XLSX)
Dataset S3 (XLSX)
Dataset S4 (XLSX)
Dataset S5 (XLSX)