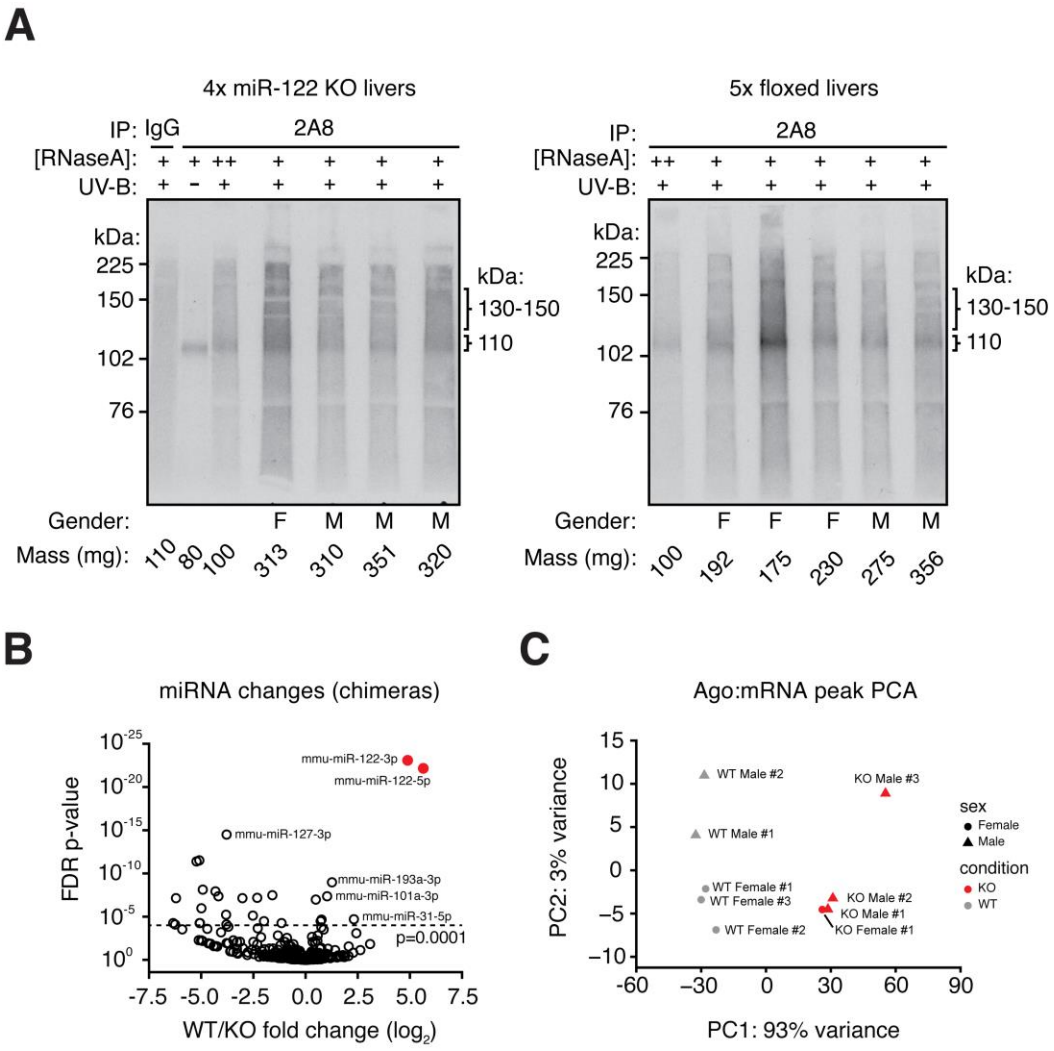
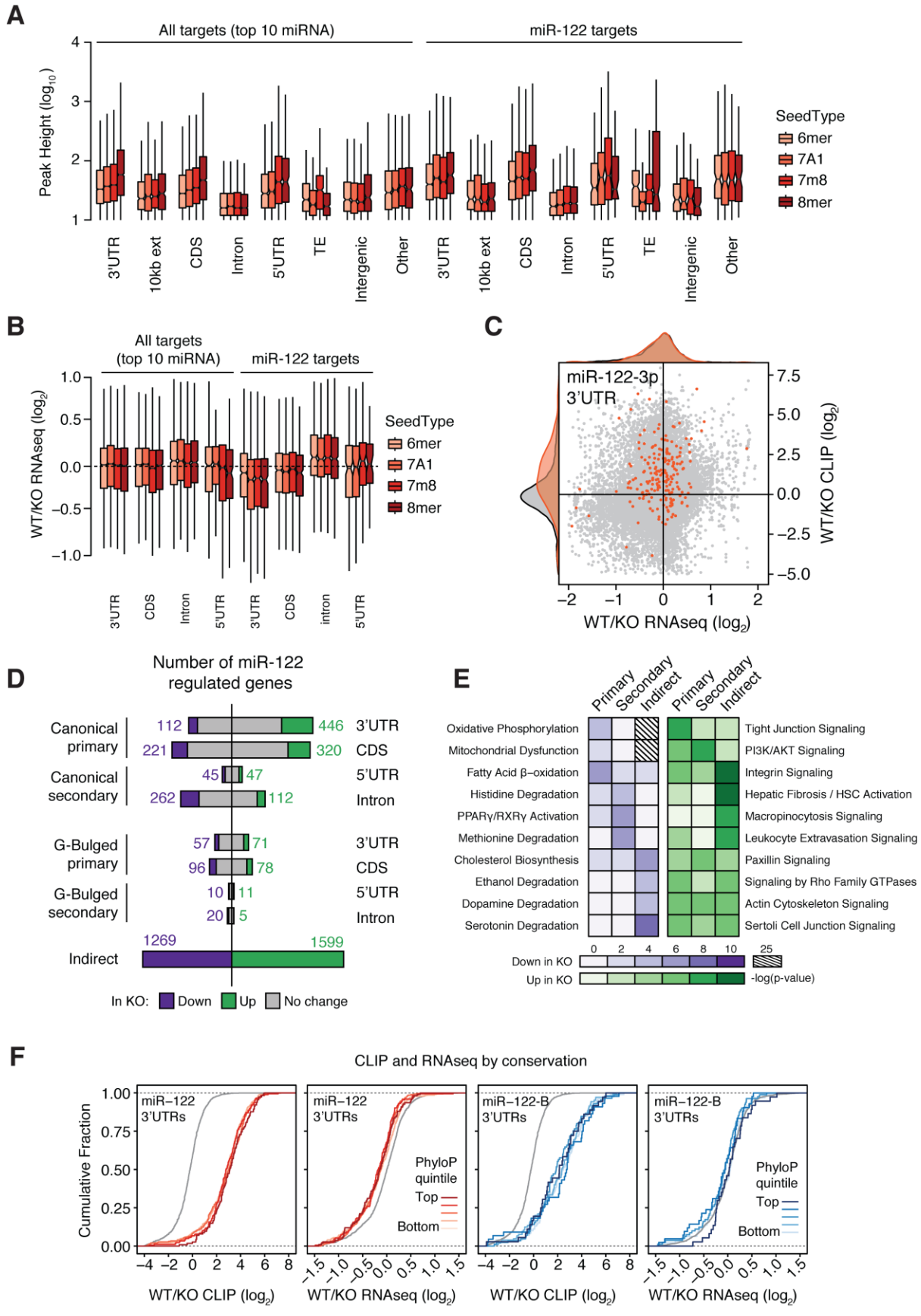
Supplemental Information

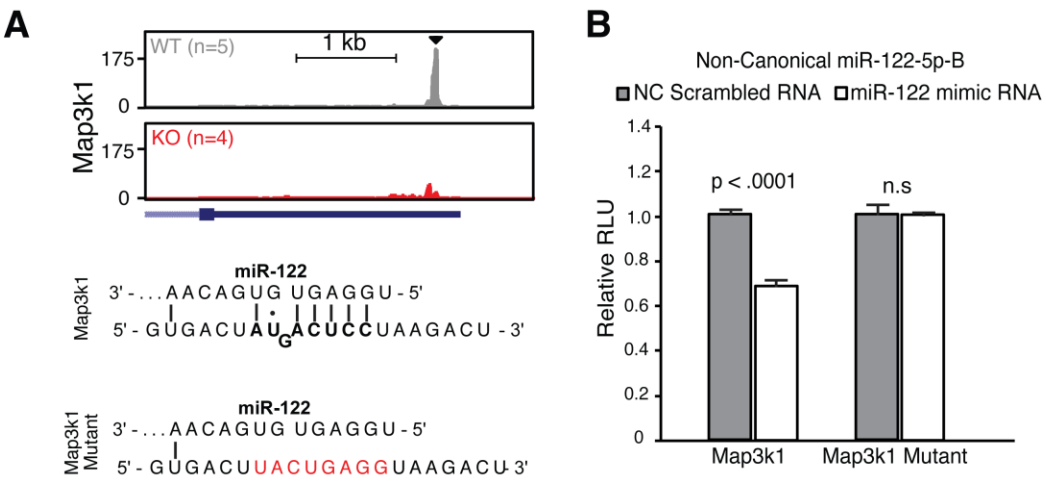
Supplemental Figures



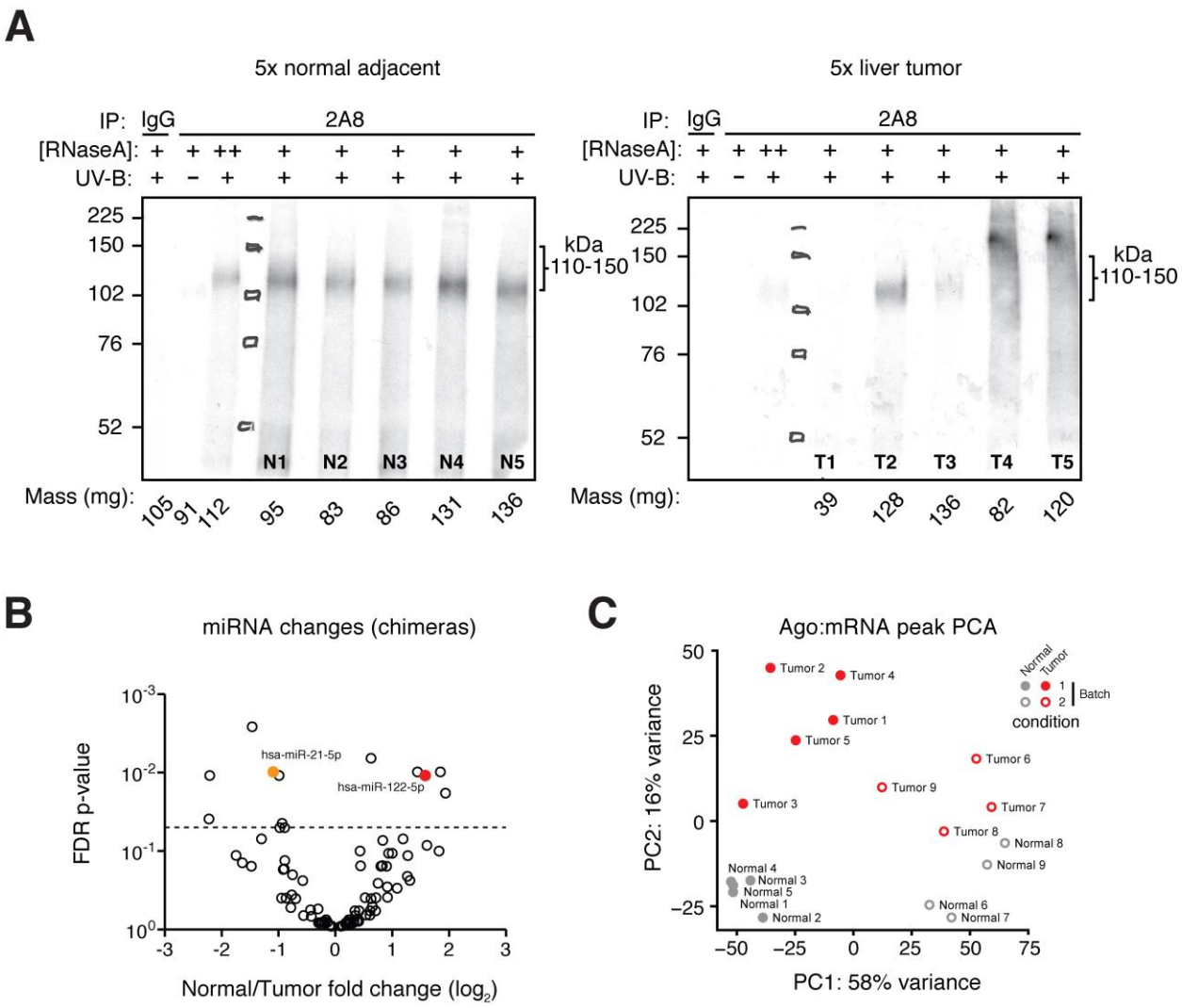
Supplemental Figure 1, related to Figures 1-2. Argonaute CLIP in WT and miR-122 KO mouse livers. (**A**) Autoradiogram of ^{32}P-labelled RNA crosslinked to IP purified AGO from four 122KO (left panel) or five WT floxed control livers (right panel) using the 2A8 anti-Ago antibody. Irrelevant antibody control "IgG", non-crosslinked control, and high "++" RNAse A control are shown. RNA-protein complexes of ~110kDa (miRNA) and ~130-150kDa (miRNA:mRNA) are indicated. A weak band of ~110kDa was typically observed under non-crosslinked conditions. (**B**) Volcano plot highlighting the $\log_2$ fold change of Ago associated miRNAs from miRNA:mRNA chimeras between WT and KO mouse livers. miR-122 arms highlighted in red. FDR, false discovery rate. **(C)** Principal component analysis (PCA) of significant Ago peaks on mRNAs segregates WT and KO livers.


A
All targets (top 10 miRNA)
miR-122 targets
Peak Height (log10)
SeedType
6mer
7A1
7m8
8mer
3'UTR
10kb ext
CDS
Intron
5'UTR
TE
Intergenic
Other
B
All targets (top 10 miRNA)
miR-122 targets
WT/KO RNAseq (log2)
C
miR-122-3p
3'UTR
WT/KO CLIP (log2)
WT/KO RNAseq (log2)
D
Number of miR-122 regulated genes
Canonical primary
Canonical secondary
G-Bulged primary
G-Bulged secondary
Indirect
3'UTR
CDS
5'UTR
Intron
In KO:
Down
Up
No change
E
Primary
Secondary
Indirect
Oxidative Phosphorylation
Mitochondrial Dysfunction
Fatty Acid β–oxidation
Histidine Degradation
PPARγ/RXRγ Activation
Methionine Degradation
Cholesterol Biosynthesis
Ethanol Degradation
Dopamine Degradation
Serotonin Degradation
Tight Junction Signaling
PI3K/AKT Signaling
Integrin Signaling
Hepatic Fibrosis / HSC Activation
Macropinocytosis Signaling
Leukocyte Extravasation Signaling
Paxillin Signaling
Signaling by Rho Family GTPases
Actin Cytoskeleton Signaling
Sertoli Cell Junction Signaling
Down in KO
Up in KO
-log(p-value)
F
CLIP and RNAseq by conservation
Cumulative Fraction
miR–122 3'UTRs
miR–122-B 3'UTRs
PhyloP quintile
Top
Bottom
WT/KO CLIP (log2)
WT/KO RNAseq (log2)

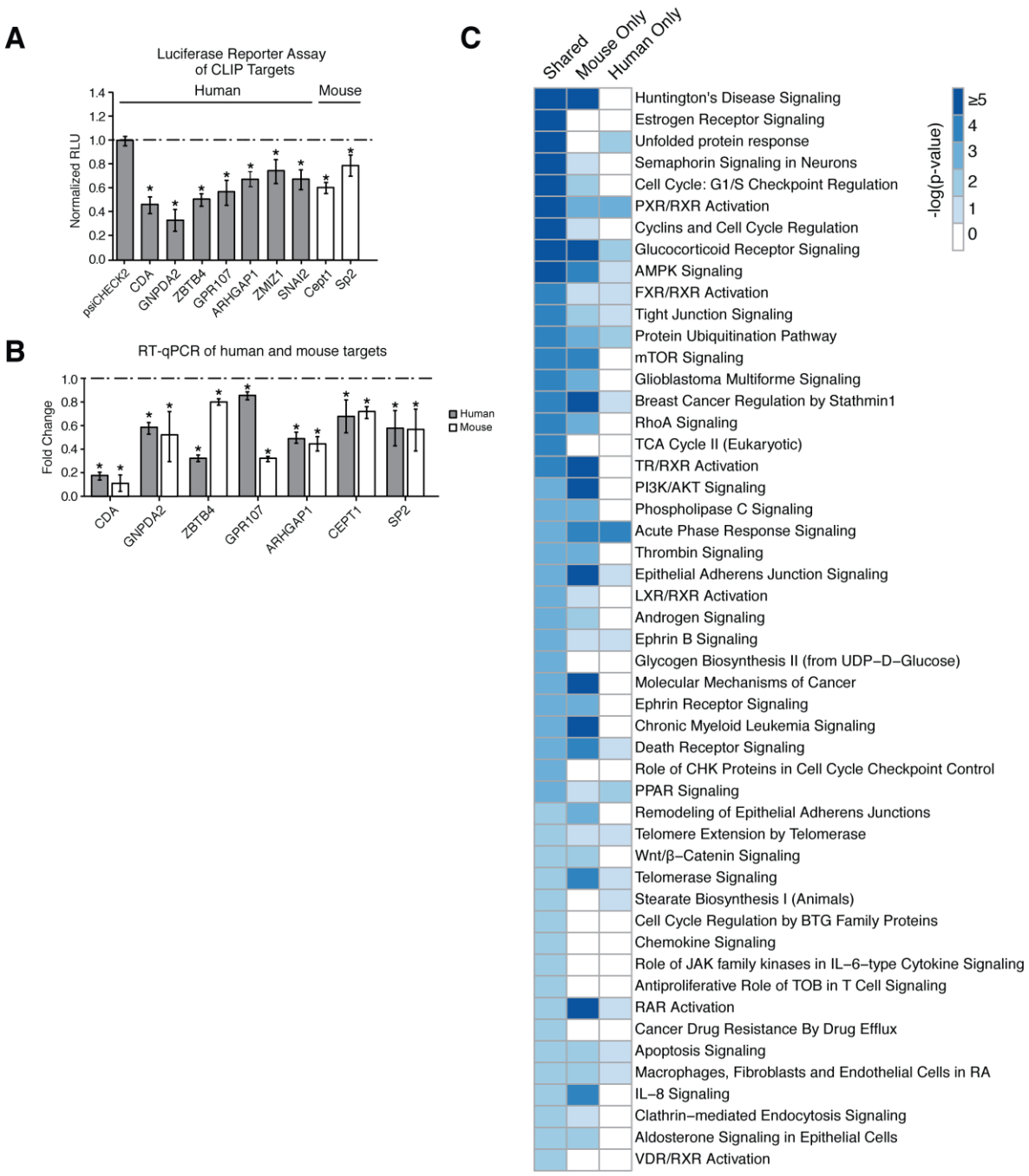
Supplemental Figure 2, related to Figure 2. General features of CLIP derived miR-122 targets in mouse liver. (A) CLIP peak height per for miRNA targets from the top 10 miRNAs or miR-122 binned by seed type. **(B)** Log_2 fold change in RNAseq expression between WT and KO livers for categories in (A). **(C)** Scatterplot with marginal histograms of log_2 fold change in CLIP binding and RNAseq expression between WT and KO livers for all 3'UTR loci containing miR-122-3p binding events, compared to top 20 miRNA targets. **(D)** Global RNAseq profiling of 3'UTR and CDS targets (primary), all other (secondary) miR-122 targets, and all significantly changing non-miR-122 targets (indirect), by number of genes. **(E)** Significantly regulated pathways from D. **(F)** Cumulative density function (CDF) of the log_2 fold change in CLIP or RNAseq binding for all canonical (red) or non-canonical (blue) miR-122 3'UTR targets binned by PhyloP conservation.



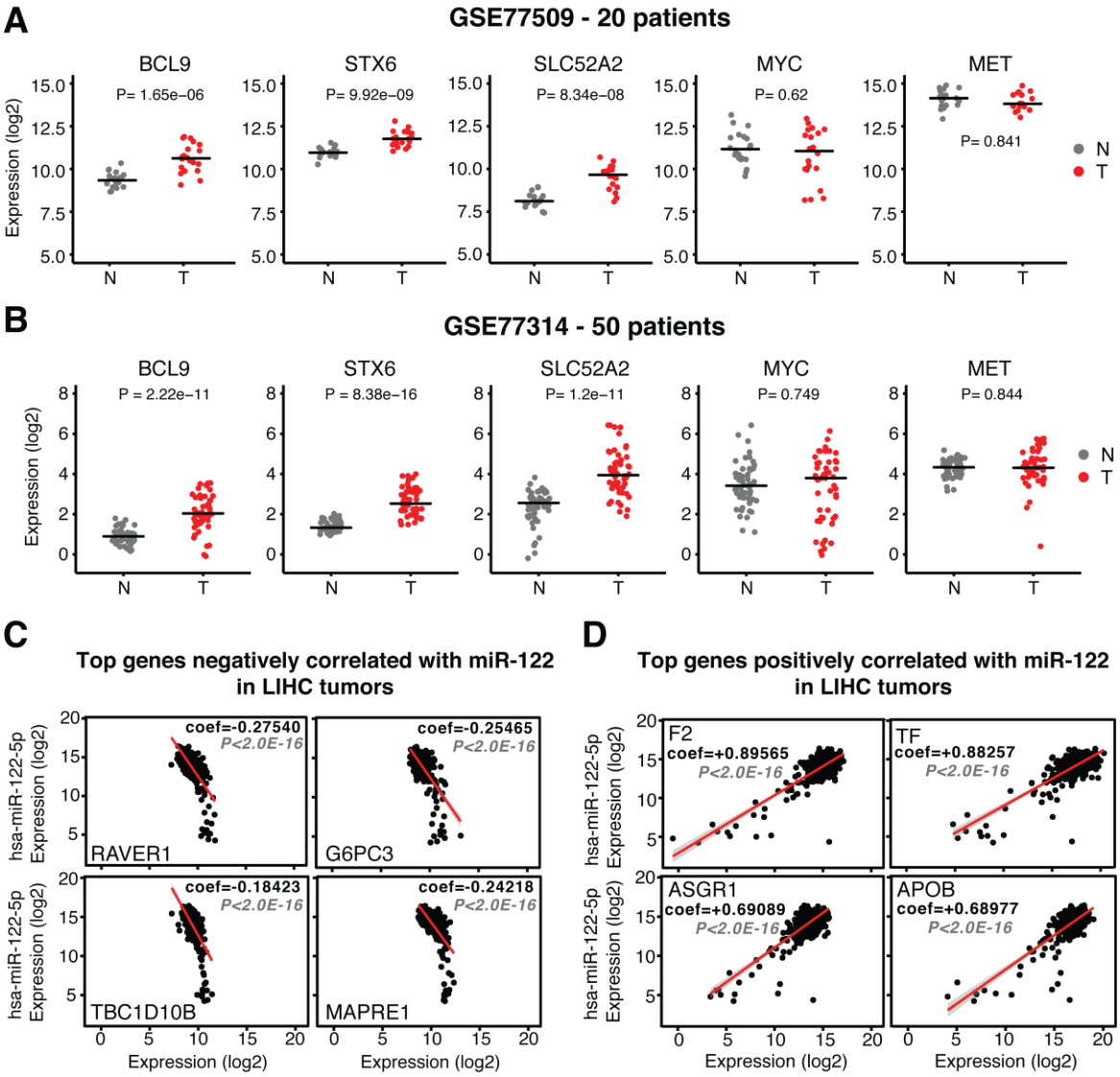
Supplemental Figure 3, related to Figure 2. **Validation of a mouse-specific G-bulged miR-122 target. (A)** Ago-CLIP binding profile of the Map3k1 transcript in WT and KO mice. miR-122 site is denoted by a black triangle. Bottom depicts the wild type and mutated bulged binding sites. **(B)** Luciferase reporter assay of mouse Map3k1. Renilla luciferase activity was normalized to Firefly luciferase (RLU) after transfection with 50nM miR-122 mimic or NC scrambled RNA. (P value < 0.001, two-tailed t-test).



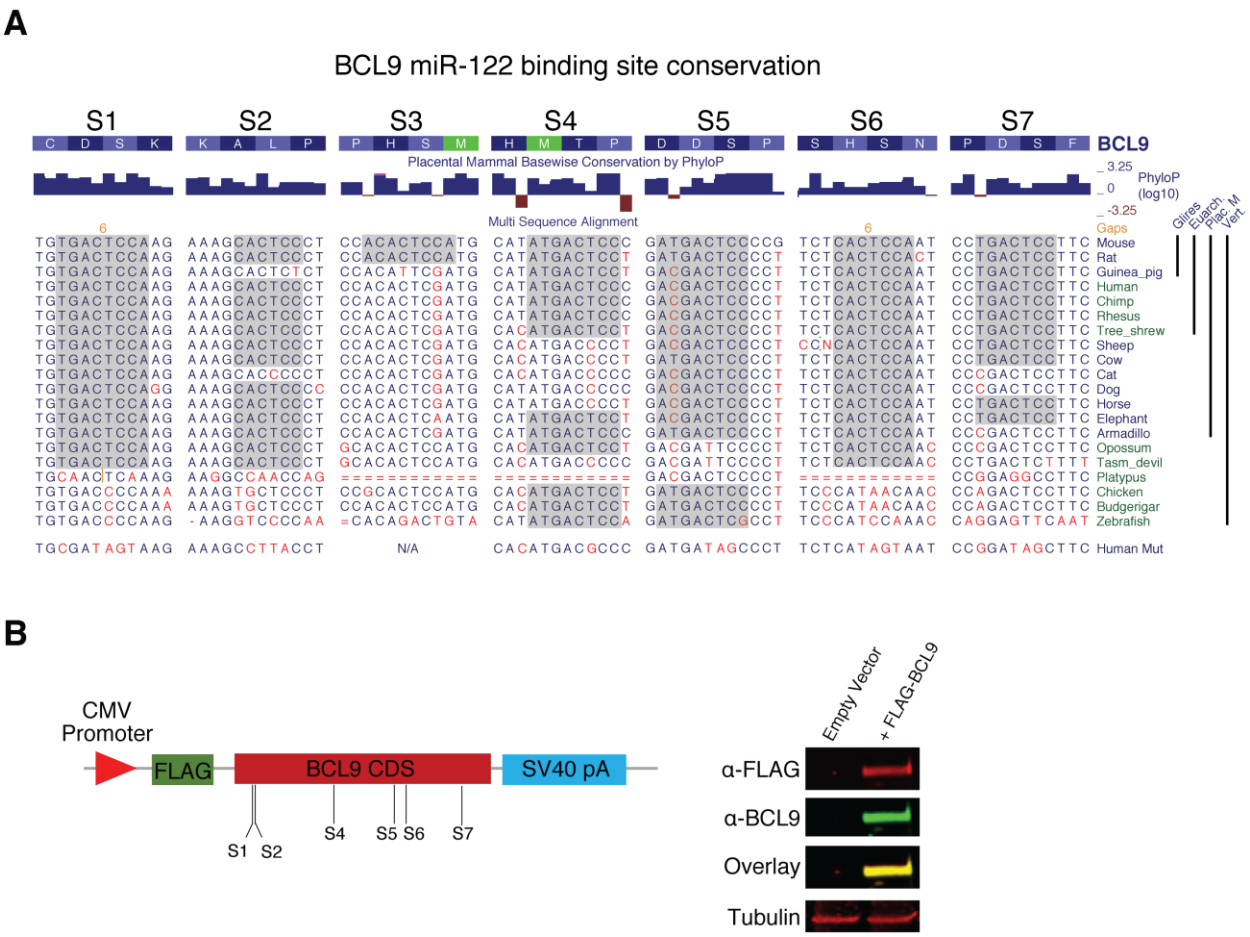
Supplemental Figure 4, related to Figure 3. **Argonaute CLIP in matched human normal adjacent and HCC tumor tissue.** (**A**) Autoradiogram of ^{32}P-labelled RNA crosslinked to IP purified AGO from five matched normal adjacent (left panel) and HCC tumor tissue (right panel) using the 2A8 anti-Ago antibody. Irrelevant antibody control "IgG", non-crosslinked control, and high "++" RNAseA control are shown. Excised RNA-protein complexes from ~110kDa-150kDa (miRNA:mRNA) are indicated. A weak band of ~110kDa was typically observed under non-crosslinked conditions. (**B**) Volcano plot highlighting the $\log_2$ fold change of Ago associated miRNAs from miRNA:mRNA chimeras between normal and tumor human tissue, across nine patients. miR-122-5p highlighted in red, miR-21-5p highlighted in orange. FDR, false discovery rate. **(C)** Principal component analysis (PCA) of significant Ago peaks on mRNAs segregates patient samples by type and by batch.



Supplemental Figure 5, related to Figure 3. Validation of miR-122 targets conserved in mouse and human, and pathway analysis (A) Luciferase Reporter Assay of mouse and human conserved STX6 and SLC52A2. Renilla luciferase activity was normalized to Firefly luciferase (RLU) after transfection with miR-122 mimic or NC scrambled RNA. (**B**) Luciferase reporter assay of select human miR-122 targets chosen for displaying conservation in mice. (**C**) Heatmap of the top 50 pathways targeted by miR-122 in mouse and human livers by IPA.



Supplemental Figure 6, related to Figure 4. miR-122 target derepression in HCC is not limited to TCGA RNAseq data. (A-B) Expression levels for BCL9, STX6, and SLC52A2 from HCC versus normal adjacent tissue RNAseq from previously published datasets: GSE77509 (A) and GSE77314 (B). MYC and MET genes were chosen as expression controls. P-values calculated by Mann-Whitney U-test. N, Normal; T, Tumor. **(C-D)** RSEM expression of miR-122 and its target genes in LIHC tumor samples. Linear regression coefficients of best fit line and significance P-values are shown**.** Genes were selected (n=4) based lowest (C) or highest (D) Pearson correlation value.



Supplemental Figure 7 related to Figures 6-7. Full BCL9 site conservation and generation of Flag-tagged BCL9 expression vector. (A) Full site conservation for all coding exon miR-122 binding sites on BCL9. Amino acid sequence, PhyloP conservation, and multi-sequence alignment are shown. Boxed regions indicate site conservation, with red nucleotides indicating deviation from mouse. Bottom human mutation sequence displays mutant miR-122 sites generated for (B). Species labeled at right from clade (glires) to phylum (vertebrates). **(B)** Schematic of human FLAG-BCL9 expression vector containing 6 conserved out of 7 total miR-122 sites. Immunoblot of H293T cells transfected with empty vector or FLAG-BCL9 expression vector and harvested after 48 hours. Blots were probed with anti-FLAG and anti-BCL9 antibodies.

Supplemental Tables

Supplementary Table 1, related to Figure 3: HCC subject characteristics

Patient	Age/Sex	Pathological Evaluation
1	58/F	HCC, moderately to poorly differentiated
2	33/M	Multi-focal HCC with intrahepatic cholangiocarcinoma.
3	59/M	Moderately differentiated HCC
4	63/M	HCC, moderately to poorly differentiated
5	60/M	HCC, moderately to poorly differentiated
6	44/F	Poorly-differentiated HCC
7	61/M	Well-differentiated HCC (grade I) with lymphovascular invasion
8	74/M	HCC clear cell variant, grade II
9	72/F	Grade III HCC with lymphatic/vascular invasion

Supplementary Table 2, related to Figures 5, 6 and 7: qPCR and cloning oligos used in this study

Primers for 3'UTR cloning into psiCHECK2		
Primer Name	**Sequence**	**5' end Location**
h-CDA-F	5'-CCG CTC GAG CAG AAG ACC CAG TGA CAG CC-3'	427
h-CDA-mut-F	5' TTT CCA GAT TAG TGA GGA GCC TGA GTC AGC 3'	562
h-CDA-mut-R	5' GCT GAC TCA GGC TCC TCA CTA ATC TGG AAA 3'	562
h-CDA-R	5' ATA AGA ATG CGG CCG CAA CAG GAT AGA ACC TTG GGA 3'	742
h-GNPDA2-F	5' CCG CTC GAG TAC TGA AGT CAC ATA GCC AC 3'	226
h-GNPDA2-R	5' ATA AGA ATG CGG CCG CTG CTA GGT ATA AAG CTC ATA 3'	703
h-GPR107-F	5' CCG CTC GAG GCT CTT AGA ACA TTA GAT AG 3'	3696
h-GPR107-R	5' ATA AGA ATG CGG CCG CCC ATC TCA GTG TGC TCC CTC 3'	4350
h-ZBTB4-F	5' CCG CTC GAG TTT TTT CCT GAT GGT TTC TCC C 3'	922
h-ZBTB4-R	5' ATA AGA ATG CGG CCG CTA AAG TCC CTG TAC CCC ATC CC 3'	1200
h-STX6-F	5' CCG CTC GAG TAA CTT TTG AAT TGC ACT TTT 3'	1561
h-STX6-R	5' ATA AGA ATG CGG CCG CAG TAC TTA GAC AAG GTT CAC A 3'	1909
h-SLC52A2-F	5' CCG CTC GAG CCC CAA CAC CTG TCT TTC CCT 3'	25
h-SLC52A2-R	5' ATA AGA ATG CGG CCG CTA CCC TAC AGC CCC CAC TGG T 3'	182
m-Arhgap1-F	5' CCG CTC GAG TGC TAA TAG GGA TGA TAA GA 3'	428
m-Arhgap1-mut1-F	5' TGG TCC TGT CGT GAG GCA GGG AGA GCC GTA 3'	586
m-Arhgap1-mut1-R	5' CGG CTC TCC CTG CCT CAC GAC AGG ACC ATT 3'	594
m-Arhgap1-mut2-F	5' AGT GTT GCT AGT GAG GAT TTA CTG AGG AGC 3'	1233

m-Arhgap1-mut2-R	5' CTC CTC AGT AAA TCC TCA CTA GCA ACA CTT 3'	1242
m-Arhgap1-R	5' ATA AGA ATG CGG CCG CTA GAT TCA CAG TAA GAG CGT 3'	1315
m-Cept1-de2-R	5' ATA AGA ATG CGG CCG CAT ACA AAA AAT AGC TAG AAT 3'	1286
m-Cept1-F	5' CCG CTC GAG CTT TCA TTC TTT GTT ACT GG 3'	116
m-Cept1-R	5' ATA AGA ATG CGG CCG CGT CTC TAG TAC CTC AGT TTA 3'	1319
m-Sp2-del-R	5' ATA AGA ATG CGG CCG CCA GGA GGA CTT AAT TGA GGG 3'	402
m-Sp2-F	5' CCG CTC GAG CAT CTG ATT GGC CCT GGG TC 3'	41
m-Sp2-R	5' ATA AGA ATG CGG CCG CTG GAA TGT TTG GAG GAG AGC 3'	524
m-Zmiz1-F	5' AGC TTT GTT TAA ACT CCA TCC AGA GCT CGC TCC A 3'	1279
m-Zmiz1-F	5' CCG CTC GAG CCC TCT GTC CCT GTG CTC CA 3'	1177
m-Zmiz1-R	5' ATA AGA ATG CGG CCG CTT CTC TGG AAG GAG ACA TTT 3'	2423
m-Zmiz1-R	5' ATA AGA ATG CGG CCG CAG GGC TTC CAT TAG GGA GAA 3'	2902
m-Stx6-F	5' CCG CTC GAG CGC CCG CCT CAG GCA CTT CTG 3'	9
m-Stx6-R	5' ATA AGA ATG CGG CCG CGC TGT CAG TTC TTC TGA GGA C 3'	292
m-Slc52a2-F	5' CCG CTC GAG CCC AGA CGT GTA GGA GTT ACT 3'	1698
m-Slc52a2-R	5' ATA AGA ATG CGG CCG CAT TGG ATG AAG ATG TTT TAT T 3'	1941
Primers for qPCR analysis		
Primer Name	**Sequence**	N/A
h-ACTB-RT-F	5'- CTGGCACCACACCTTCTACAATG -3'	N/A
h-ACTB-RT-R	5'- TAGCACAGCCTGGATAGCAACG -3'	N/A
h-STX6-RT-F	5'- ATCGTGGAACAGCAGGATGAGC -3'	N/A
h-STX6-RT-R	5'- AAGAGGATGGCTATGGCACACC -3'	N/A
h-SLC52A2-RT-F	5'- TTCCAGGGTCTTCTGCTGCTGTTG -3'	N/A
h-SLC52A2-RT-R	5'- ACTCTTCCACCTCTTCCTCTGCTC -3'	N/A
h-CDA-RT-F	5'- ACCCAGTGACAGCCAGAGAATG -3'	N/A
h-CDA-RT-R	5'- GAAACATCATCTTTGCCCAGTCC -3'	N/A
h-GNPDA2-RT-F	5'- TGCGTCACCGTAATGAGGCTTG -3'	N/A
h-GNPDA2-RT-R	5'- ACTCCCTGTTGGTAAACCCAGTG -3'	N/A
h-ZBTB4-RT-F	5'- CGGTTTCATCTTCAGCAAAAGC -3'	N/A
h-ZBTB4-RT-R	5'- TGAGCAGGGAACTTGGTGTCTC -3'	N/A
h-GPR107-RT-F	5'- TGGTGGTGAATGTCAGTAGCCTC -3'	N/A
h-GPR107-RT-R	5'- CAAGGACTTTCTCATCCCTGCTG -3'	N/A
h/m-BCL9-RT-F	5'- AGCCCTAAGTCAAAGCAGGAGG -3'	N/A
h/m-BCL9-RT-R	5'- CATTTCCAGCCCCATTCTTCAG -3'	N/A
h/m-ARHGAP1RT-F	5'- TCCTCTTCAAGCCCCTCATCAG -3'	N/A
h/m-ARHGAP1RT-R	5'- CGAGACTCCAAACTGCTGGTTG -3'	N/A
h/m-CEPT1-RT-F	5'- GTCAACAAGGAAACGATGTGGAG -3'	N/A

h/m-CEPT1-RT-R	5'- GATGGTGATGAGATTTGGGGC -3'	N/A
h/m-SP2-RT-F	5'- TGGAGGGCAGTTTGTGTTTGC -3'	N/A
h/m-SP2-RT-R	5'- GCTTGATGGGGACAGGCTTATG -3'	N/A
m-Actb-RT-F	5'- ACAACGGCTCCGGCATGT -3'	N/A
m-Actb-RT-R	5'- TCTTGCTCTGGGCCTCGTCAC -3'	N/A
m-Stx6-RT-F	5'- TGGAGGACCCCTTCTTTGTAGTG -3'	N/A
m-Stx6-RT-R	5'- TCTATGCTGCGGAGATTGTTCC -3'	N/A
m-Slc52a2-RT-F	5'- GAATCAAGTCAGCCAGCCACAC -3'	N/A
m-Slc52a2-RT-R	5'- AGATGAGCGAAAGAGCCAGTCC -3'	N/A
m-Cda-RT-F	5' TGACTACGAATGTGCCATCCG 3'	N/A
m-Cda-RT-R	5' CATAGAAAACGCCTGCTACCCAC 3'	N/A
m-Gnpda2-RT-F	5'- TTCGGTCGTTGGATGAGGCTGTAG -3'	N/A
m-Gnpda2-RT-R	5'- CAGGGTGGTTTCTTGGAAGTCC -3'	N/A
m-Zbtb4-RT-F	5'- CATCTTCTGCTGGGACACCTTTG -3'	N/A
m-Zbtb4-RT-R	5'- TGGGTTTGTAGCCTCCATTGGG -3'	N/A
m-Gpr107-RT-F	5'- CATCTCTGGAAGCATAGTCAAGGTC -3'	N/A
m-Gpr107-RT-R	5'- AGCCTCACTGTCTTTGGGGTTG -3'	N/A
BCL9 cloning primers		
Primer Name	**Sequence**	
BCL9-FLAG-F	5' ATAAGAATGCGGCCGCTAAACATTCCAGTAACCCTAAAG TGA 3'	
BCL9-FLAG-R	5' GCCCTCTAGATGCATGCTCGA 3'	
seq_FLAG_BCL9_F	5' CGGAATTGTACCCGCGGG 3'	
seq_FLAG_BCL9_R	5' CCCACTCTTGGAGTCACAG 3'	
BCL9_S1.2_Mut_F	5' CATACCCCTAAAGCCTTACCTGGCCCAGGTGGGAGCAT G 3'	
BCL9_S1.2_Mut_R	5' GCCCCCACTCTTACTATCGCAGGGGGATGGCTGGGATT G 3'	
BCL9_S4_Mut_F	5' ACGCCCGAGCAGATAGCGTGGCTG 3'	
BCL9_S4_Mut_R	5' CATGTGGTCAAGGTGGTCGGGTCC 3'	
BCL9_S5_Mut_F	5' AGCCCTCCAGCTCGTTCTCCCAAC 3'	
BCL9_S5_Mut_R	5' ATCGTCATCTGAGCTGGCCACAGT 3'	
BCL9_S6_Mut_F	5' TAATCAGATGCCCTCTCCAAATG 3'	
BCL9_S6_Mut_R	5' CTATGAGAAAGTGGCTGGGTCAT 3'	
BCL9_S7_Mut_F	5' TAGCTTCACTGTCCTGGGGAACAGCATGCC 3'	
BCL9_S7_Mut_R	5' TCCGGACCCCCGGGGCCTCCAGG 3'	