

SUPPLEMENTAL FIGURES

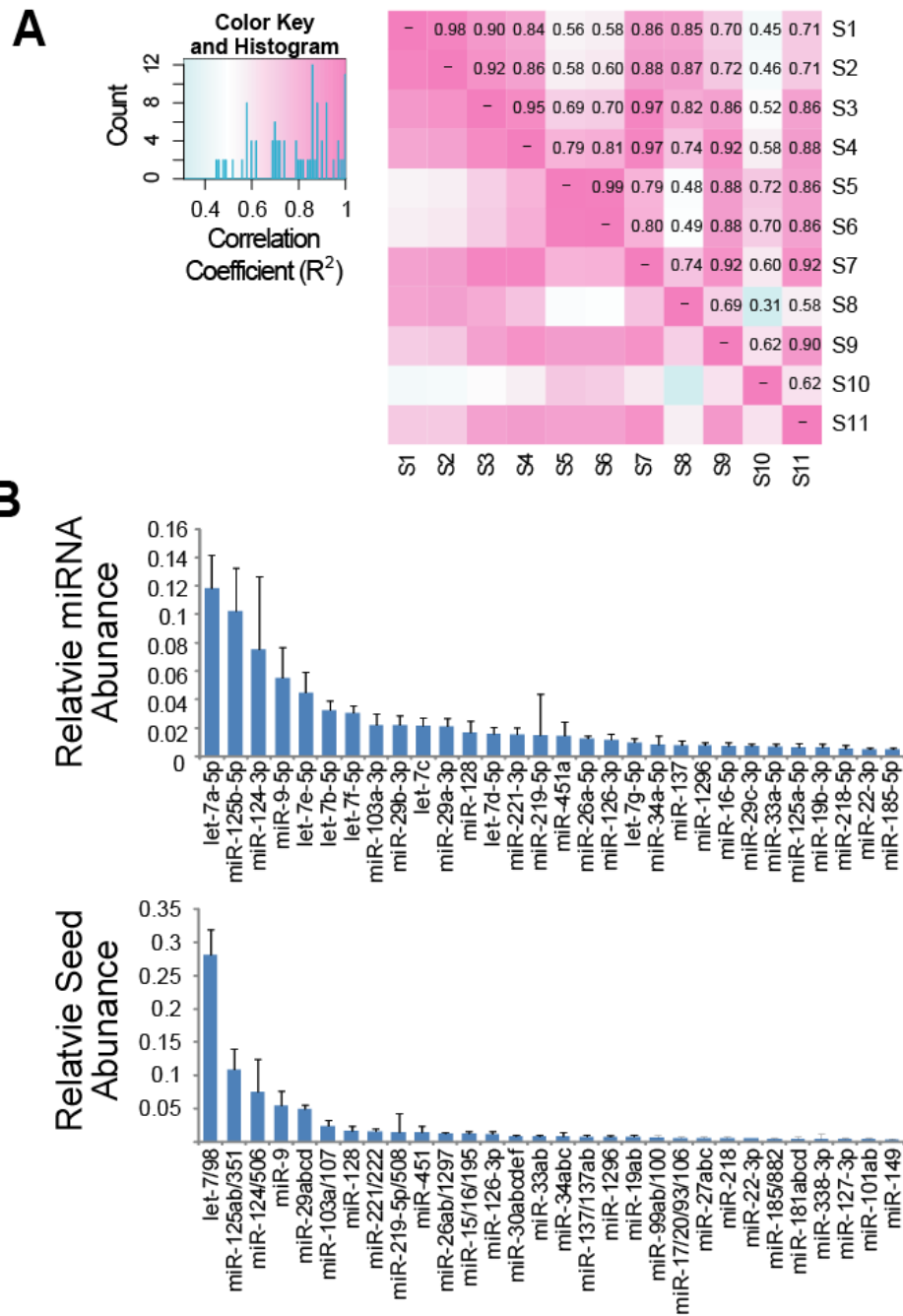


Figure S1. Reproducibility and levels of human brain miRNAs across samples. (A) Pairwise correlations of normalized, relative miRNA levels across each of the 11 samples (S1-S11) were performed and a heatmap plotting the correlation coefficient for each comparison is shown. Median $R^2=0.79$. (B) Bar graphs are provided to show the relative mean levels among the 11 samples for the top 30 most abundant miRNAs (top) and miRNA seeds (bottom). For the latter, miRNAs were grouped into seed families based on shared seed (7m8) sequences. Error bars show standard deviation.

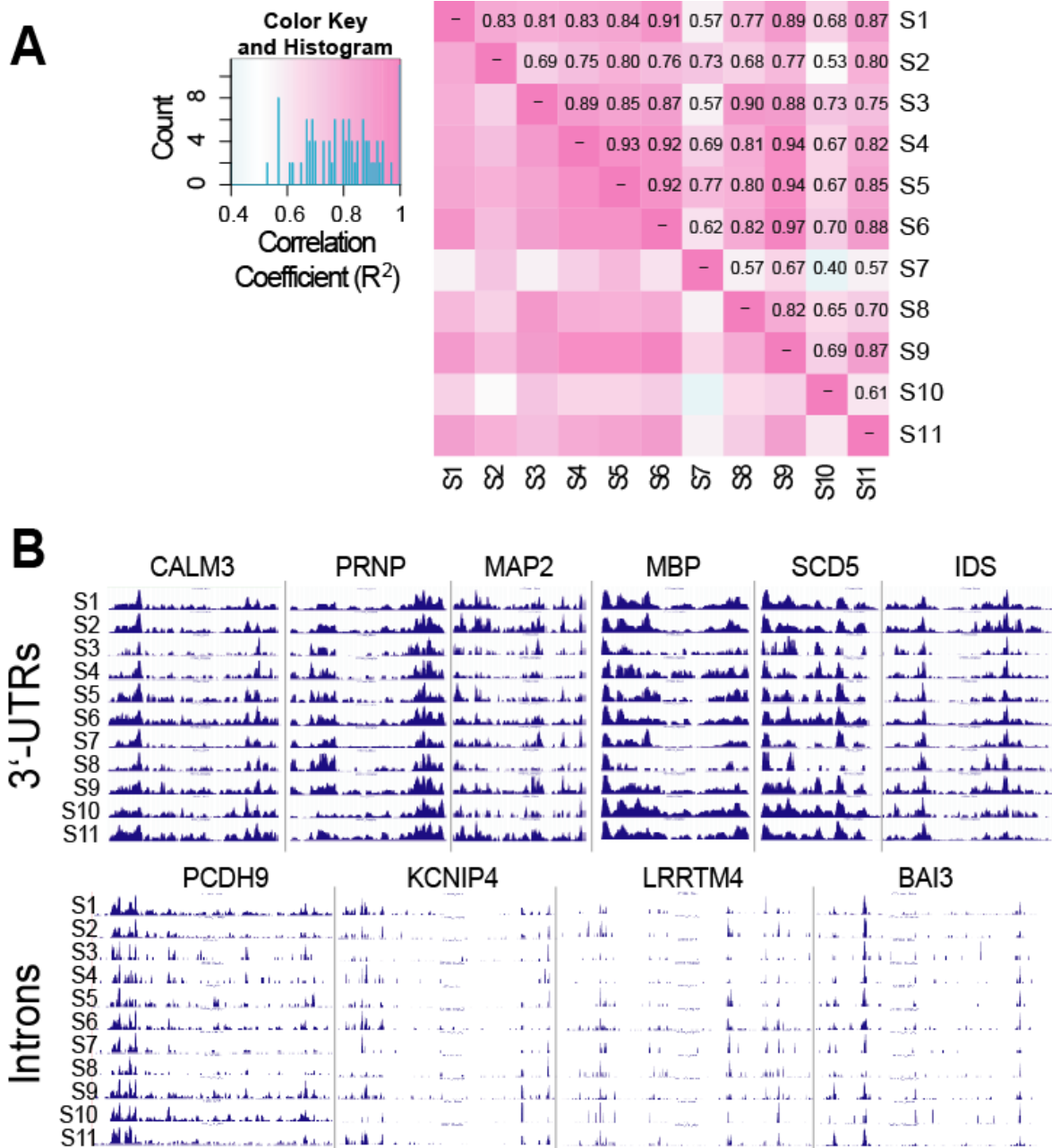


Figure S2. Reproducibility of Ago2 binding profiles across samples. (A) Human brain Ago2 HITS-CLIP RNA tag reads were mapped to the genome and corresponding wiggle files (i.e. the read coverage at each genomic position) were generated. Correlation of wiggle files was performed and the resulting distribution of R-squared values for all pair-wise comparisons is shown by histogram and heatmap. Median $R^2=0.80$. (B) UCSC Genome Browser images of positional read coverage (i.e. wiggle files) for each of the 11 samples (S1-S11) are provided to illustrate data reproducibility across samples. Images span 3'-UTR (top) or intronic (bottom) regions within the indicated genes. In total, the 3'-UTRs depicted span >10kb while the intronic regions span >40kb. Overall, the data exhibits evident visual reproducibility. Interestingly, the intronic regions shown are parts of introns which had a notably high density of Ago2 clusters as compared to most intronic regions.

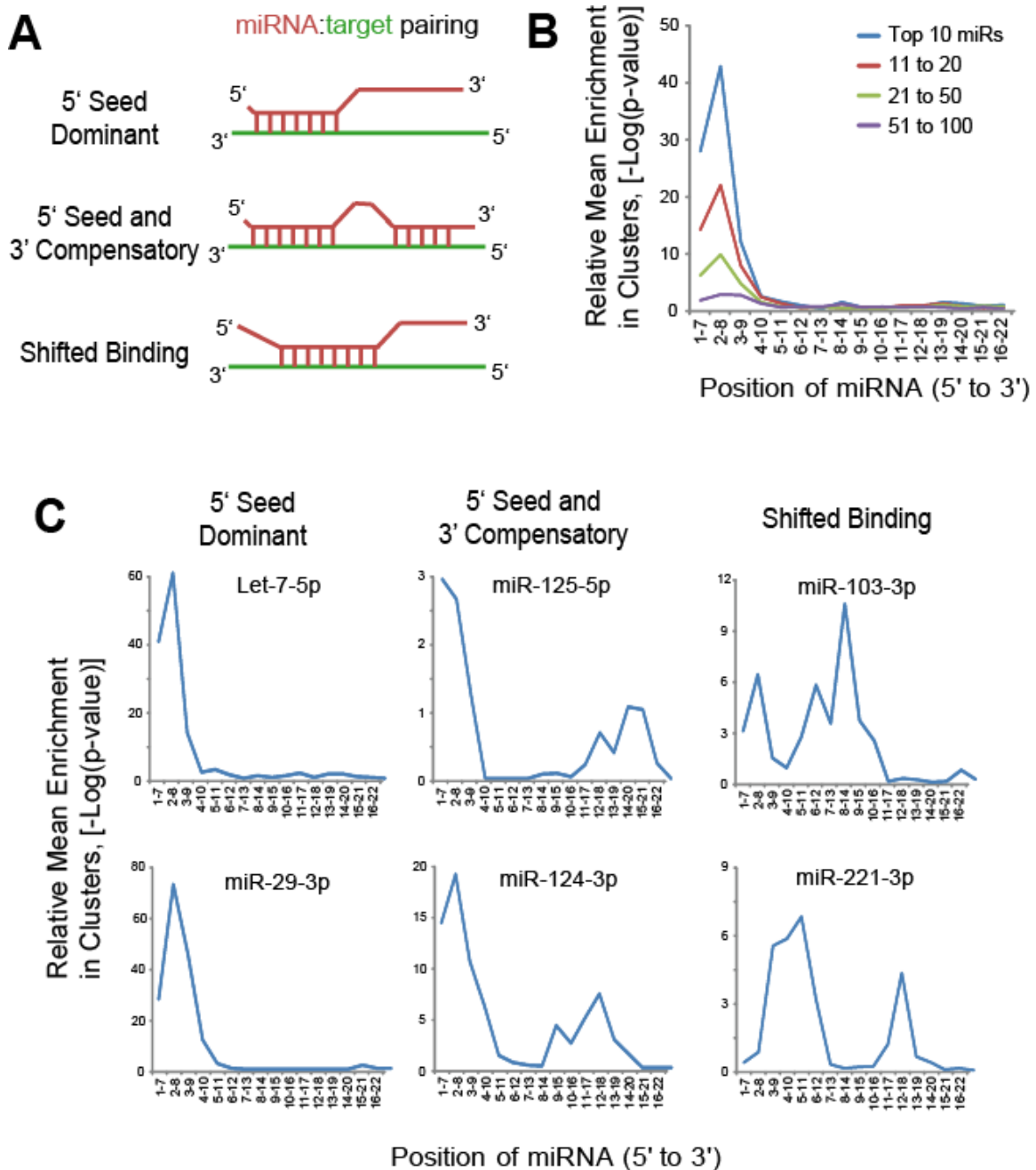


Figure S3. Positional analysis of miRNA sequence motif enrichments. (A) Schematic of various models for miRNA:target recognition. To determine which positions within miRNAs account for the most highly enriched heptamers present within the human Ago2 brain clusters, we evaluated the significance of enrichment for each possible heptamer along the length of the miRNA. (B) The relative enrichment of heptamers corresponding to each position along the miRNA was determined for the top 100 most highly expressed miRNAs. miRNAs were parsed into the indicated groups by expression level (e.g. top 10 miRNAs are the most highly expressed), and the mean positional motif enrichment within groups was calculated and plotted. (C) Positional motif enrichment profiles are shown for miRNAs which serve as examples for the various miRNA binding models depicted in panel A.

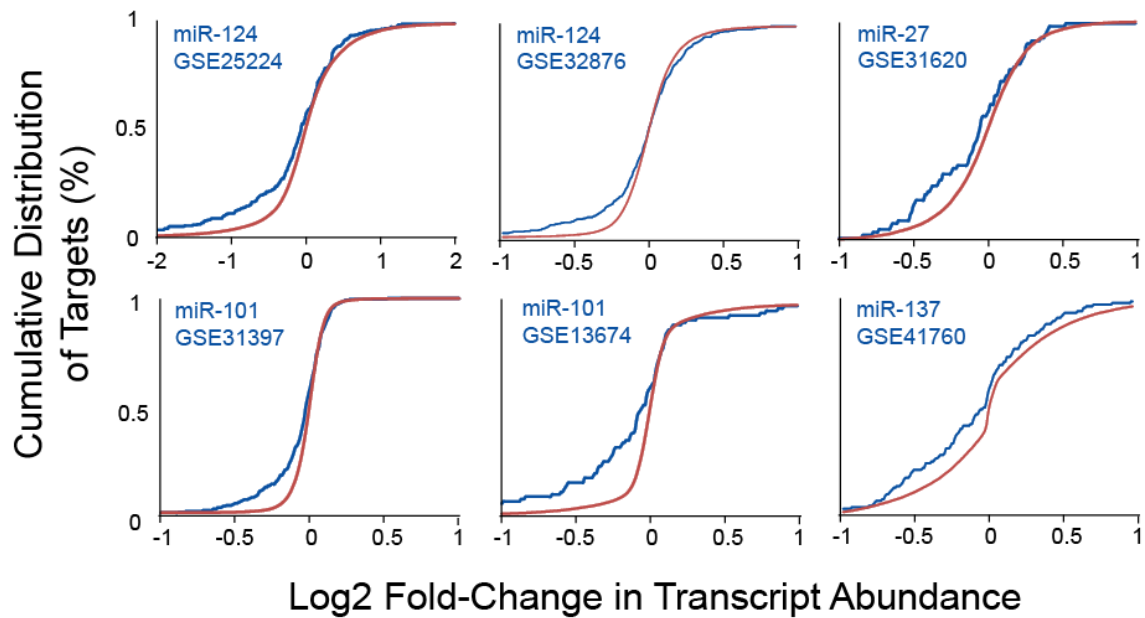


Figure S4. Functionality of miRNA:target gene pairs identified by Ago2 HITS-CLIP. Functionality of miRNA:target gene pairs identified by Ago2 HITS-CLIP was evaluated using microarray data for gene expression changes following miRNA overexpression for the indicated miRNAs (GEO Accession indicated for each plot). Cumulative fraction curves for miRNA:target pairs (e.g. genes with 3'-UTR Ago2 HITS-CLIP clusters harboring seed sites for the relevant miRNAs) are plotted. Curve displacement to the left or right, relative to background genes (i.e. no site), indicates down- or up-regulation, respectively. As expected, miRNA overexpression results in down-regulation of many targets identified among our Ago2 HITS-CLIP data.

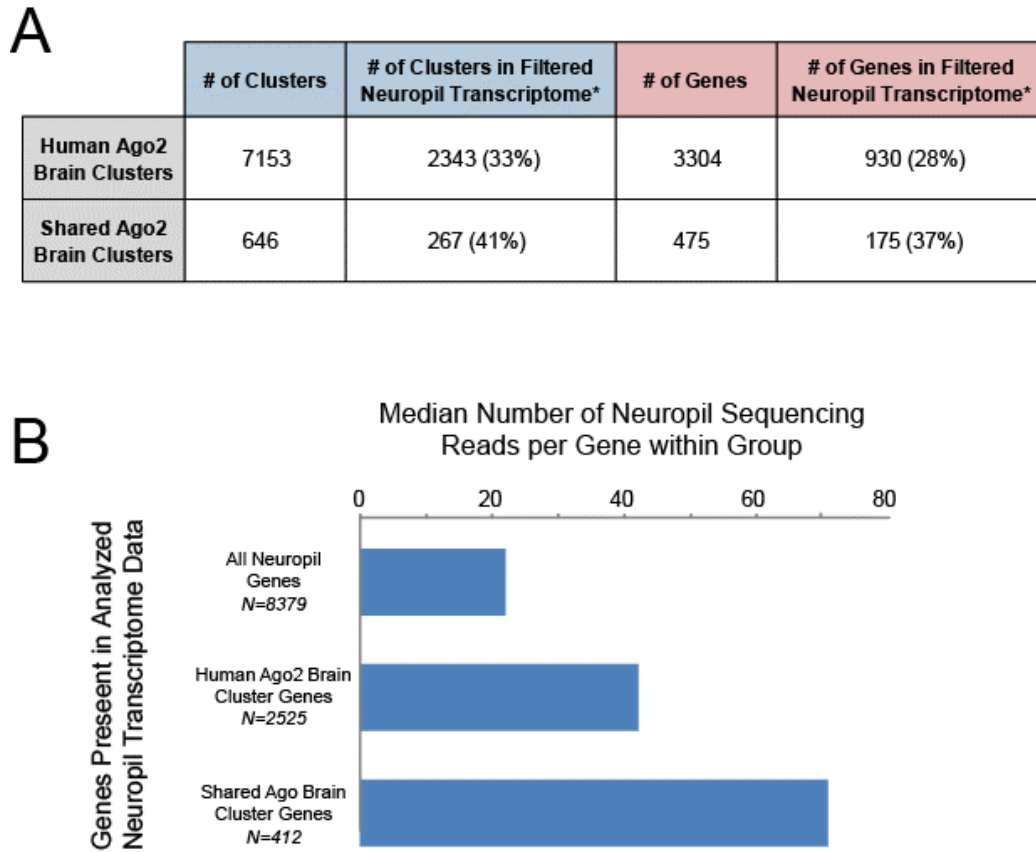


Figure S5. Enrichment of Ago brain clusters in the local neuropil transcriptome. (A) Table summarizing the overlap of Ago brain clusters and associated genes with the previously published local transcriptome of the synaptic neuropil. “Shared” refers to Ago clusters present in both human and mouse. *The filtered neuropil transcriptome consists of 2550 genes whose transcripts localize to the synapse (see Cajigas et al. 2012). Shared Ago brain clusters show >25% enrichment of clusters and genes overlapping with filtered neuropil genes ($p=0.000012$), Chi-squared test with Yate’s correction). (B) Genes containing Ago brain clusters also show enrichment of neuropil sequencing coverage in the Cajigas et al. data.

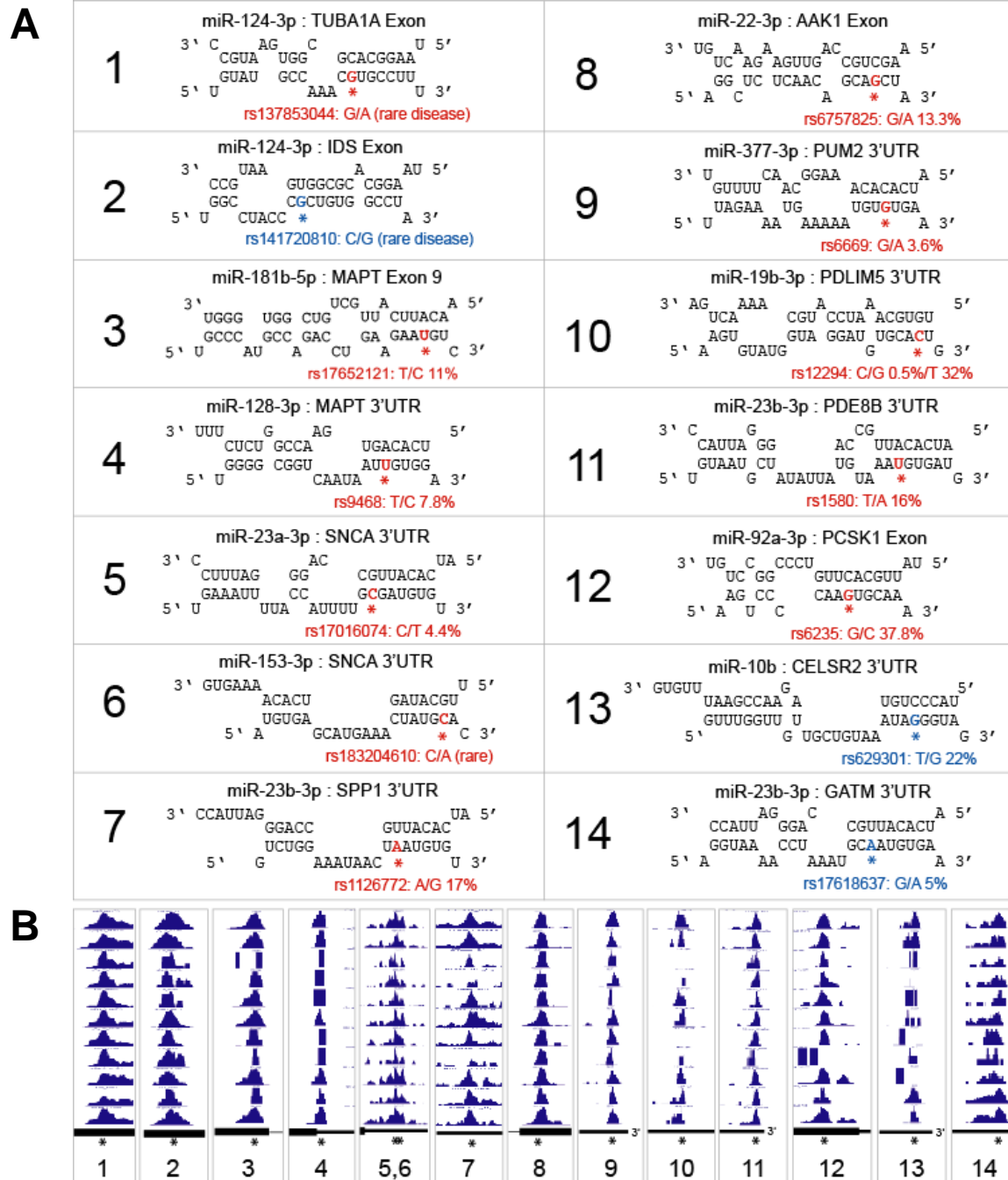


Figure S6. Disease-relevant SNPs alter miRNA targets in Ago2 HITS-CLIP sites. (A) Diagrams of the miRNA:target interactions (as predicted by RNAhybrid web-server, see Experimental Procedures) are shown. The position of the SNP within the seed region and the SNP ID along with alleles (major/minor) and minor allele frequencies are highlighted in red (site disruption) or blue (site enhancement/creation). **(B)** Data reproducibility in regions with disease-relevant SNPs. UCSC Genome Browser images of positional read coverage information (i.e. wiggle files) for each of the 11 samples are provided to illustrate the presence of reproducible peaks overlaying disease-relevant SNPs (denoted by asterisks). Thick and thin black boxes represent exons and UTRs, respectively, while lines indicate introns. Ends of 3'-UTRs are signified with 3'. Numbers are provided to associate the sites between A and B panels.

AUTHOR CONTRIBUTIONS

R.L.B. and B.L.D. conceived the project, supervised the experiments, analyzed the data and wrote the paper. R.L.B., P.J., Y.X. and B.L.D. designed the research. B.L.G and R.L.B. performed the biochemical aspects of HITS-CLIP and luciferase validation studies. P.J. led the bioinformatic and statistical analyses, developed a novel “peak calling” strategy and wrote bioinformatic methods. R.M.S and R.L.B performed additional bioinformatic and statistical analyses. Y.X. supervised the bioinformatics and statistical analyses. J.A.N. and R.T. generated the anti-Ago2-3148 antibody. C.A.R. coordinated the experimental approach to the brain samples.