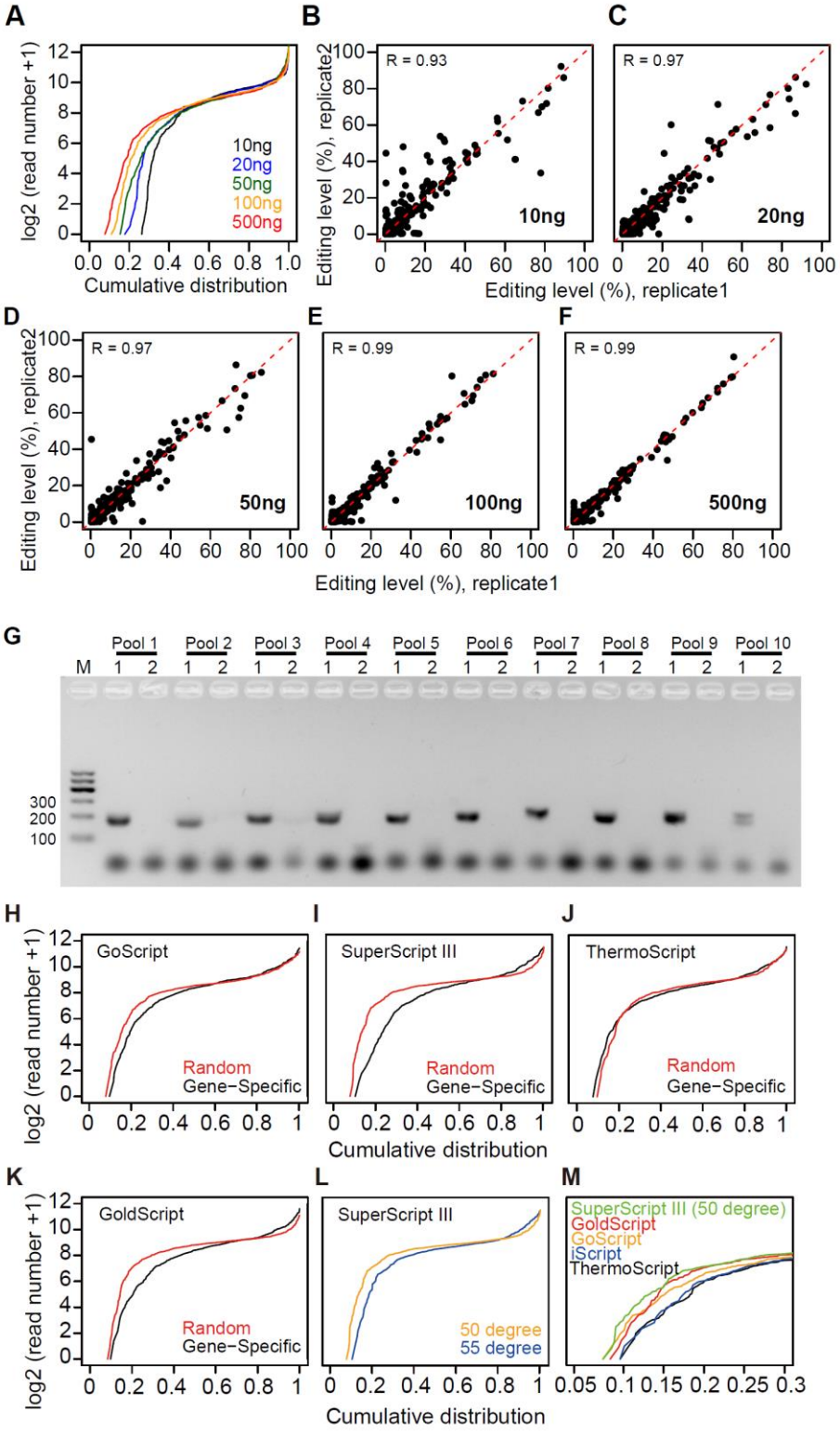


Supplemental Figure S1. The development and performance of miR-mmPCR-seq.

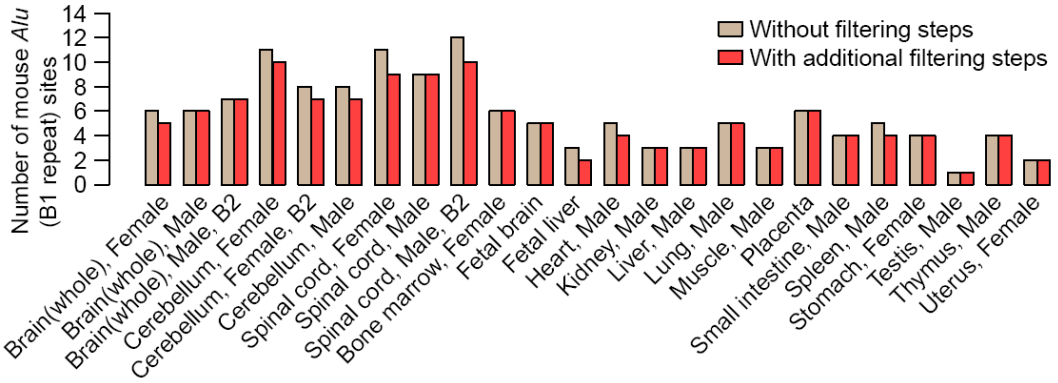
(A) The cumulative distribution of amplicon coverage using different amounts of preamplified cDNA input. The y axis shows the number of reads in log2 scale. Read numbers are normalized to 0.25 million mapped reads per sample.

(B-F) Relationship between the reproducibility of RNA editing measurements and the amount of preamplified cDNA input. The Pearson correlation coefficient (R) is indicated.

(G) Ten pools of mouse multiplex primers were randomly selected. Mouse brain RNA was treated with one or two rounds of DNase I. 1,1 round of treatment; 2, 2 rounds of treatment; M, marker.

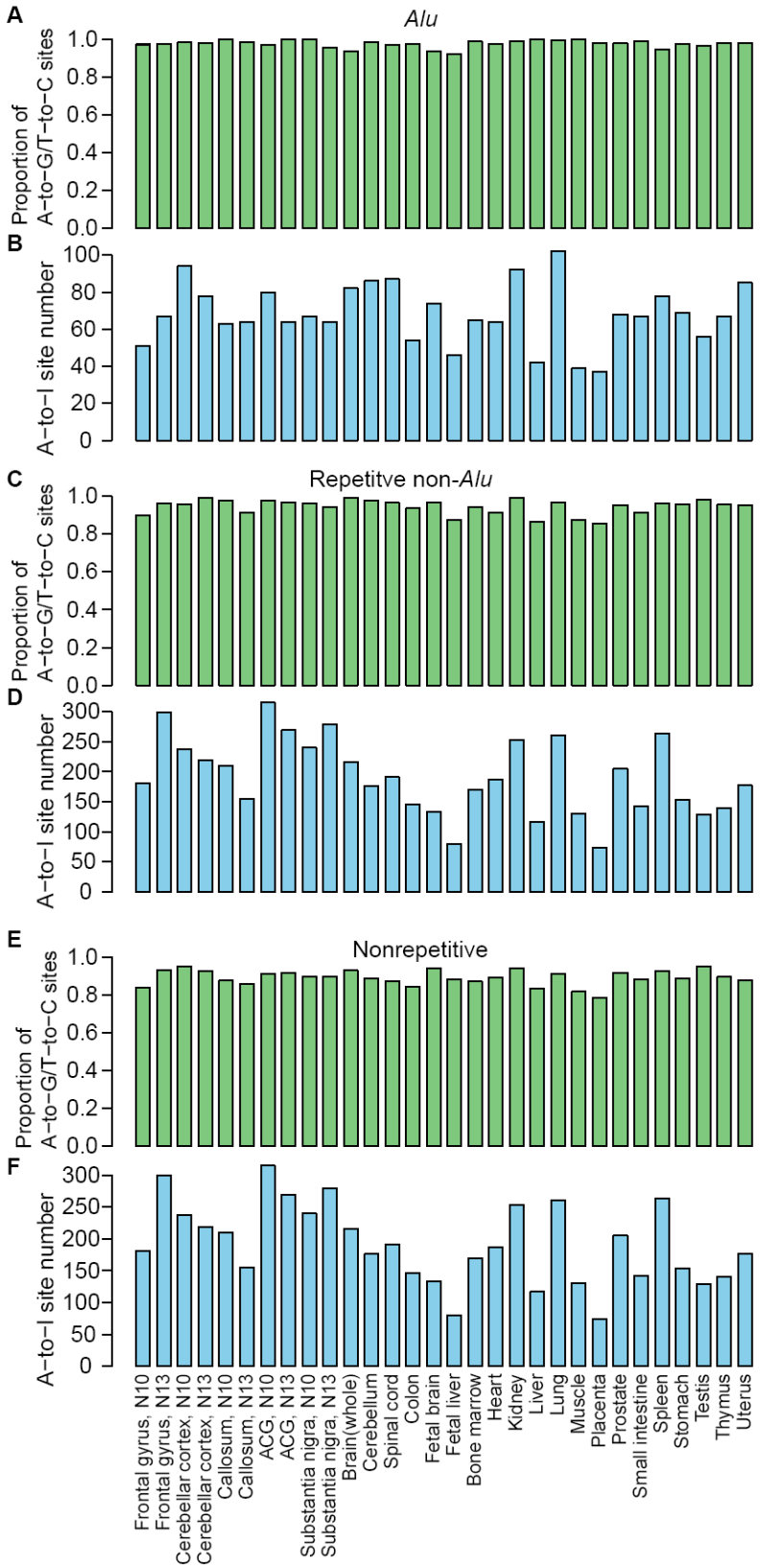
(H-L) The cumulative distributions of amplicon coverage using different reverse transcriptases. Four different reverse transcriptases in which transcriptases are not preblended with primers were tested using either gene-specific or random hexamer primers (**H-K**). SuperScript III was tested with 50°C and 55°C reaction temperatures (**L**).

(M) Comparison of amplicon coverage using five different reverse transcriptases.



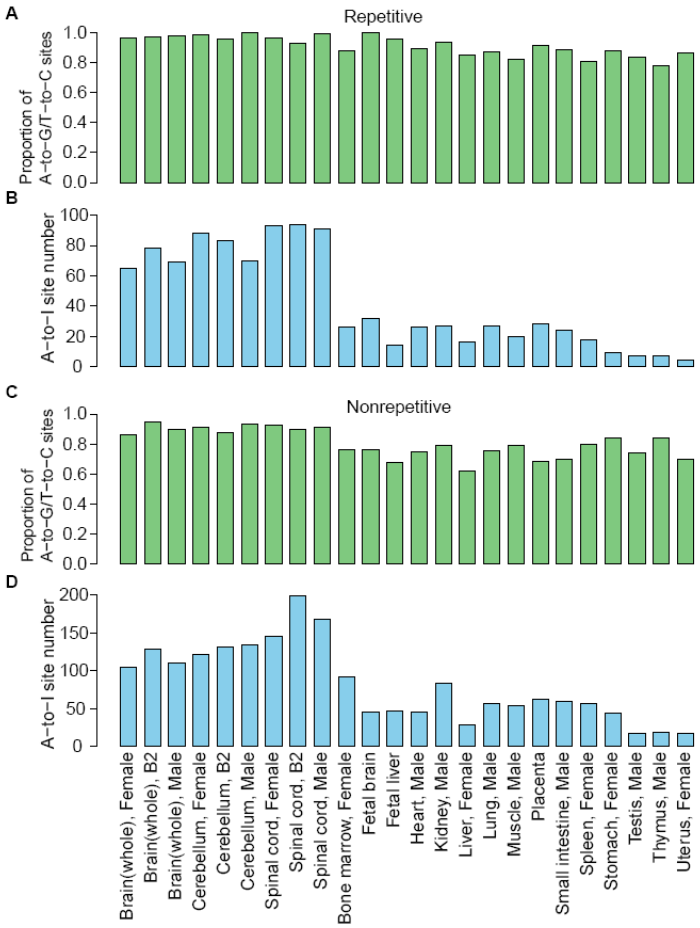
Supplemental Figure S2. **The number of the B1-derived A-to-I sites in different mouse tissues, without or with additional filtering steps.**

Sites with editing level ≥5% were used for analysis.



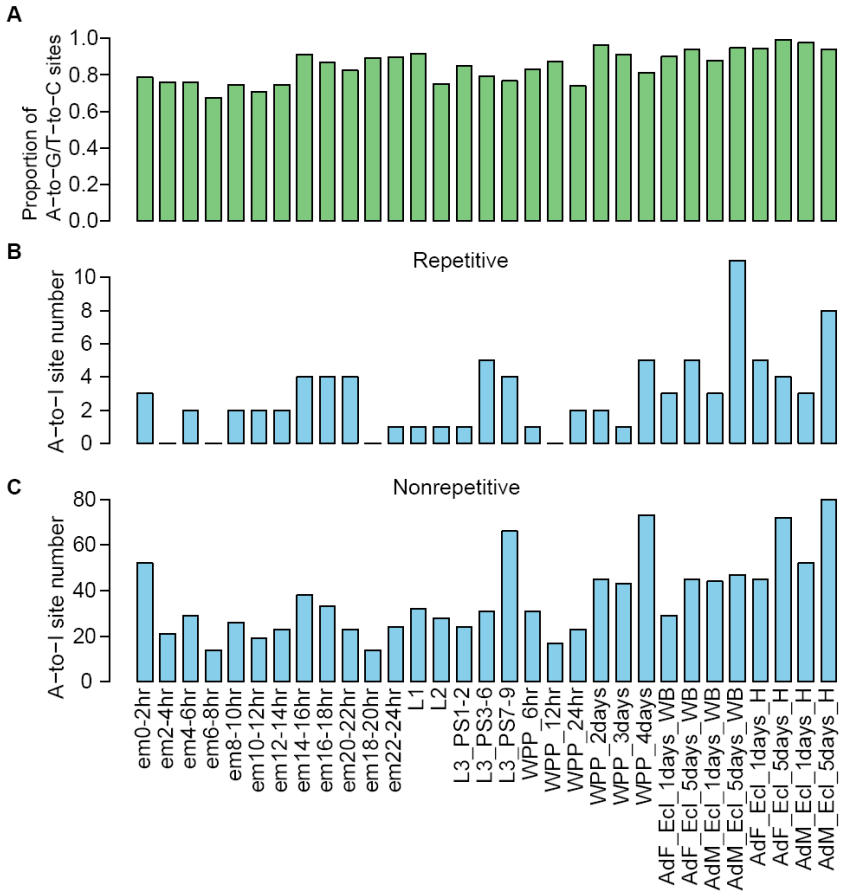
Supplemental Figure S3. Identification of RNA editing sites in human.

The proportion of detected sites that are A-to-G/T-to-C and the A-to-I site number in different human tissues for *Alu*-derived (**A-B**), repetitive non-*Alu*-derived (**C-D**) and nonrepetitive-derived (**E-F**) sites. Sites with editing level ≥5% were used for analysis. ACG, anterior cingulate gyrus.



Supplemental Figure S4. Identification of RNA editing sites in mouse.

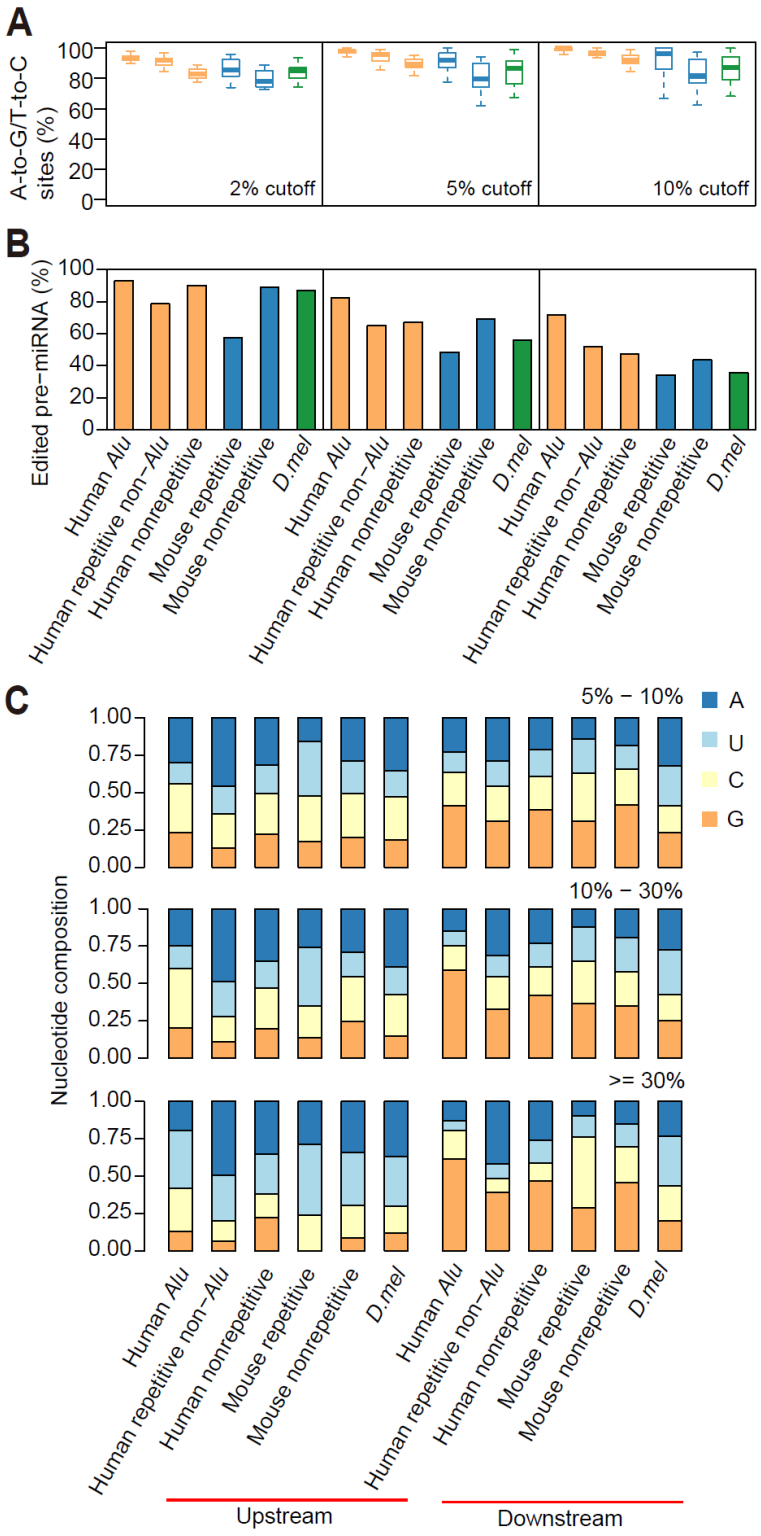
The proportion of detected sites that are A-to-G/T-to-C and the A-to-I site number in different mouse tissues for repetitive-derived (**A-B**) and nonrepetitive-derived (**C-D**) sites. Sites with editing level ≥5% were used for analysis.



Supplemental Figure S5. Identification of RNA editing sites in *D.mel*.

(**A**) The proportion of detected sites that are A-to-G/T-to-C in *D.mel*. Due to the limited number of repetitive RNA variants, repetitive-derived and nonrepetitive-derived sites were combined together for calculation.

(**B-C**) The A-to-I site number for repetitive-derived (**B**) and nonrepetitive-derived (**C**) sites. Sites with editing level ≥5% were used for analysis.



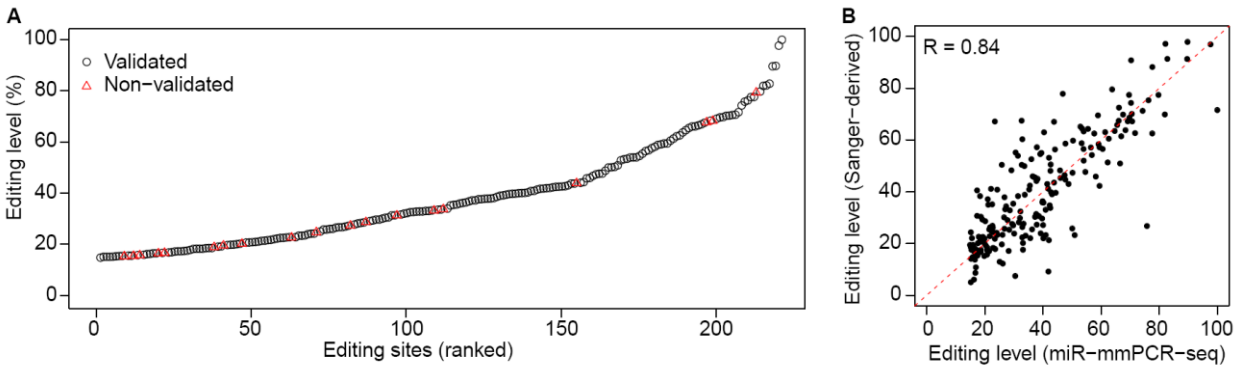
Supplemental Figure S6. pri-miRNA editing identification.

(A) Relationship between the fraction of detected sites that are A-to-G/T-to-C and the minimum

editing level. The y axis shows the proportion of A-to-G/T-to-C sites (the number of A-to-G/T-to-C variants divided by the number of all detected variants). The editing level of each site is defined as the fraction of reads with the altered nucleotide.

(B) Relationship between the proportion of edited pre-miRNAs and the minimum editing level.

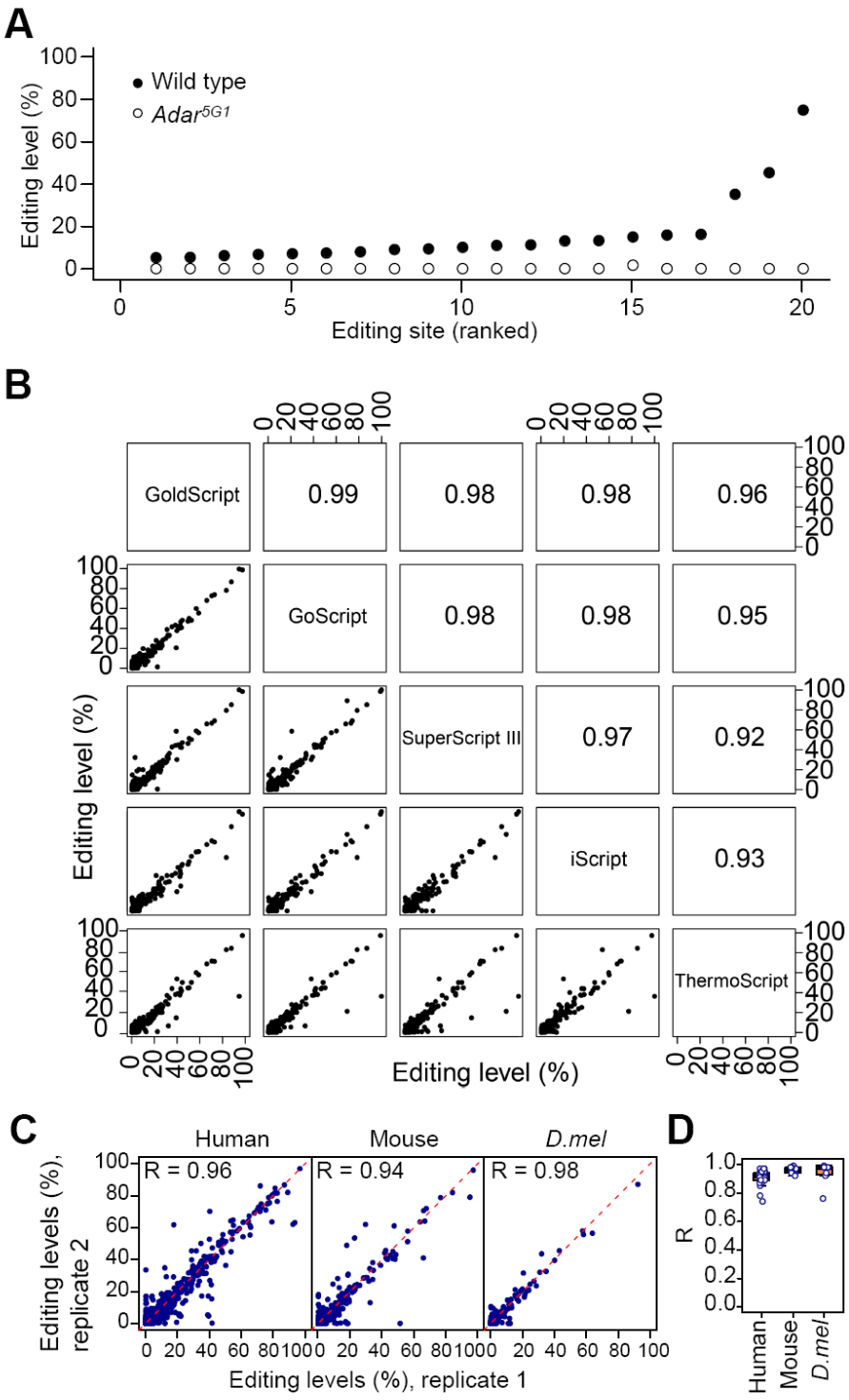
(C) Nucleotide composition in the position immediately upstream and downstream of editing sites. The fractions of A, U, C, and G were shown for sites edited at different levels (5-10%, 10-30%, and ≥30%).

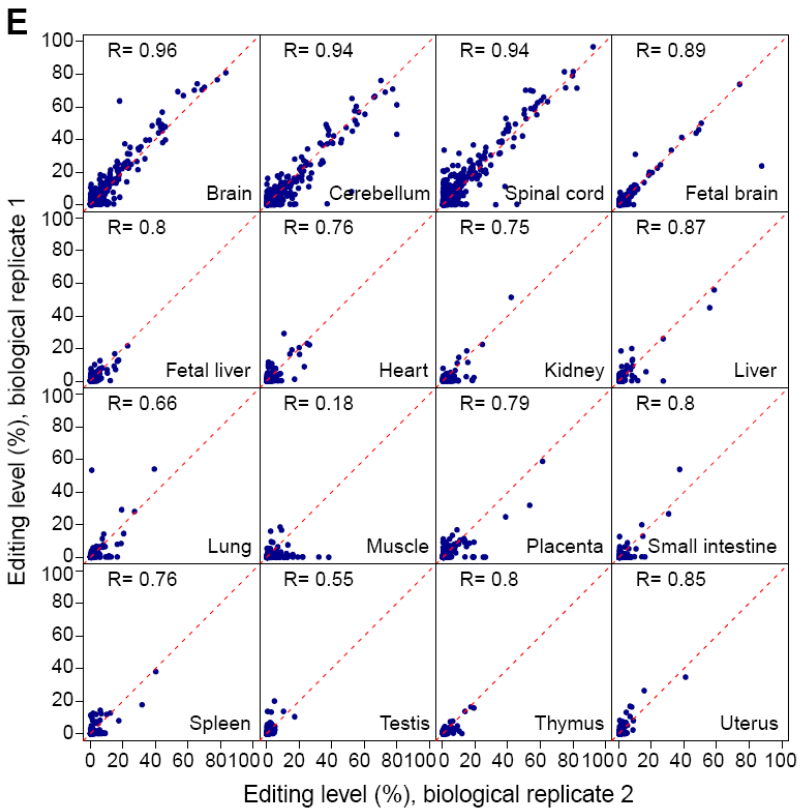


Supplemental Figure S7. Sanger validation of A-to-I RNA editing sites.

(A) The summary of Sanger validation of A-to-I RNA editing sites. Circle, validated; triangle, non-validated. Sites were sorted by editing levels measured from miR-mmPCR-seq.

(**B**) The comparison between Sanger-derived editing levels and the editing levels measured from miR-mmPCR-seq. Sanger-derived editing levels were inferred via QSVanalyser (Carr et al. 2009). In brief, the chromatograms were quantified by measuring T and C peak heights in strands opposing the edited strand because A/G mixed peaks have more inconsistent heights (Eggington et al. 2011).


A
Wild type
Adar5G1
Editing level (%)
Editing site (ranked)
B
GoldScript
GoScript
SuperScript III
iScript
ThermoScript
0.99
0.98
0.98
0.96
0.98
0.98
0.95
0.97
0.92
0.93
Editing level (%)
C
Human
Mouse
D.mel
R = 0.96
R = 0.94
R = 0.98
Editing levels (%), replicate 1
Editing levels (%), replicate 2
D
R



Supplemental Figure S8. The accuracy and reproducibility of editing site identification and quantification.

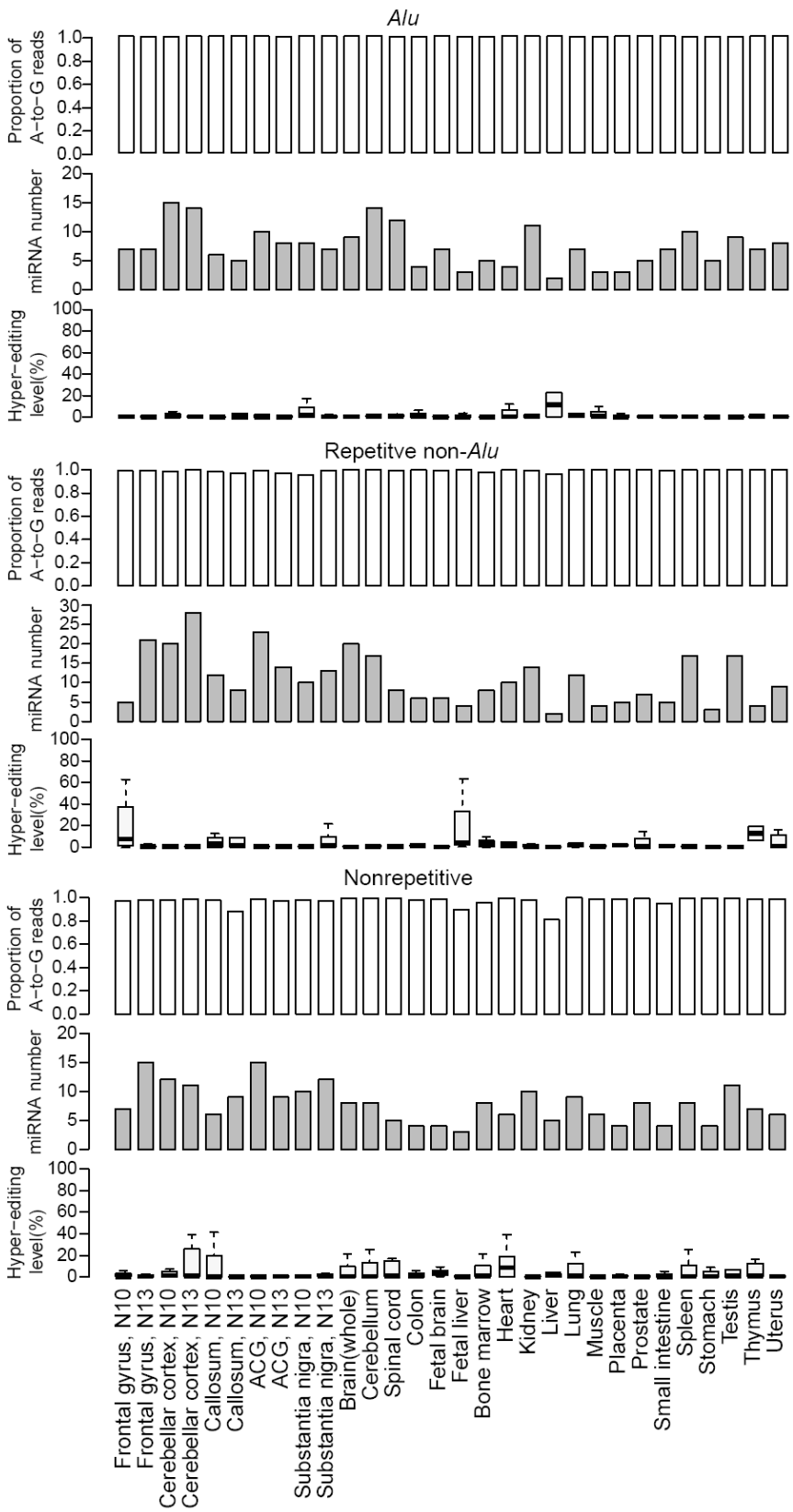
(A) RNA editing levels of editing sites measured from the heads of wild-type and *Adar*^{5G1} mutant flies.

(B) Reproducibility of RNA editing level measurements. Examples of editing level comparison between the mouse brain RNA samples reverse-transcribed using different reverse transcriptases with random primers. The Pearson correlation coefficient (R) is indicated.

(C) Examples of editing level comparison between technical replicates in human (brain sample), mouse (brain sample) and *D.mel* (female adult head sample).

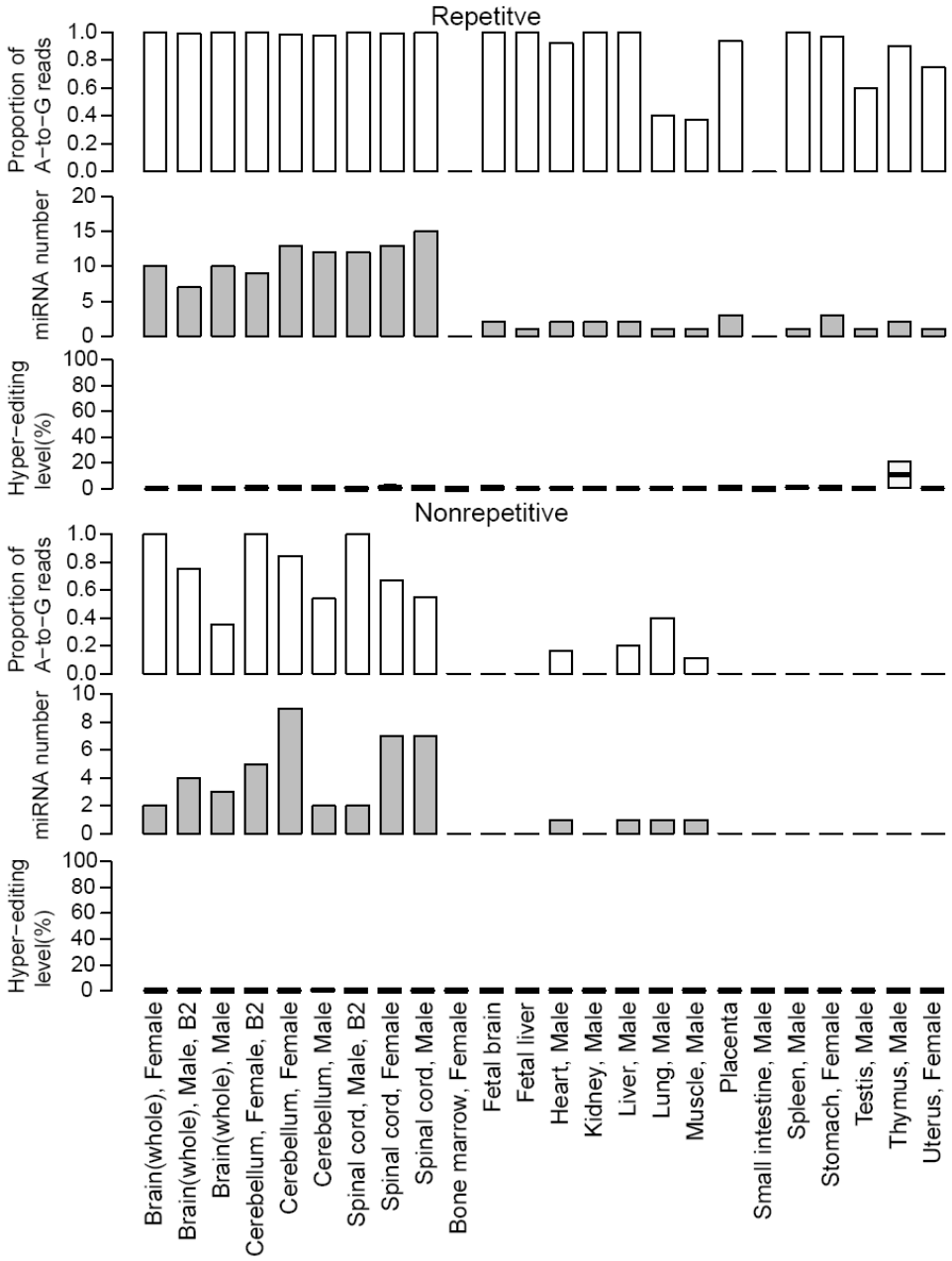
(D) Reproducibility of RNA editing level measurements, quantified as R between technical replicates for samples. Only samples with more than 40 sites (human or mouse) or 30 sites (*D.mel*) that are edited (≥5% editing levels) were used for analysis.

(E) Editing level comparison between biological replicates in different mouse tissues.



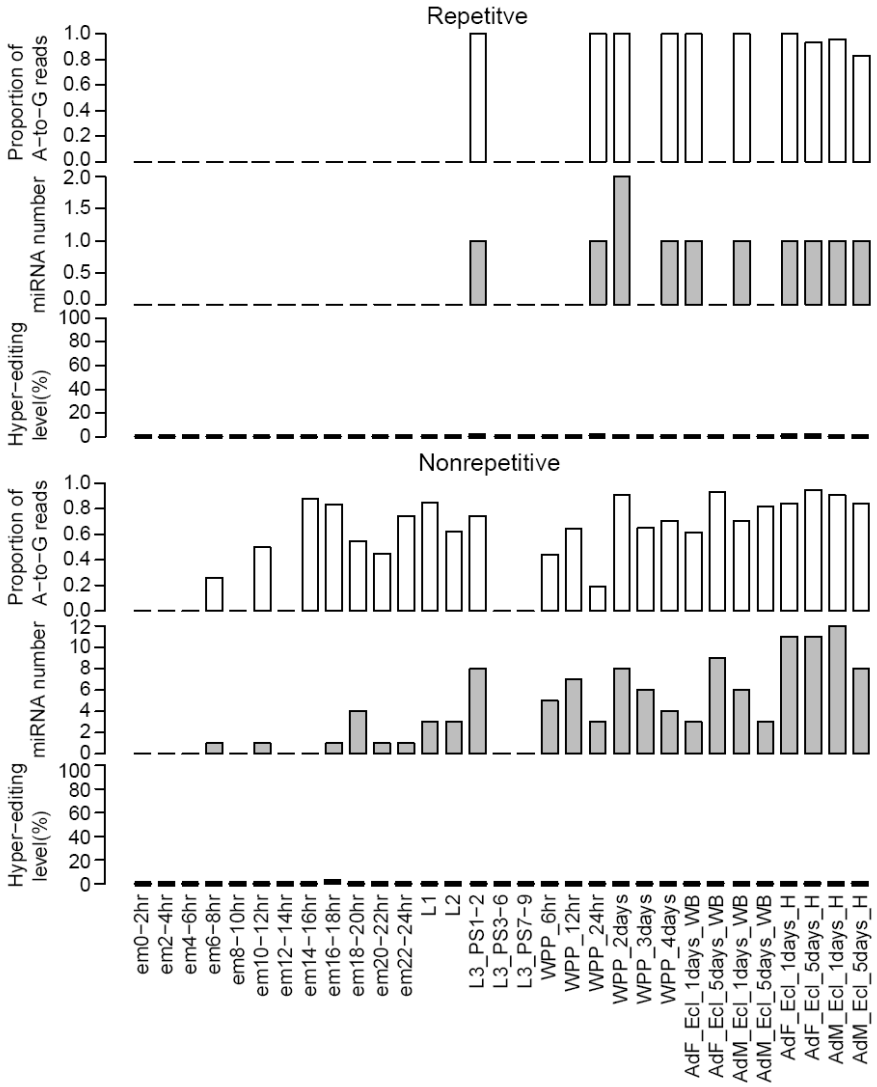
Supplemental Figure S9. Identification of hyper-editing sites in human.

The proportion of hyper-edited reads that are A-to-G type, the number of miRNAs with A-to-I hyper-editing events and the boxplot of editing levels of hyper-edited miRNAs in different human tissues for *Alu*-derived, repetitive non-*Alu*-derived and nonrepetitive-derived miRNAs were shown, separately. ACG, anterior cingulate gyrus. The editing level of a miRNA is defined as the number of A-to-G hyper-edited reads divided by the total number of reads mapped to this miRNA.



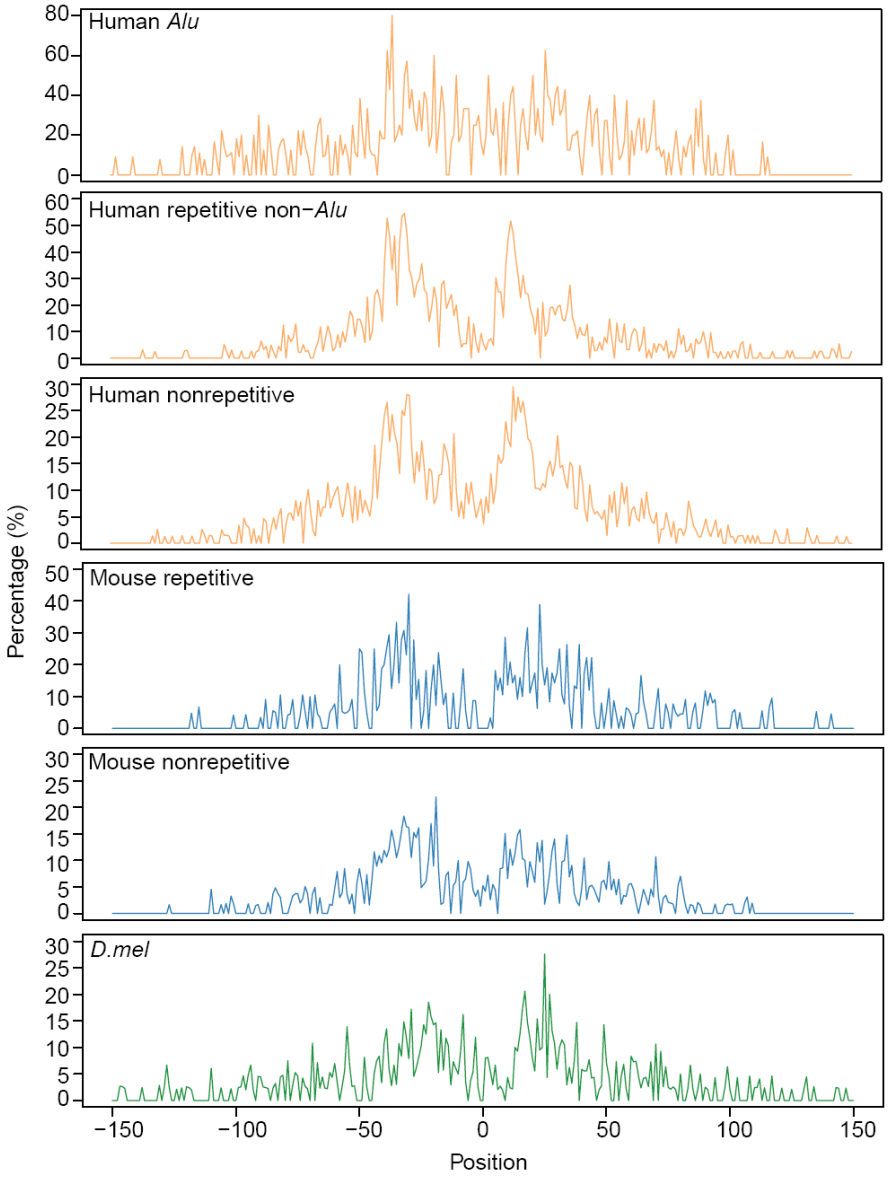
Supplemental Figure S10. Identification of hyper-editing sites in mouse.

The proportion of hyper-edited reads that are A-to-G type, the number of miRNAs with A-to-I hyper-editing events and the boxplot of editing levels of hyper-edited miRNAs in different mouse tissues for repetitive-derived and nonrepetitive-derived miRNAs were shown, separately.



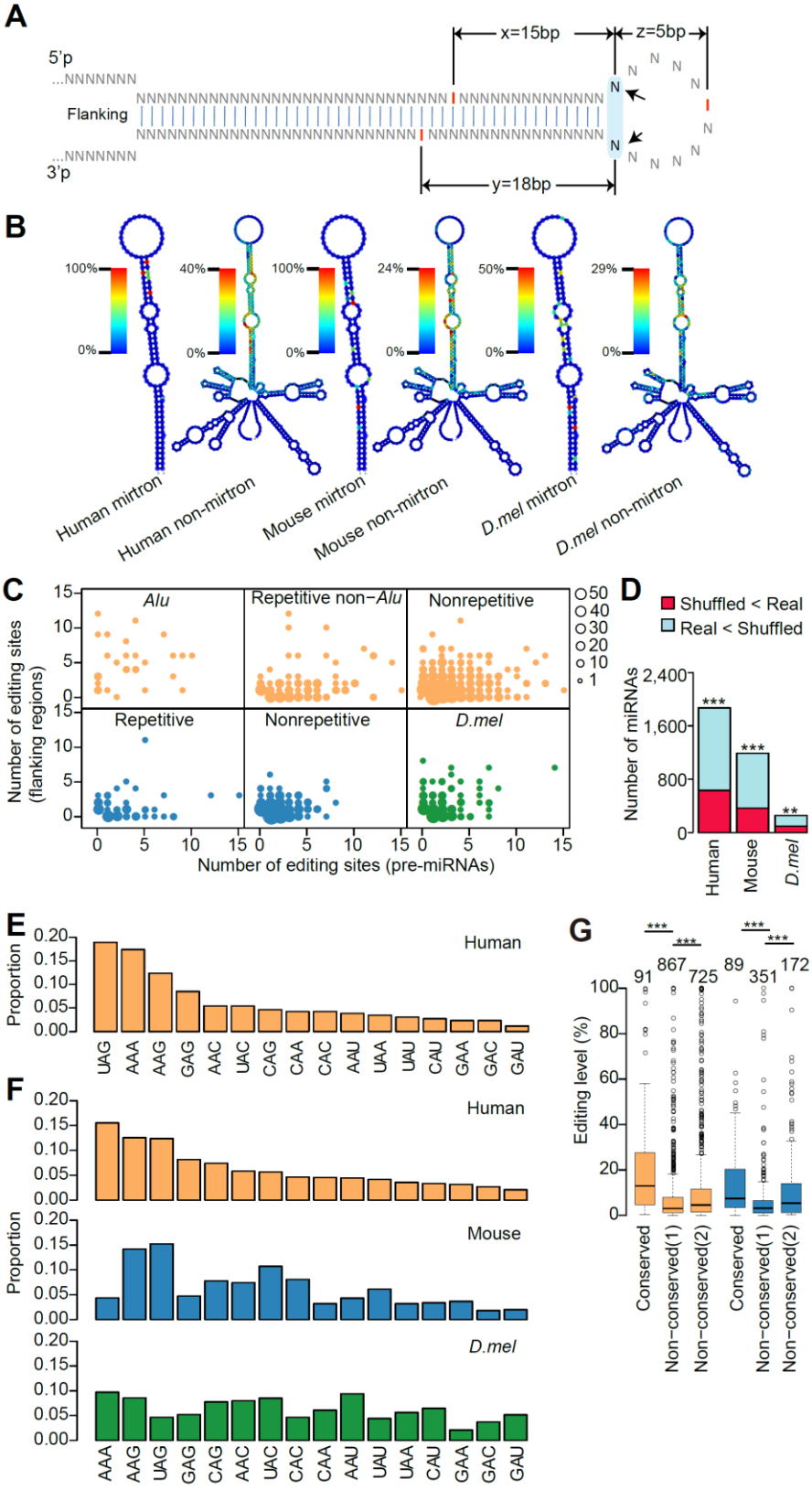
Supplemental Figure S11. Identification of hyper-editing sites in *D.mel*.

The proportion of hyper-edited reads that are A-to-G type, the number of miRNAs with A-to-I hyper-editing events and the boxplot of editing levels of hyper-edited miRNAs in *D.mel* for repetitive-derived and nonrepetitive-derived miRNAs were shown, separately.



Supplemental Figure S12. The distribution of RNA editing sites across pri-miRNAs.

Metagene profiles depicting the editing site distribution across pri-miRNAs. Position 0 of a miRNA is defined as the central position of the loop region of a pre-miRNA. The negative value, the distance between an upstream position and the central position; the positive value, the distance between a downstream position and the central position. Percentage of As that were edited at each position was calculated.



Supplemental Figure S13. Characterization and regulation of miRNA editing sites.

(A) Illustration of the mapping of a given editing site to the secondary structure of a miRNA. For a 5'p editing site, we calculated the distance between this site and the 5'start of the central loop region (x); for a 3'p site, we calculated the distance between this site and the 3'end of the central loop region (y); for a site within the loop, we calculated the distance between this site and the 5' start of the loop region (z). Finally, the information of all edited sites was projected to the secondary structure of a selected miRNA and the flanking sequence. Percentage of As that were edited at each position was calculated.

(B) Comparison of editing site distribution between mirtron and non-mirtron miRNAs. The structure predicted using *dme-mir-2492* precursor and the flanking 150 nt sequence is used as the representative non-mirtron secondary structure. The structure predicted using mirtron *dme-mir-2494* is used as the representative mirtron secondary structure.

(C) Editing density comparison between pre-miRNAs and their flanking regions. For each pre-miRNA, the flanking region with the same length as pre-miRNAs was used for analysis. The size of the circles represents the number of miRNAs.

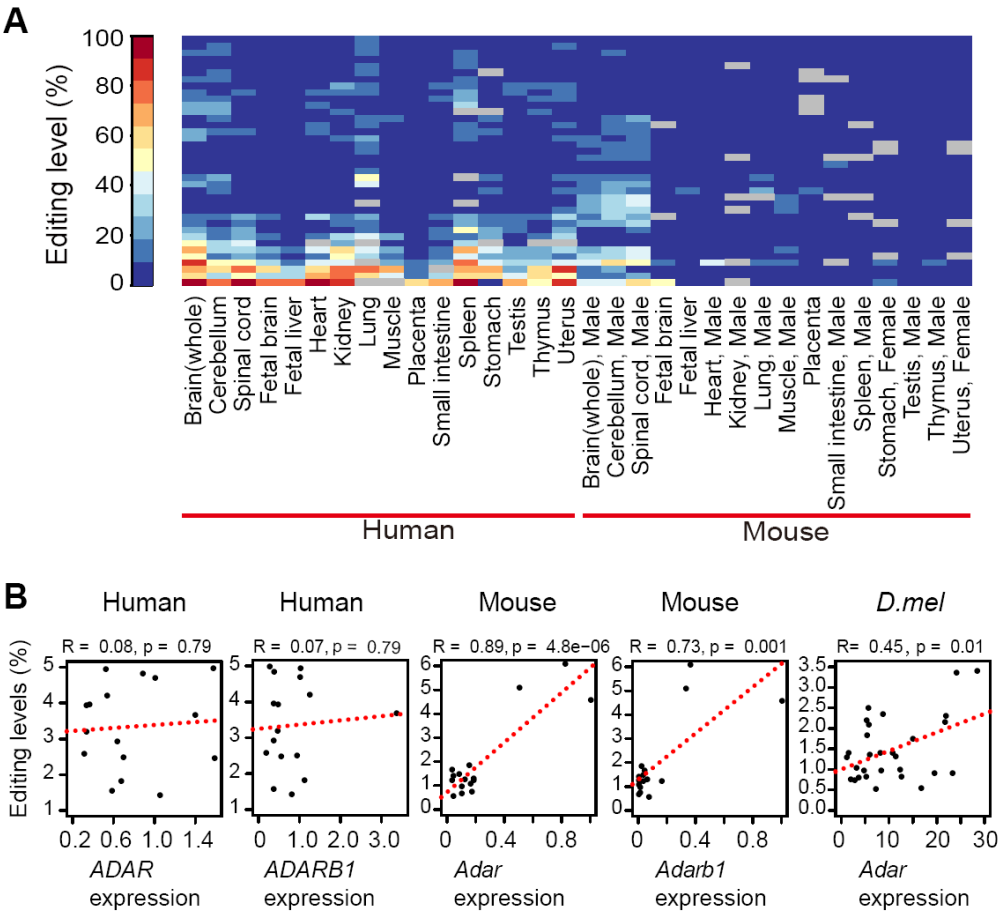
(D) Free energy comparison between pre-miRNA flanking regions and shuffled sequences. Upstream and downstream 150 nt sequence joined by a contiguous 15 nt A stretches were analyzed. Red, the free energy of real sequences < the mean free energies of sequences shuffled 100 times. Blue, the free energy of real sequences > the mean free energies of sequences shuffled 100 times. P values were calculated using Paired t-test. **P<0.01, ***P<0.001.

(E) The proportion of different editing triplets in human. Sites with representative editing level >20% were used for analysis.

(F) The proportion of different editing triplets, weighted by their editing levels, on pre-miRNAs in different species.

(G) Editing level comparison between different categories of editing sites in human and mouse. For this analysis, we used the representative editing level of each editing site, which is the maximum editing level across the 16 tissues profiled in both human and mouse. Conserved, sites edited in both human and mouse; Nonconserved (1), sites located at conserved miRNAs but only

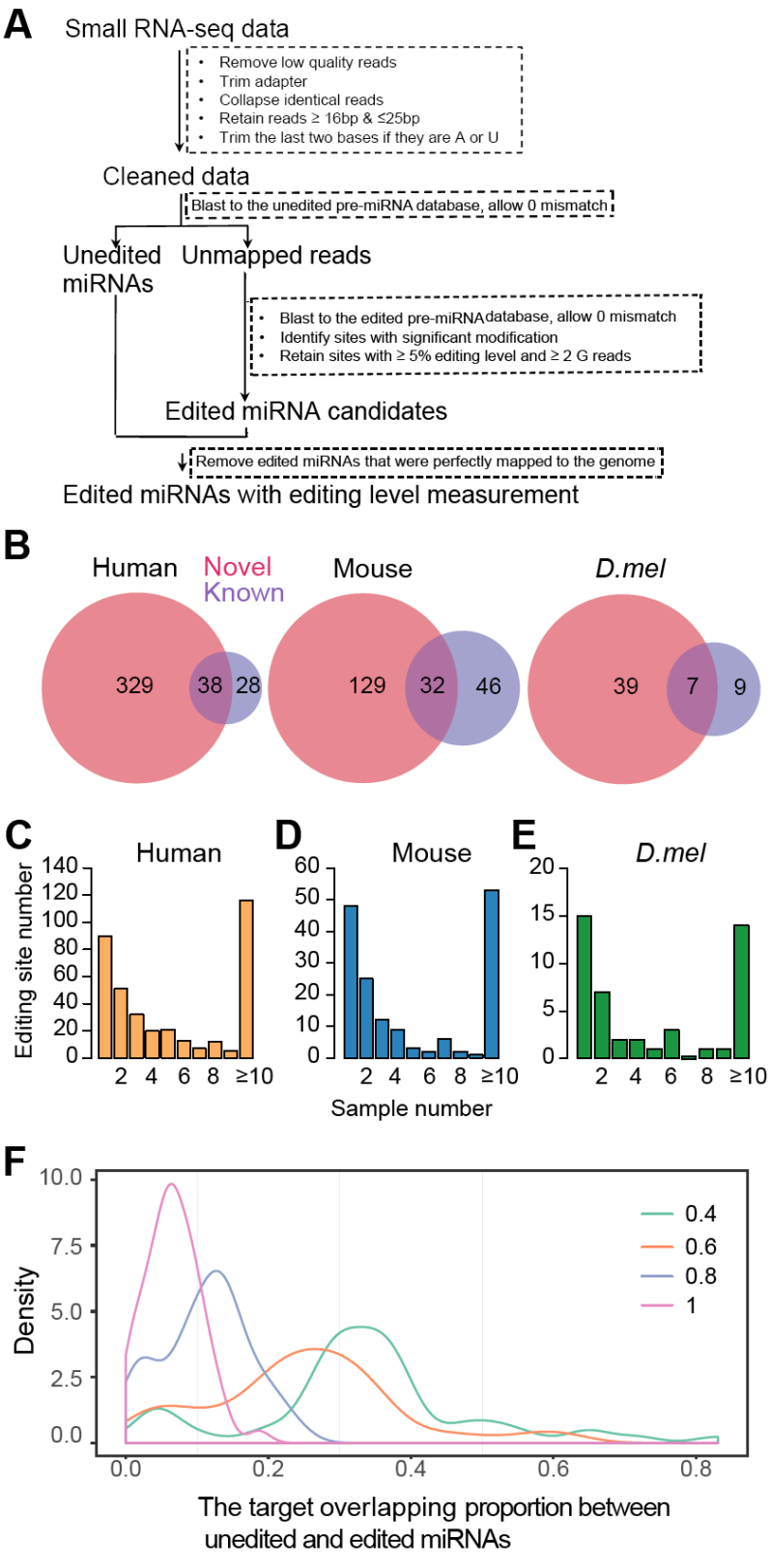
edited in one species; Nonconserved (2), sites located at lineage-specific miRNAs. P values were calculated using Wilcoxon test. **P<0.01, ***P<0.001.



Supplemental Figure S14. Editing level profiles of conserved sites.

(A) Heatmap showing editing levels of conserved sites for 16 human and mouse tissues.

(B) Correlations between *ADAR* expression levels and overall editing levels of each sample in human, mouse and *D.mel*. R values were calculated by linear regressions on overall editing levels and gene expressions measured by real-time PCR. The overall editing level of a sample is defined as the mean editing level of all sites.



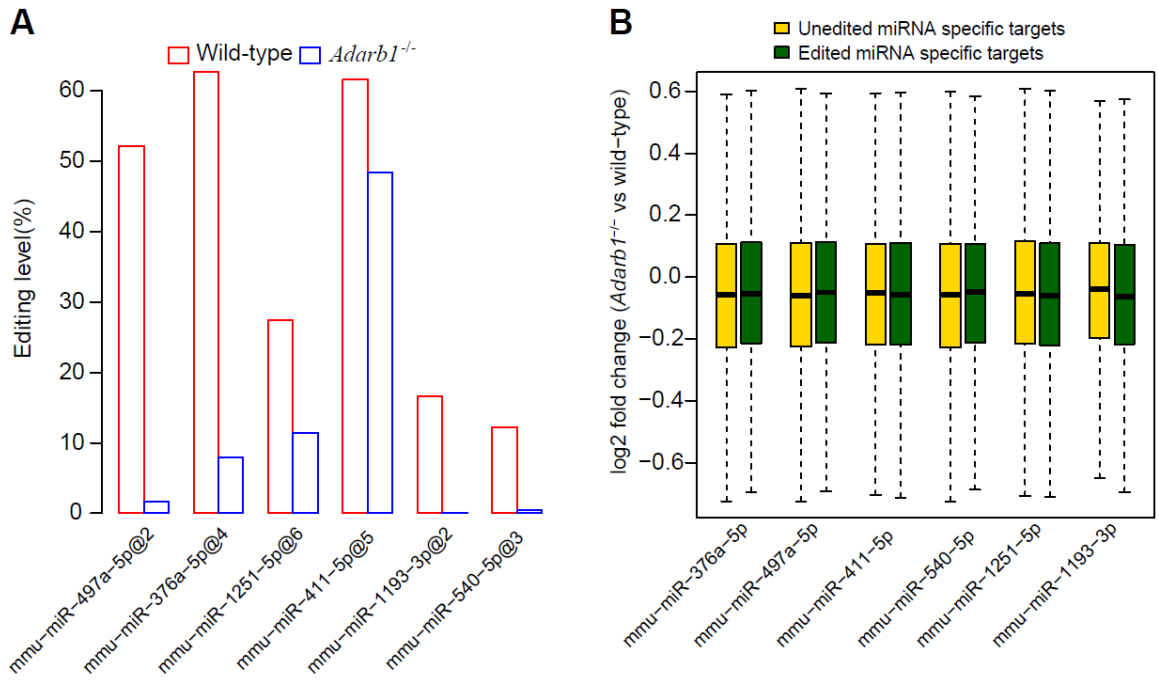
Supplemental Figure S15. Identification and analysis of edited mature miRNAs.

(A) Pipeline to identify edited mature miRNAs.

(B) The overlapping between previously unidentified edited mature miRNAs and known mature miRNAs.

(C-E) The distribution of the number of editing sites shared by multiple samples in human (**C**), mouse (**D**) and *D.mel* (**E**).

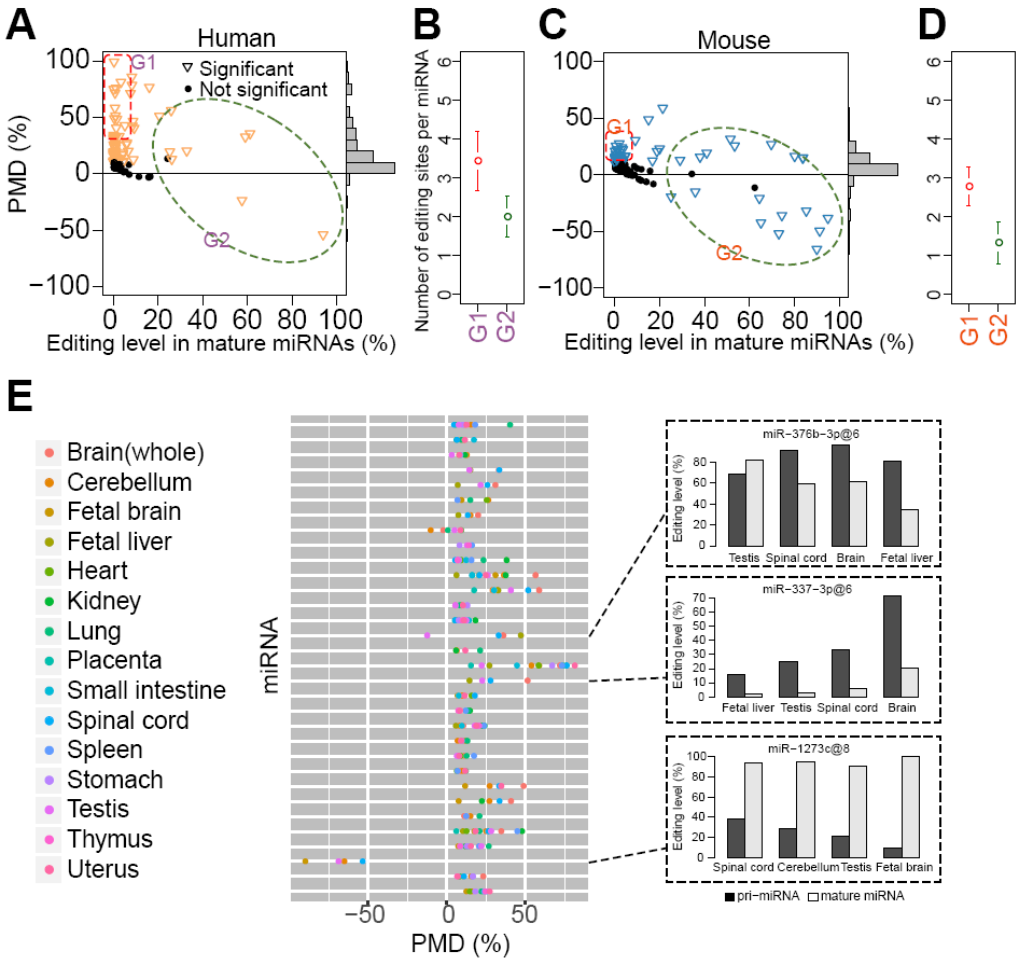
(F) The distribution of the overlapping proportions between the unedited and edited version of miRNAs with different TargetScan context++ cutoffs.



Supplemental Figure S16. The impact of miRNA editing on target gene expression changes.

(A) The editing level difference between wild-type and *Adarb1*$^{-/-}$ for the 6 selected miRNAs with editing sites in the seed region. The x axis is the miRNA_name@edited position.

(B) The comparison of target gene expression changes between the unedited miRNA specific targets and the edited miRNA specific targets. The y axis is the log2 fold change between *Adarb1*$^{-/-}$ and wild-type mice.



Supplemental Figure S17. RNA editing modulates miRNA processing.

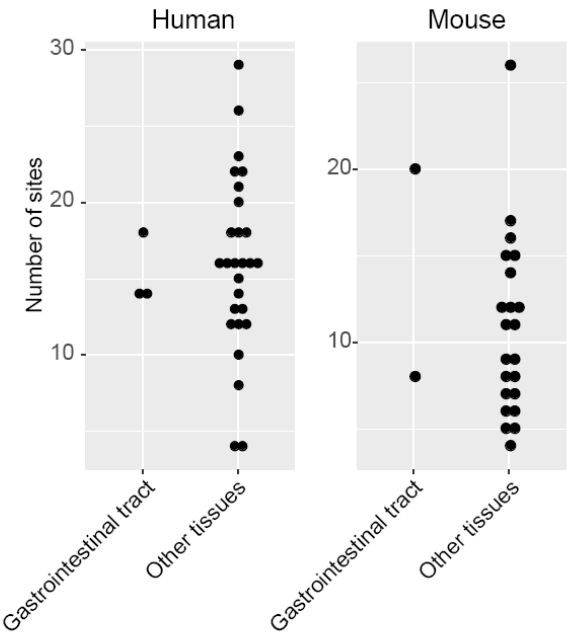
(A) The editing level difference between pri- and mature miRNAs in human. Sites that were (1) covered by at least 50 mmPCR-seq reads and 20 small RNA-seq reads, (2) with editing level ≥ 5% in pri-miRNAs, (3) located at mature miRNAs that can be mapped to a unique position in the genome were used for all PMD related analysis. For human data, we used 16 tissues (**Supplemental Table S7**) from which we generated both miR-mmPCR-seq and small RNA-seq data; for mouse data, we used 3 neural tissues (**Supplemental Table S7**) from which we generated both miR-mmPCR-seq and small RNA-seq data.

(B) The comparison of the number of editing sites per miRNA between group1 (G1, red rectangle) and group2 sites (G2, green oval) in **A**.

(C) The editing level difference between pri- and mature miRNAs in mouse.

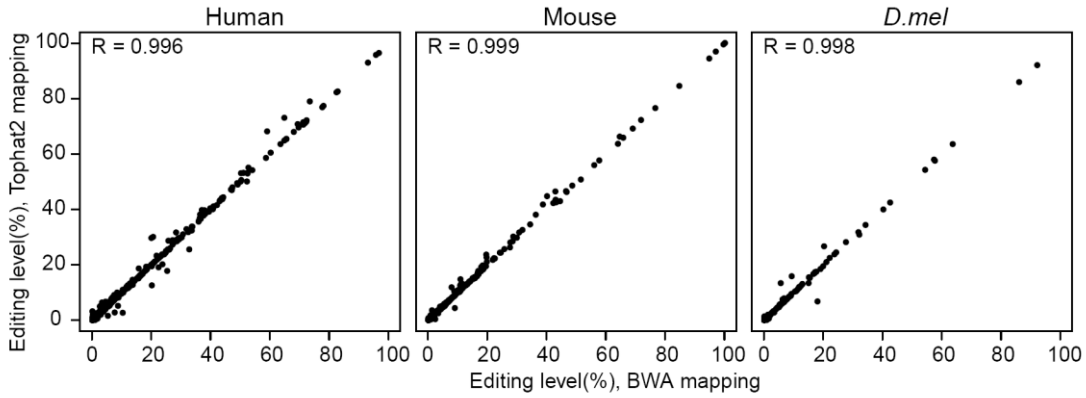
(D) The comparison of the number of editing sites per miRNA between group1 (G1, red rectangle) and group2 sites (G2, green oval) in **C**.

(E) PMD values in multiple tissues in human. Sites detected in at least 4 tissues were shown.



Supplemental Figure S18. The comparison of the number of putative C-to-U sites between the gastrointestinal tract and other tissues.

Sites with variant frequency >0.05 were used for analysis. The gastrointestinal tract tissues in human - small intestine, colon and stomach; in mouse - small intestine and stomach.



Supplemental Figure S19. Comparison of editing levels quantified by BWA or Tophat2. Three samples from different species were shown as examples: human (brain sample), mouse (brain sample), *D.mel* (female adult head sample).

Supplemental Notes

Supplemental Note 1. Gene-specific primers vs random hexamer primers.

Gene-specific primers typically work better than random hexamer primers to reverse-transcribe specific genes. However, we found that more pri-miRNA amplicons, as well as a more uniform amplification, were obtained using cDNAs generated with random primers. It is known that reverse transcription of dsRNA region may be difficult since a competition exists between the cDNA extension and the RNA hairpin formation. So it is possible that random hexamers may bind to the dsRNA region and prevent hairpin formation, which is helpful for reverse transcription of the pre-miRNA region.

Supplemental Note 2. Editing site identification in mouse B1 repeats.

A previous study suggested that separating filtering criteria for identifying *Alu* and non-*Alu* sites increases the sensitivity of editing site identification in *Alu* elements (Ramaswami et al. 2012), which are the major substrates of ADARs in human (Kim et al. 2004; Levanon et al. 2004). Mouse B1 repeats are also in *Alu* family. Both B1 and *Alu* are derived from the 7SL RNA, *Alu* is longer than the equivalent mouse B1, and the dsRNAs formed in human are longer. To ask if we could detect more mouse editing sites by separated filtering criteria for identifying B1 and non-B1 sites, we called editing sites in B1 without or with additional filtering steps, as described previously (Ramaswami et al. 2012). We found an average of 7 editing sites in B1 per tissue. Additionally, we detected a similar number of editing sites without or with additional filtering steps (**Supplemental Fig. S2**). This result is consistent with a previous study (Neeman et al. 2006) that found no enrichment of editing sites in known mouse repeat elements, including B1.

Supplemental Note 3. ADAR-binding sequencing motif analysis.

ADARs target dsRNA of any sequence but have preferences for certain neighboring nucleotides, in particular, the triplet containing nucleotides immediately adjacent to the edited

adenosine (Polson and Bass 1994; Li et al. 2009; Eggington et al. 2011). We therefore examined the triplet for all editing sites. As expected, we observed the under- and overrepresentation of guanosines at the positions directly preceding and following the edited site, respectively (**Supplemental Fig. S6C**), consistent with the known nucleotide neighbor preferences of ADARs. This result suggests that the newly identified A-to-G sites are enriched with *bona fide* A-to-I RNA editing sites.

Supplemental Note 4. Evaluation of the performance of miR-mmPCR-seq.

To further validate whether the identified A-to-G sites were A-to-I editing events, we analyzed nascent RNA-seq for the *D. mel* wild-type strain and the $Adar^{5G1}$ mutant that eliminates RNA editing (**Methods**). We examined all A-to-G sites that were edited in the wild-type sample. As expected, we achieved high accuracy; 96.7% of all A-to-G sites showed only A in the $Adar^{5G1}$ sample (**Supplemental Fig. S8A**). Finally, editing level measurements by mmPCR-seq are highly reproducible between samples reverse-transcribed using different transcriptases, technical replicates (**Supplemental Fig. S8B-D**), and biological replicates in different mouse tissues (**Supplemental Fig. S8E**), suggesting the editing measurement is accurate, as previously described (Zhang et al. 2014).

Supplemental Note 5. The impact of miRNA editing on target gene expression changes.

miRNA editing sites, particularly the ones in the seed region, can alter miRNA targeting and affect target gene expression. Such effect has been previously validated via *in vitro* assays such as the reporter assay focused on individual targets, or the overexpression of an edited miRNA in cell lines and subsequent gene expression profiling (Kawahara et al. 2007; Kume et al. 2014). Here we sought to ask if edited miRNAs may affect the expression levels of target genes *in vivo*, using mouse brain models depleted ADARB1 editing. We first analyzed the small RNA-seq data of adult brain samples from wild-type and *Adarb1* knockout mice. We identified 6 miRNAs which have editing sites in the seed region and the editing levels were mainly controlled by

ADARB1 (**Supplemental Fig. S16A,** editing level difference between wild-type and knockout mice is >10%). We next used the TargetScan algorithm to predict miRNA target sites for the unedited and edited versions of these 6 miRNAs. We then compared the target gene expression changes between unedited miRNA specific targets and edited miRNA specific targets. No statistically significant difference was observed (**Supplemental Fig. S16B**). ADARs do not only mediate miRNA editing, but also regulate miRNA processing, mRNA editing and mRNA expression. For example, the knockdown of ADARs dramatically affects the expression of thousands of genes, independent of RNA editing (Wang et al. 2013). So this may be the reason that we are not able to detect the subtle gene expression changes induced by the change of individual miRNA editing event using *Adarb1* knockout samples.

REFERENCES

Carr IM, Robinson JI, Dimitriou R, Markham AF, Morgan AW, Bonthron DT. 2009. Inferring relative proportions of DNA variants from sequencing electropherograms. *Bioinformatics* **25**(24): 3244-3250.

Eggington JM, Greene T, Bass BL. 2011. Predicting sites of ADAR editing in double-stranded RNA. *Nat Commun* **2**: 319.

Kawahara Y, Zinshteyn B, Sethupathy P, Iizasa H, Hatzigeorgiou AG, Nishikura K. 2007. Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science* **315**(5815): 1137-1140.

Kim DD, Kim TT, Walsh T, Kobayashi Y, Matise TC, Buyske S, Gabriel A. 2004. Widespread RNA editing of embedded alu elements in the human transcriptome. *Genome Res* **14**(9): 1719-1725.

Kume H, Hino K, Galipon J, Ui-Tei K. 2014. A-to-I editing in the miRNA seed region regulates target mRNA selection and silencing efficiency. *Nucleic Acids Res* **42**(15): 10050-10060.

Levanon EY, Eisenberg E, Yelin R, Nemzer S, Hallegger M, Shemesh R, Fligelman ZY, Shoshan A, Pollock SR, Sztybel D et al. 2004. Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat Biotechnol* **22**(8): 1001-1005.

Li JB, Levanon EY, Yoon JK, Aach J, Xie B, Leproust E, Zhang K, Gao Y, Church GM. 2009. Genome-wide identification of human RNA editing sites by parallel DNA capturing and sequencing. *Science* **324**(5931): 1210-1213.

Neeman Y, Levanon EY, Jantsch MF, Eisenberg E. 2006. RNA editing level in the mouse is determined by the genomic repeat repertoire. *RNA* **12**(10): 1802-1809.

Polson AG, Bass BL. 1994. Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *Embo J* **13**(23): 5701-5711.

Ramaswami G, Lin W, Piskol R, Tan MH, Davis C, Li JB. 2012. Accurate identification of human Alu and non-Alu RNA editing sites. *Nat Methods* **9**: 579-581.

Wang IX, So E, Devlin JL, Zhao Y, Wu M, Cheung VG. 2013. ADAR regulates RNA editing, transcript stability, and gene expression. *Cell Rep* **5**(3): 849-860.

Zhang R, Li X, Ramaswami G, Smith KS, Turecki G, Montgomery SB, Li JB. 2014. Quantifying RNA allelic ratios by microfluidic multiplex PCR and sequencing. *Nat Meth* **11**(1): 51-54.