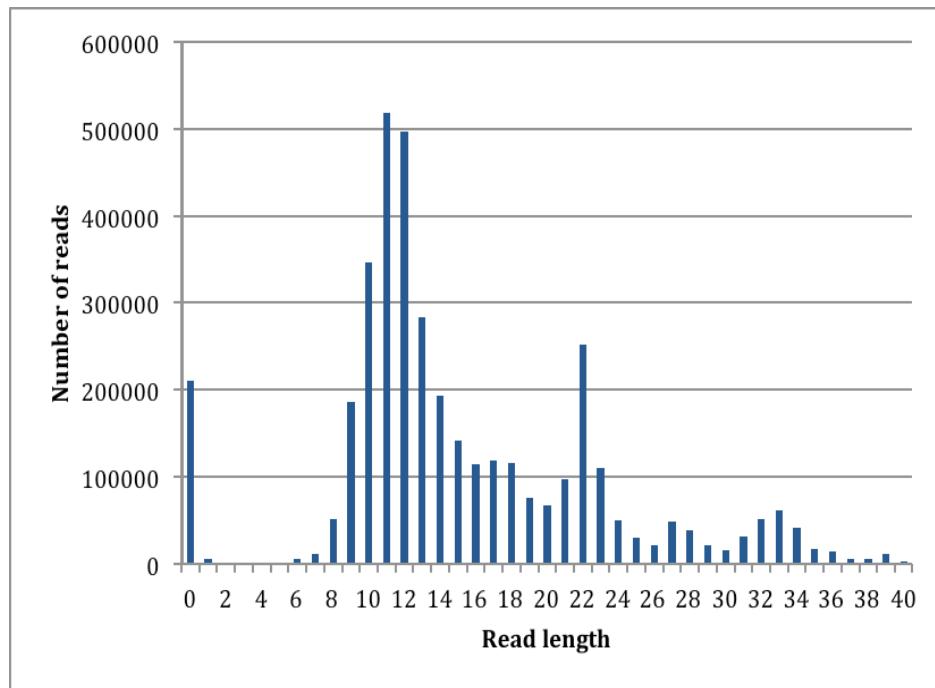
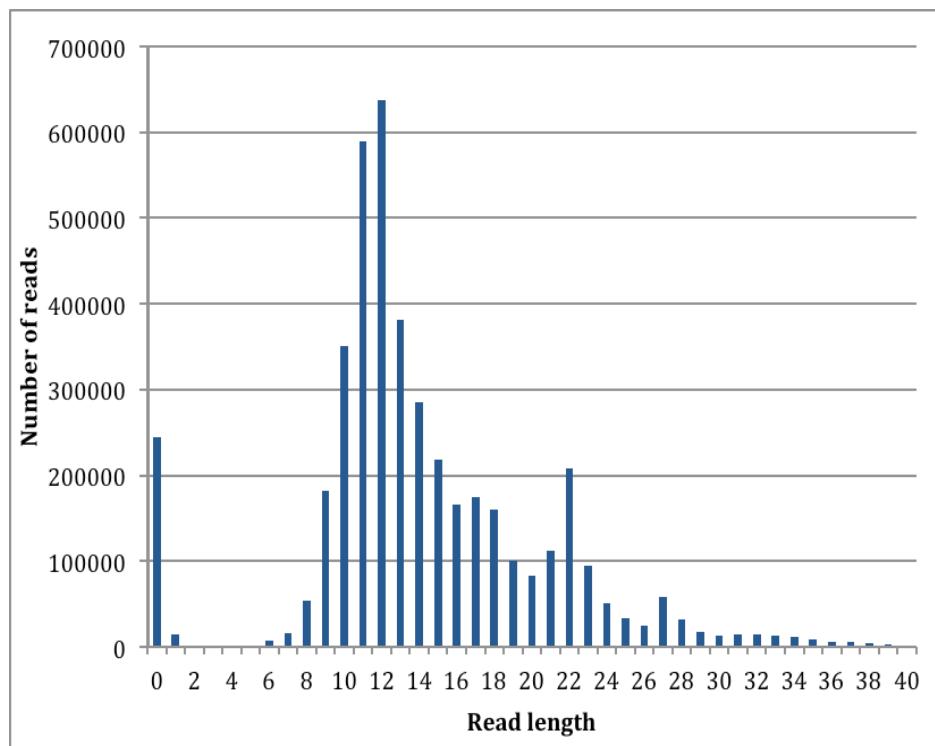


Supplementary material Ekdahl et al.

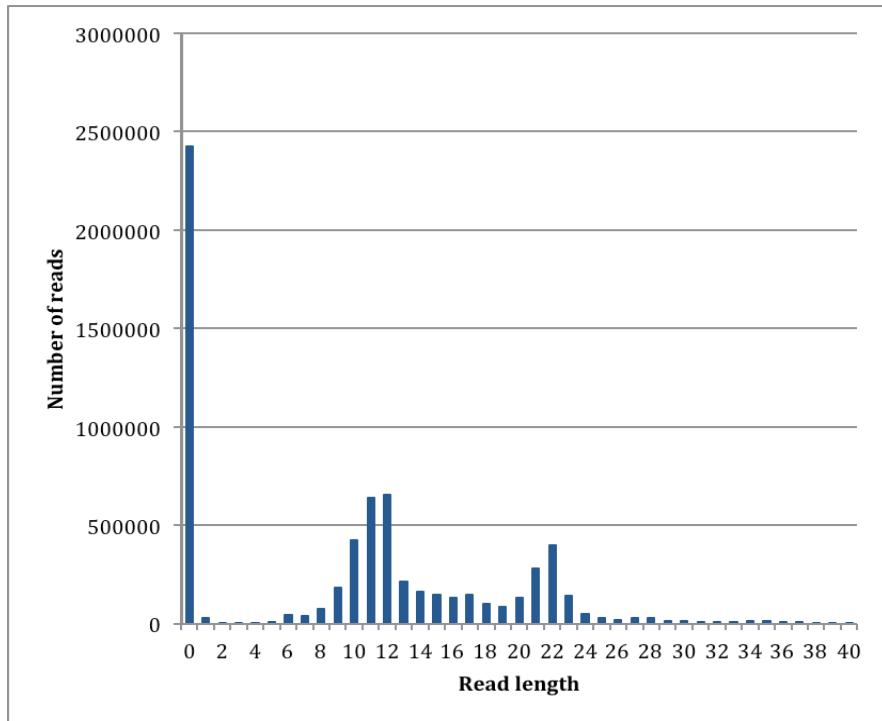
S1. Distribution of lengths of SOLiD reads after removing the adapter sequences



Supplemental Figure 1: E15 mouse, read lengths



Supplemental Figure 2: P2 mouse, read lengths

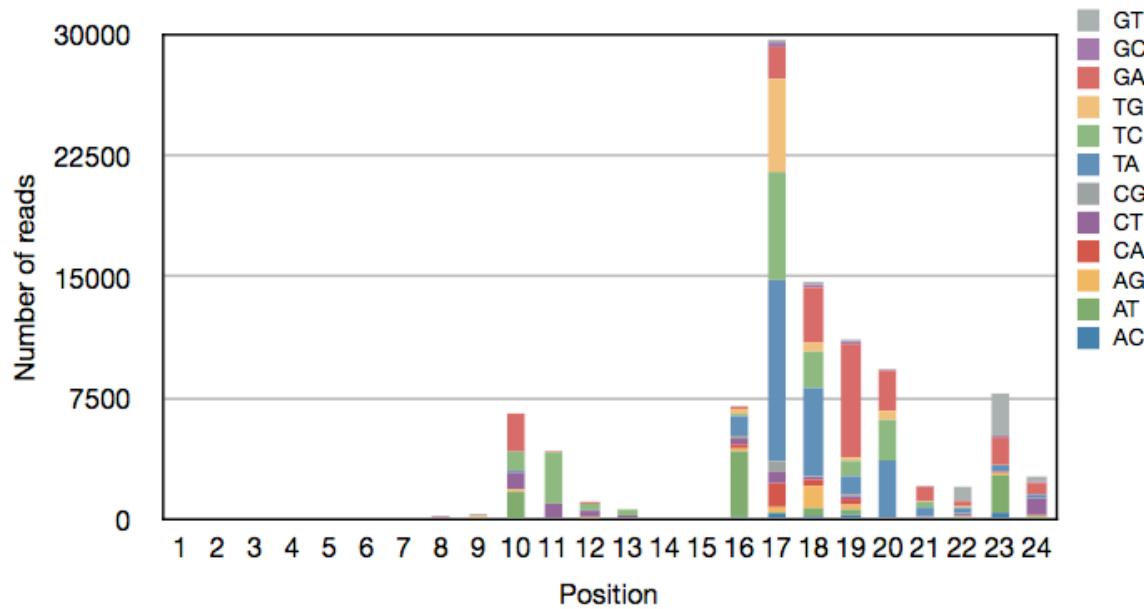


Supplemental Figure 3: P21 mouse, read lengths

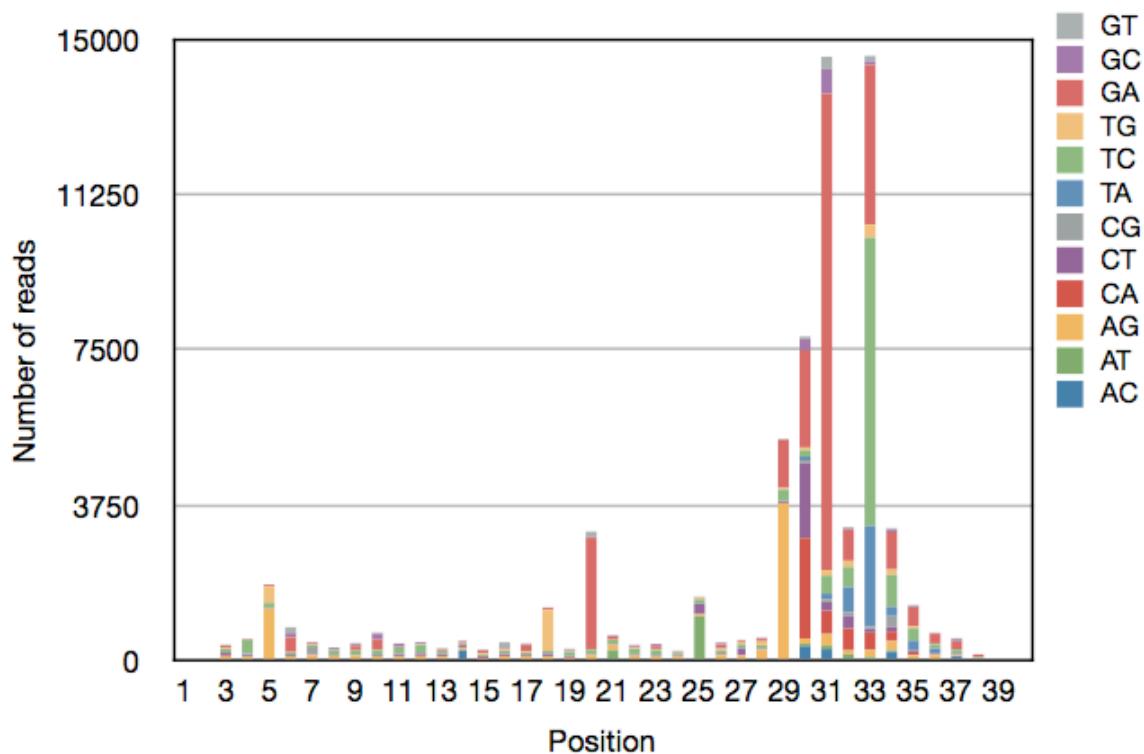
S2. Position distribution of mismatches in the SOLiD and Illumina reads

We also analyzed mismatches in SOLiD exonic mRNA reads of (1) length 19 to 24 and (2) length 35 to 40. All reads of the right length were mapped to the mouse genome and, for sequences mapped to exons, the distribution of mismatches across their lengths was recorded. For exonic reads of length 19 to 24, the distribution of A:G mismatches and all other types of mismatches are shown in Supplemental Figures 6 and 7, respectively. Supplemental Figure 5 shows the distribution of all types of mismatches in reads of length 35 to 40. As easily seen, both these groups of exonic reads have a mismatch peak at the 3' end, with distributions similar to that of miRNA reads.

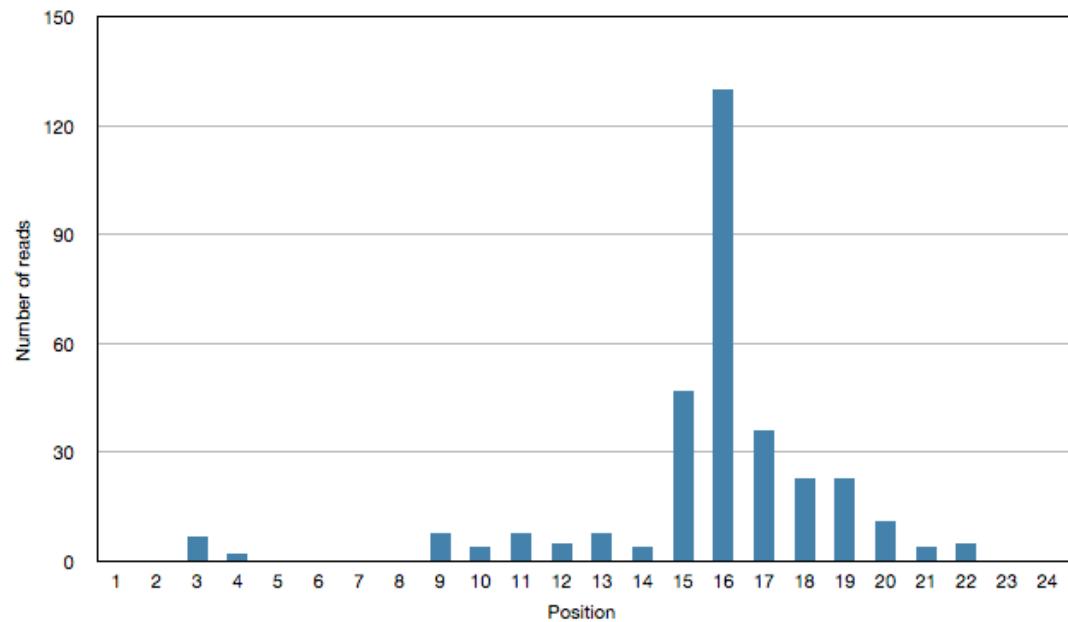
As a comparison, we investigated mismatches in Illumina miRNA reads from (Chi et al. 2009). As seen in Supplemental Figure 8, the Illumina miRNA reads do not have a mismatch peak at the 3' end. Based on this multiple evidence, we conclude that the increased frequency of mismatches in the 3' end of the miRNAs was caused by technical errors specific to the SOLiD platform. Of course, it may be considered conceivable that the 3' mismatch peak was due to a biological phenomenon, responsible for modifying RNA sequences at the 3' end. Given that the mismatches do not appear in the Illumina reads, this would imply that the Illumina reads contain many sequencing errors where, in more or less, each case the nucleotide by chance has been read as the unmodified nucleotide appearing in the genome. Since the mismatches are of all types, this cannot be due to a systematic error and the P-value for all these assumed sequencing errors is extremely low.



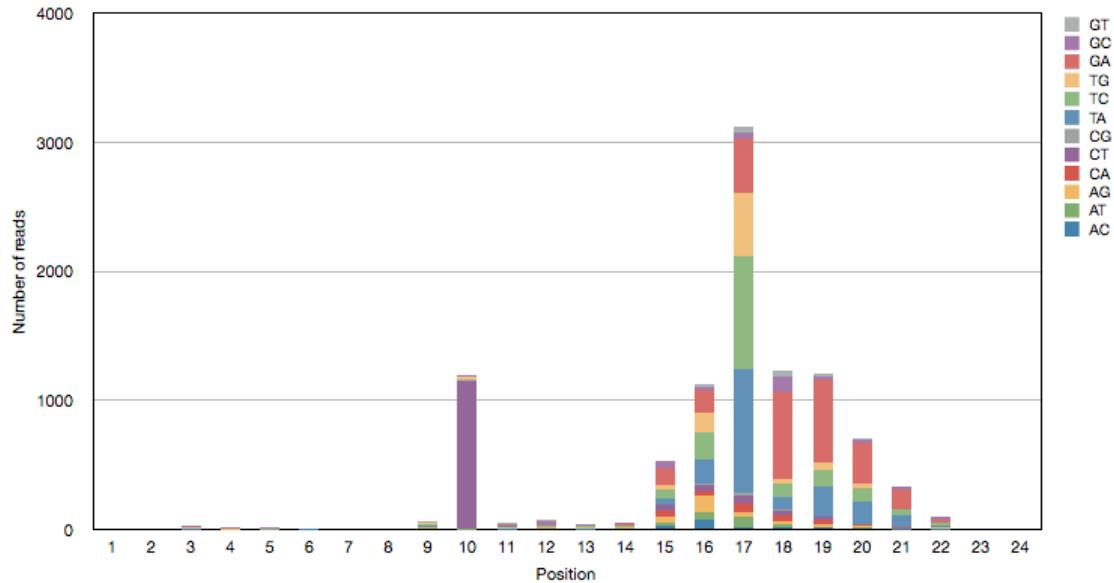
Supplemental Figure 4: Bar chart for all mismatch types in miRNA reads at developmental stage E15 from SOLiD. The chart shows alignment of the reads against the miRBase.



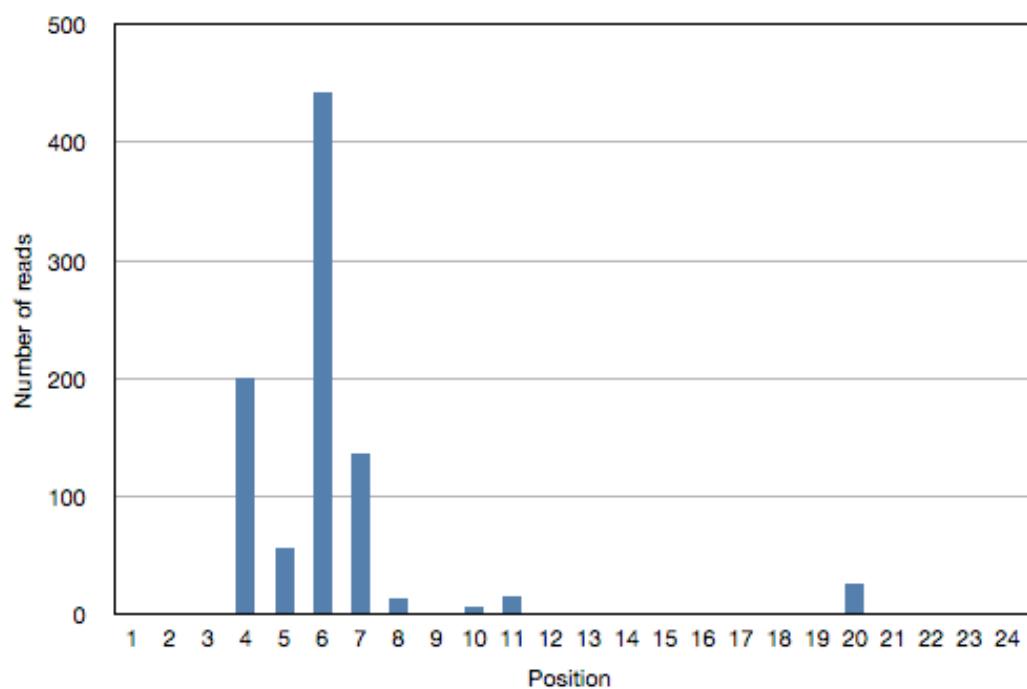
Supplemental Figure 5: Number of all types of mismatches in 35-40 nt long reads from SOLiD of non-specific RNA from developmental day E15.



Supplemental Figure 6: Exonic A:G mismatches in read of length 19-24 in the SOLiD data for E15



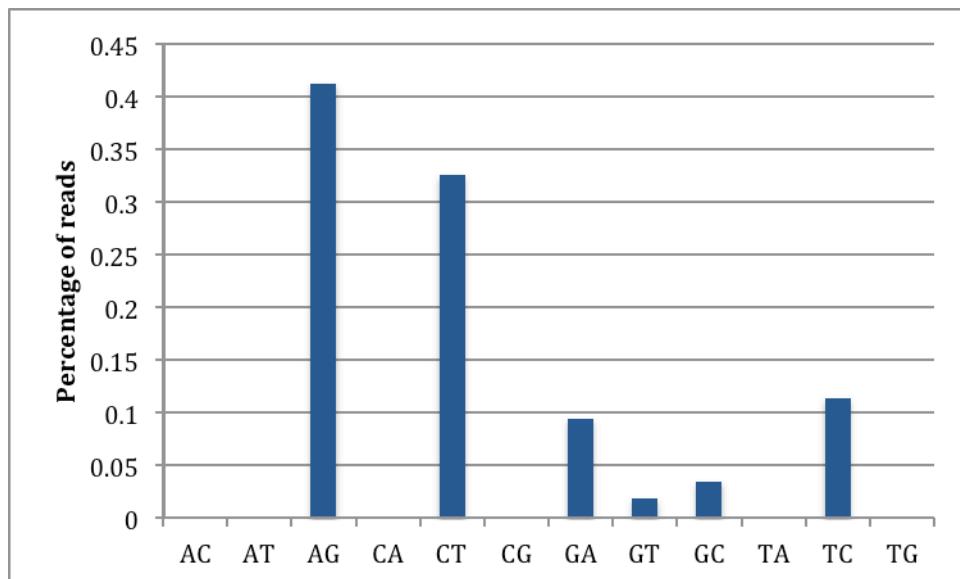
Supplemental Figure 7: Exonic mismatches for all mismatch types in reads of length 19-24 in the SOLiD data for E15



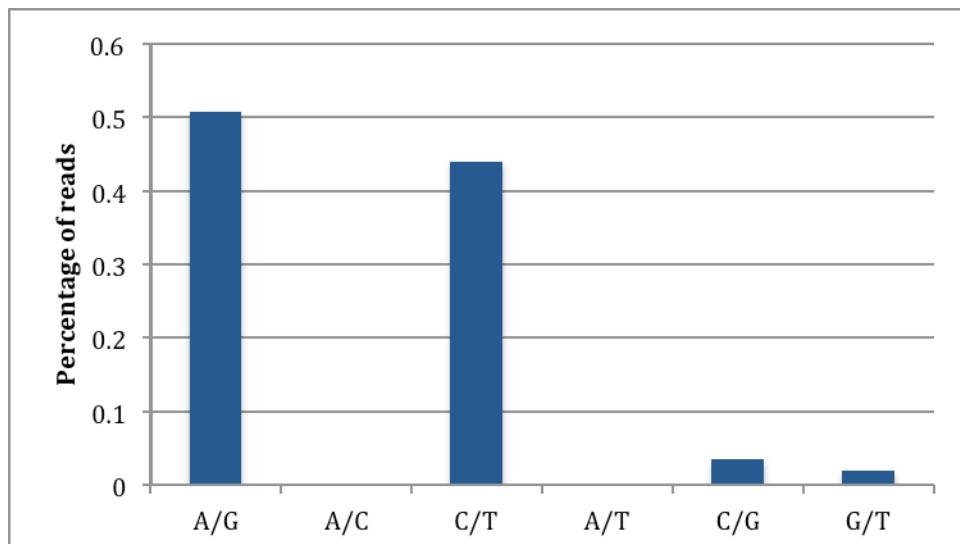
Supplemental Figure 8: A:G sites in brain A Illumina data

S3. Mismatch frequencies of miRNAs

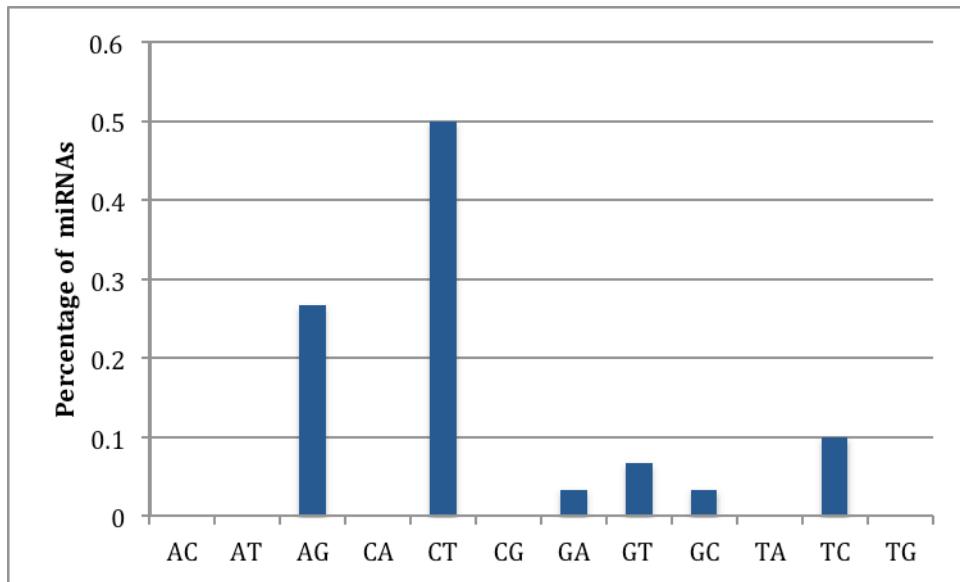
A/G means A:G and G:A



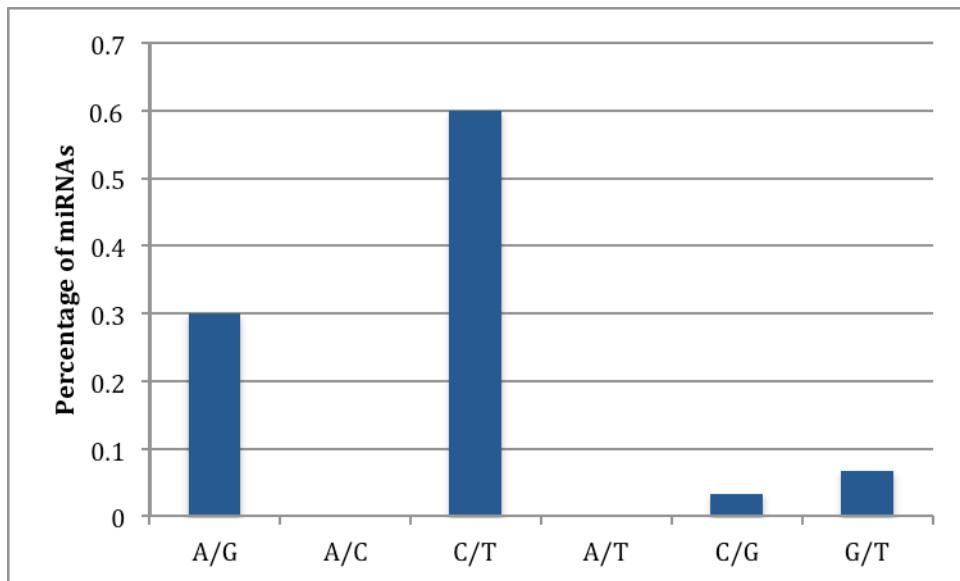
Supplemental Figure 9: Distribution of reads with different mismatches in intersection of SOLiD and Illumina data with 1-safety filter



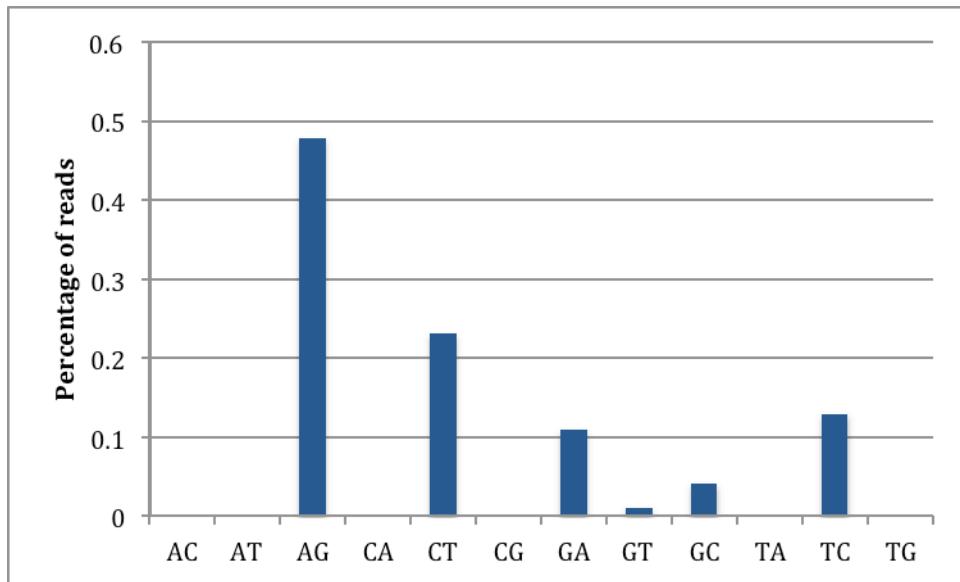
Supplemental Figure 10: Distribution of reads with different mismatches in intersection of SOLiD and Illumina data with 1-safety filter. A/G is sum of A:G and G:A



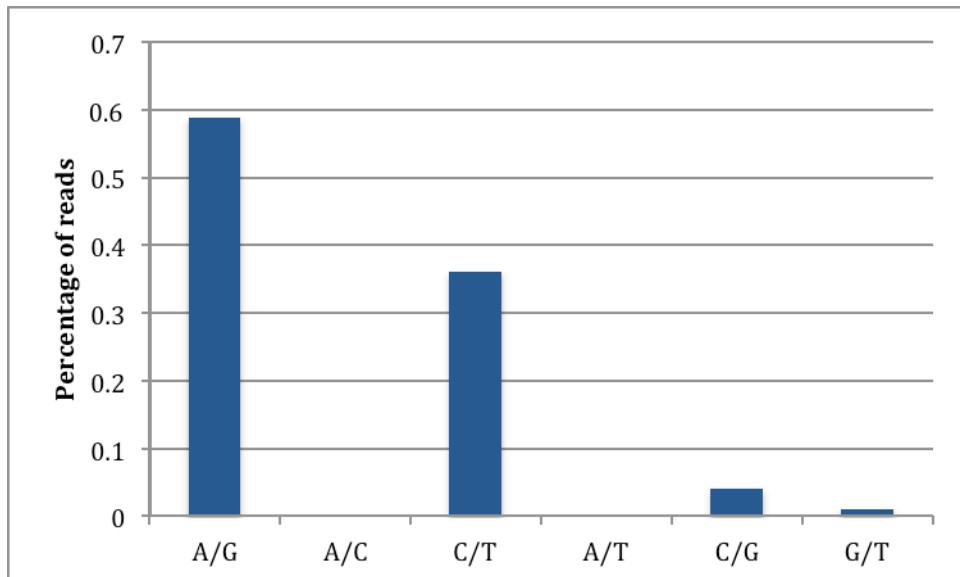
Supplemental Figure 11: Distribution of miRNA species with different mismatches in intersection of SOLiD and Illumina data with 1-safety filter



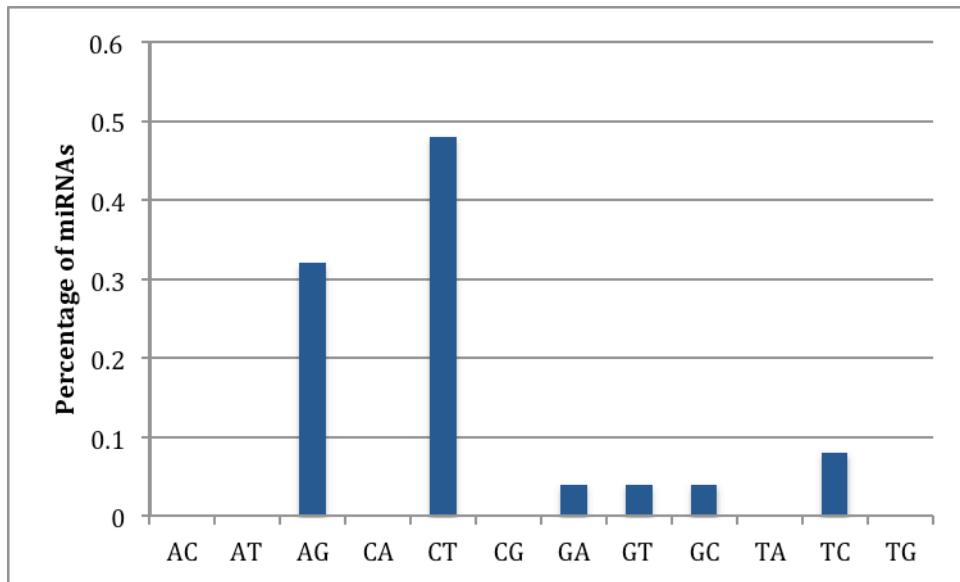
Supplemental Figure 12: Distribution of miRNA species with different mismatches in intersection of SOLiD and Illumina data with 1-safety filter



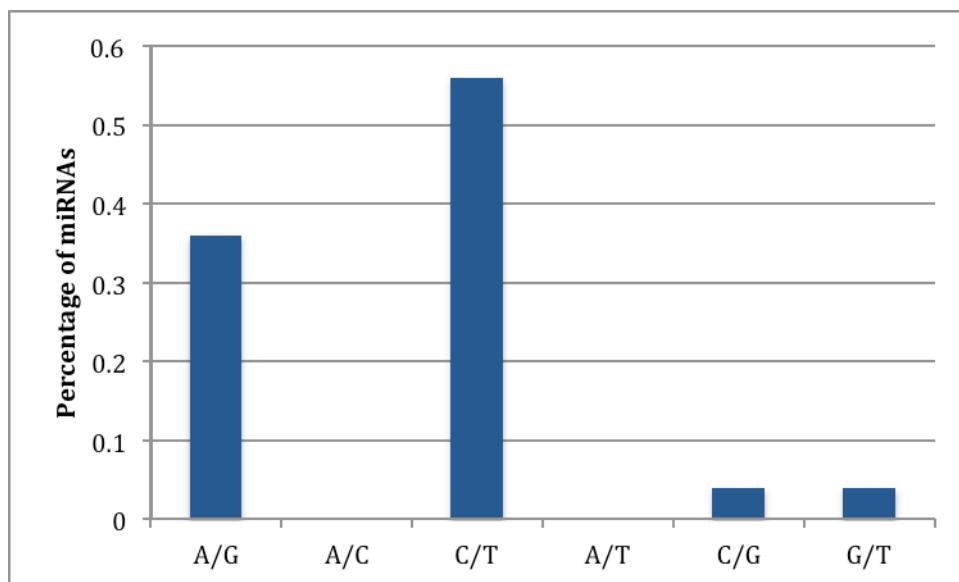
Supplemental Figure 13: Distribution of reads with different mismatches in intersection of SOLiD and Illumina data with 4-safety filter



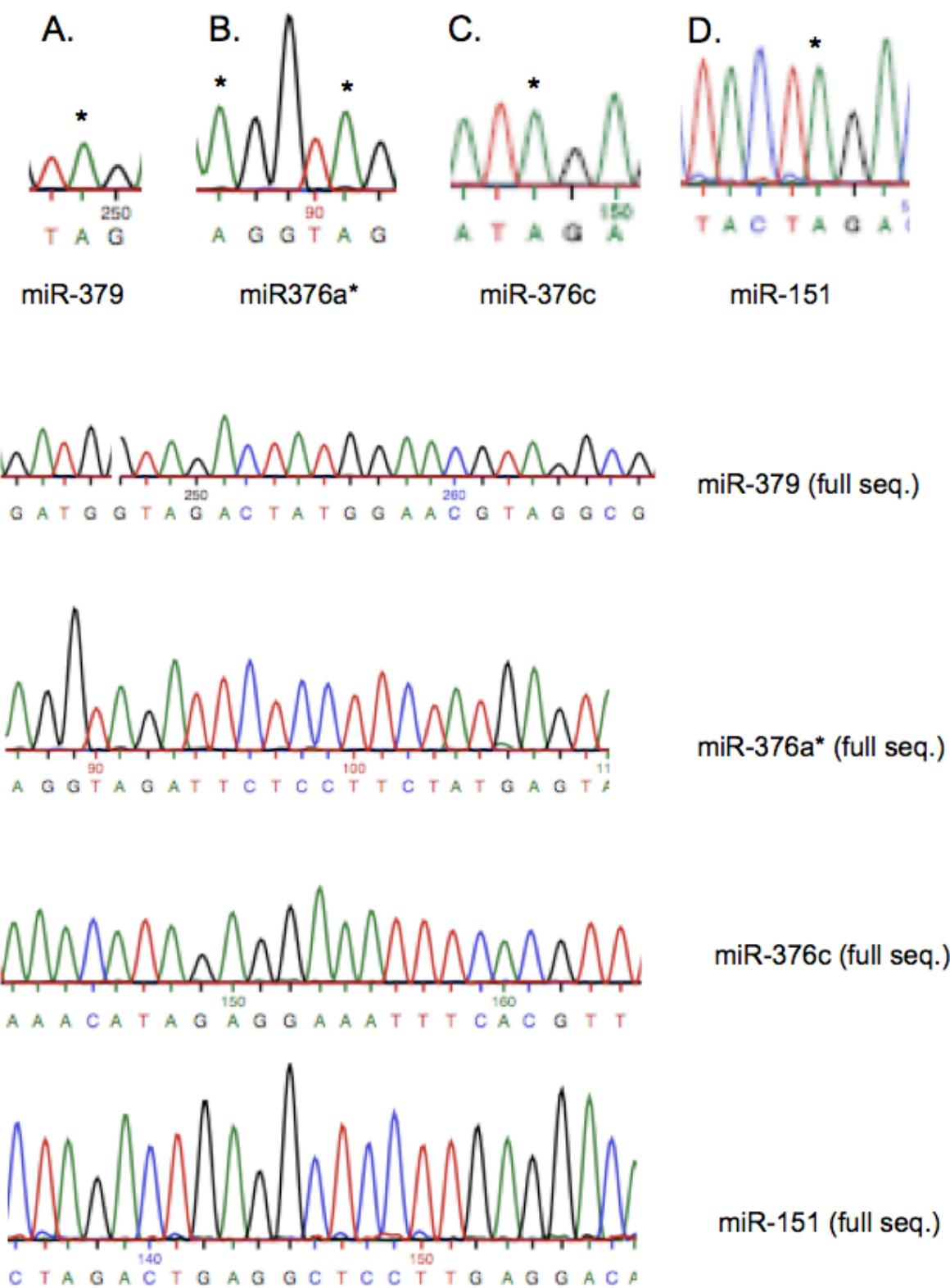
Supplemental Figure 14: Distribution of reads with different mismatches in intersection of SOLiD and Illumina data with 4-safety filter



Supplemental Figure 15: Distribution of miRNA species with different mismatches in intersection of SOLiD and Illumina data with 4-safety filter



Supplemental Figure 16: Distribution of miRNA species with different mismatches in intersection of SOLiD and Illumina data with 4-safety filter



Supplemental Figure 17: Genomic verification of a subset of edited miRNAs. Edited miRNAs experimentally verified by sequencing DNA from the same individual. A-to-I editing sites were identified as an A in the genomic sequence. A. miR-379. B. miR376a*. C. miR-376c D. miR-151

References

Chi, S.W., Zang, J.B., Mele, A. & Darnell, R.B. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 460, 479–486 (2009).