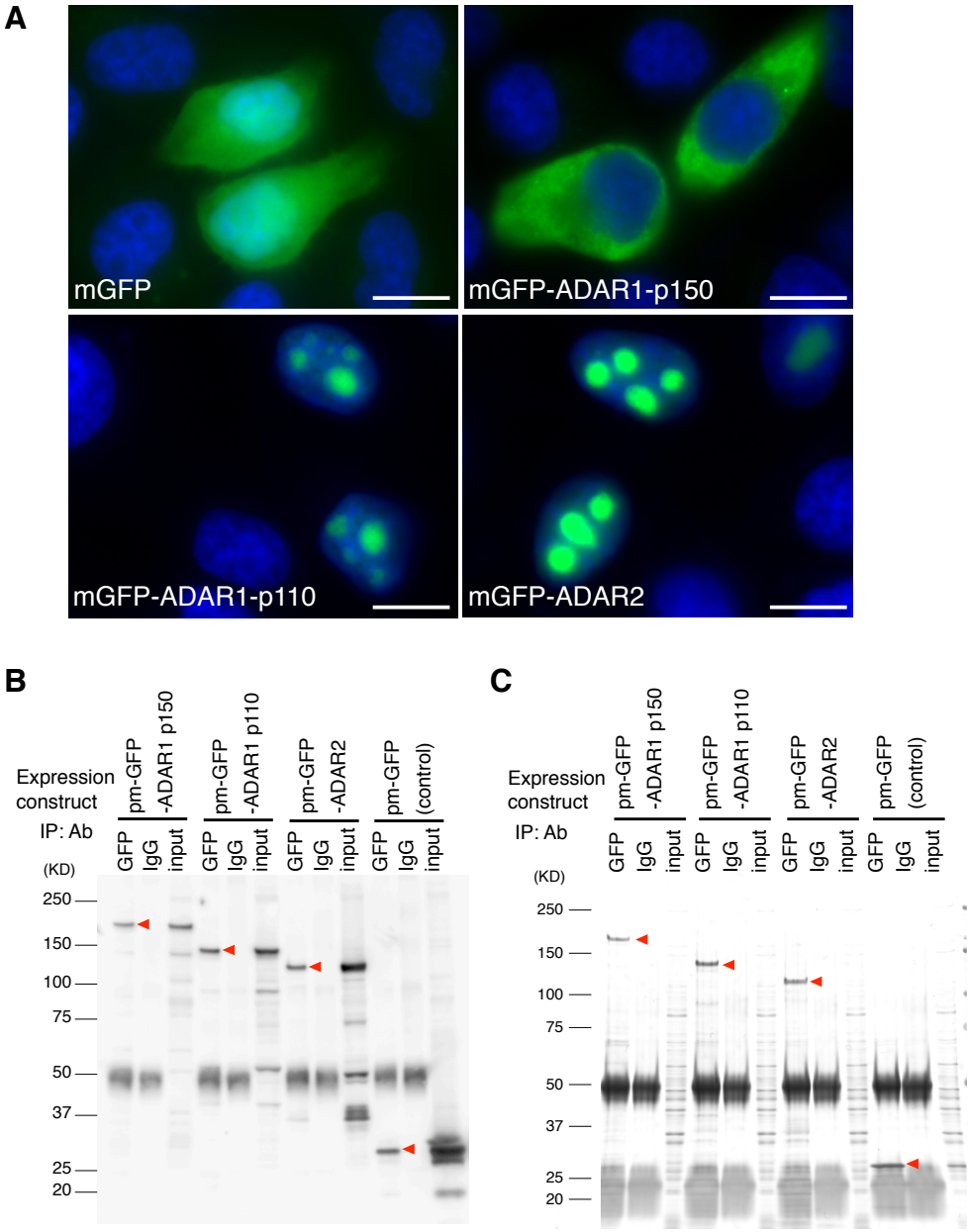


Figure S1. Subcellular localization of human ADAR isoforms, and verification of samples prior to RIP-seq by western blot and silver staining. (A) Fluorescence microscopy showing GFP signal of each overexpessed mycGFP-tagged ADAR isoform or mycGFP expression constructs in HeLa cells 48h post-transfection. Green indicates GFP signal, and blue indicates the nucleus stained by DATI. Bar, 20 μm.Western blot (B) and silver staining (C) of input samples and IP samples immunoprecipitated with anti-GFP antibody (GFP) or control IgG from cells transfected with each ADAR-expressing construct (pmGFP-ADAR1-150, pmGFP-ADAR1-110, and pmGFP-ADAR2) or pmGFP control. RIP-seq was carried out using input samples and IP samples with anti-GFP antibody. Red arrowheads indicate the positions of the expected molecular weights for each mGFP-tagged protein.

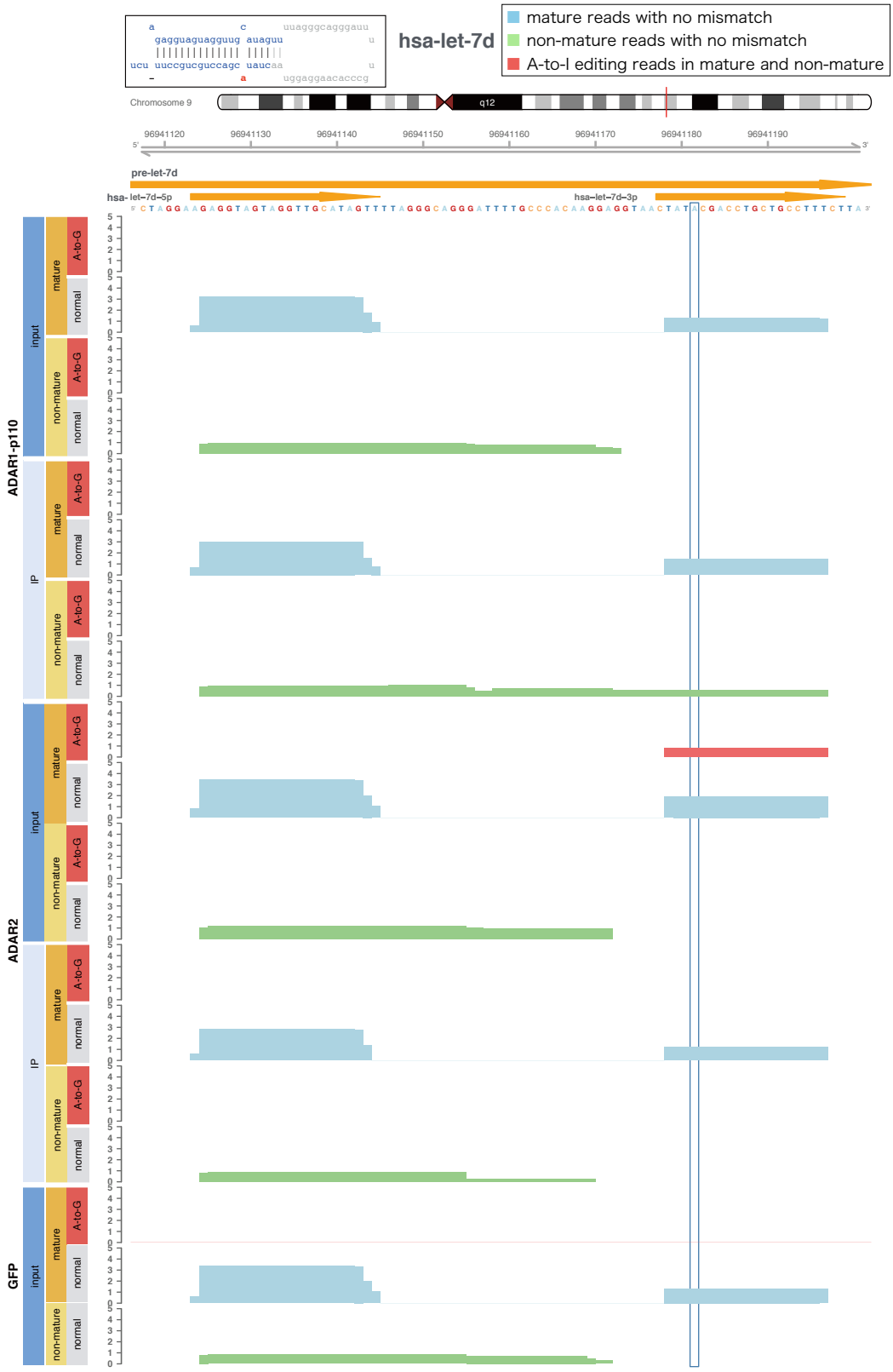
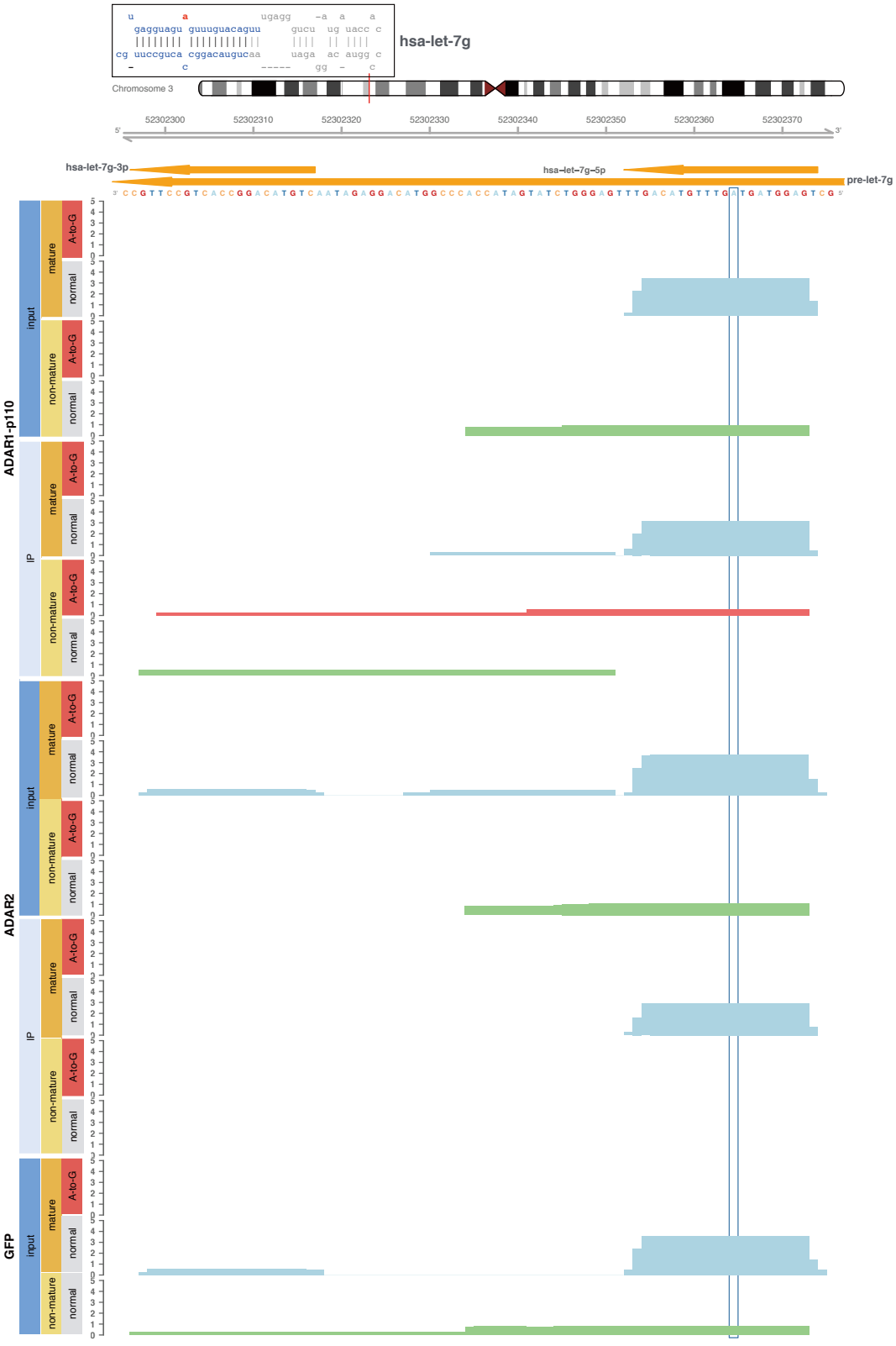
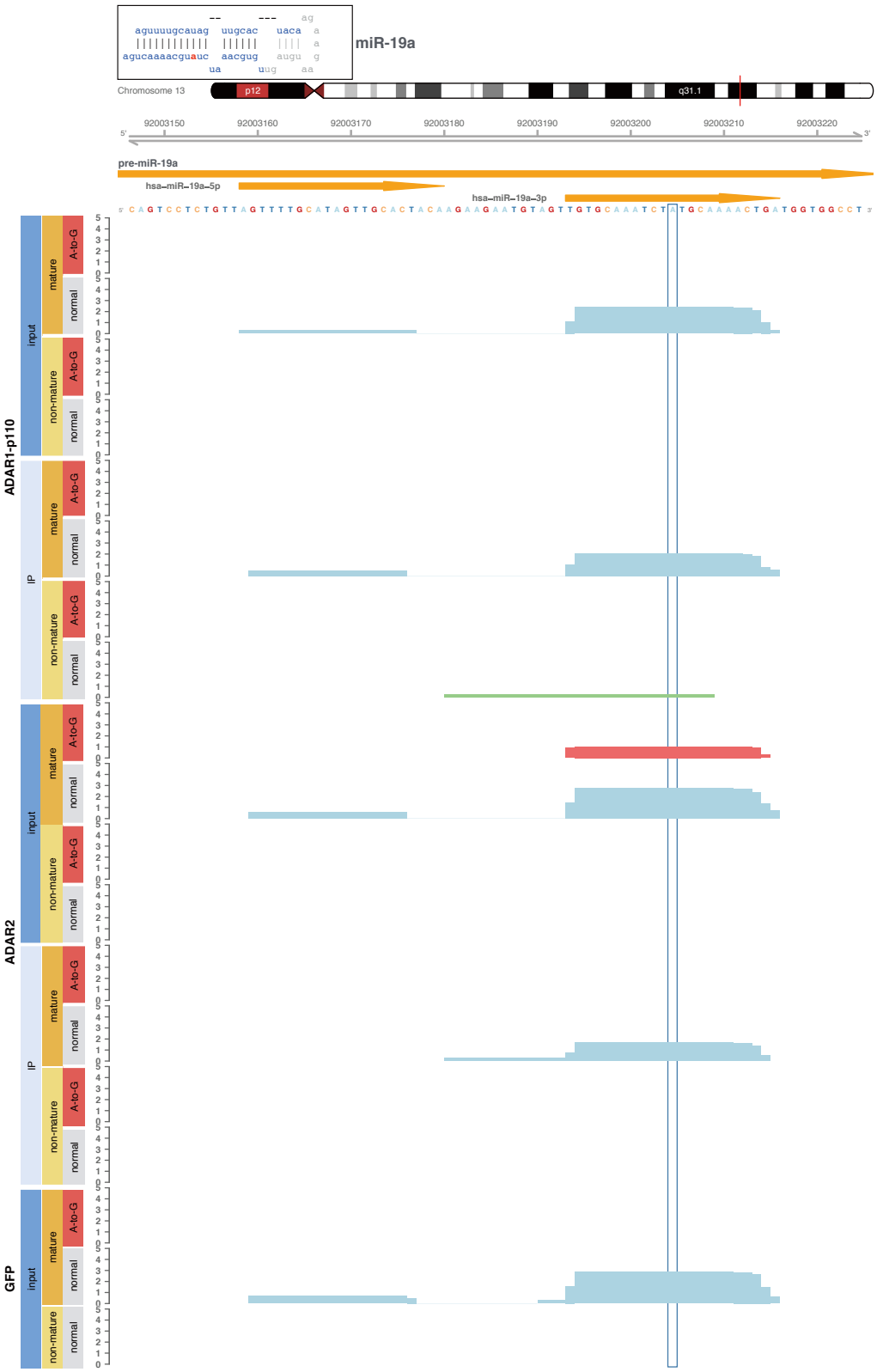
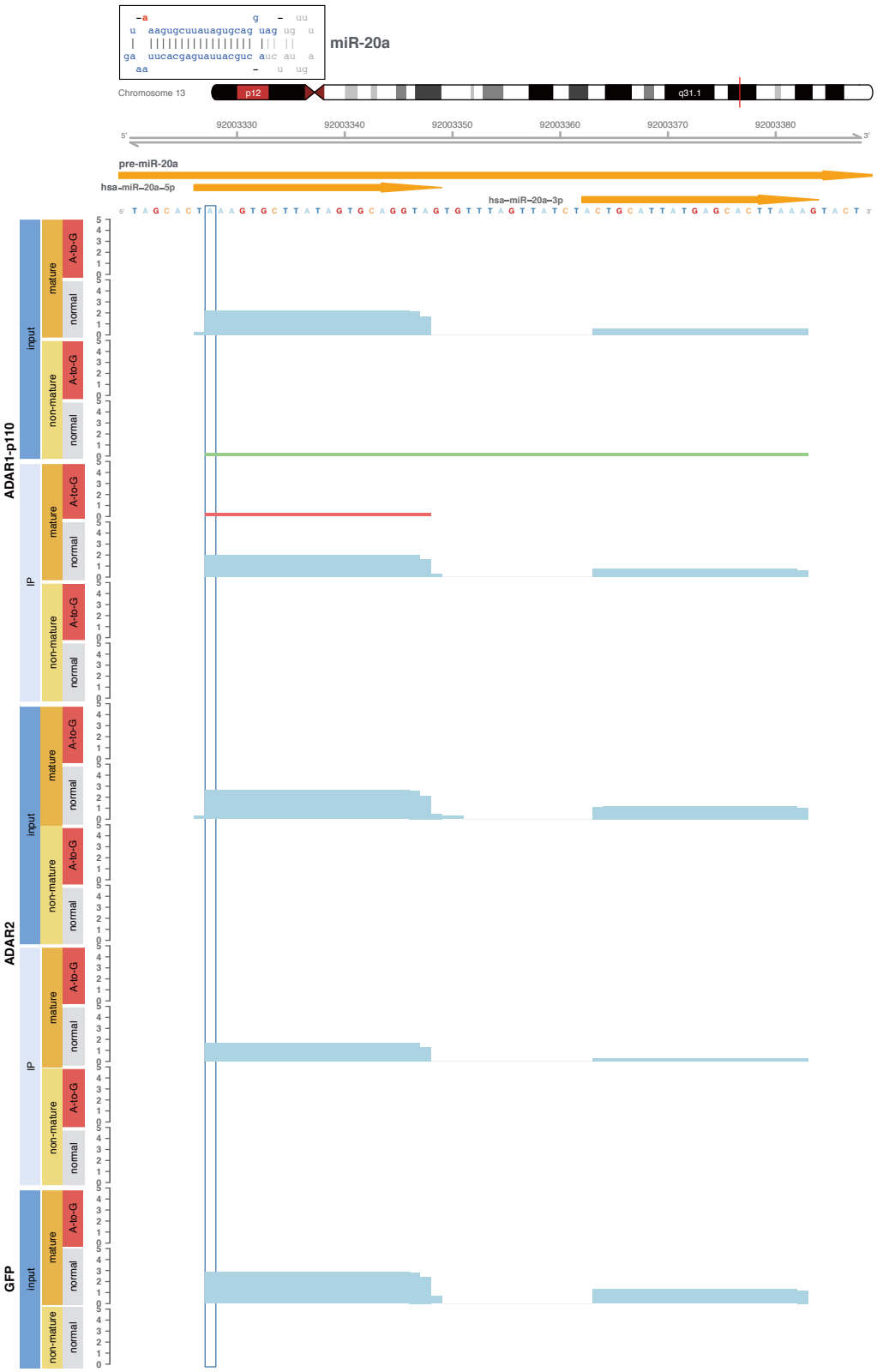
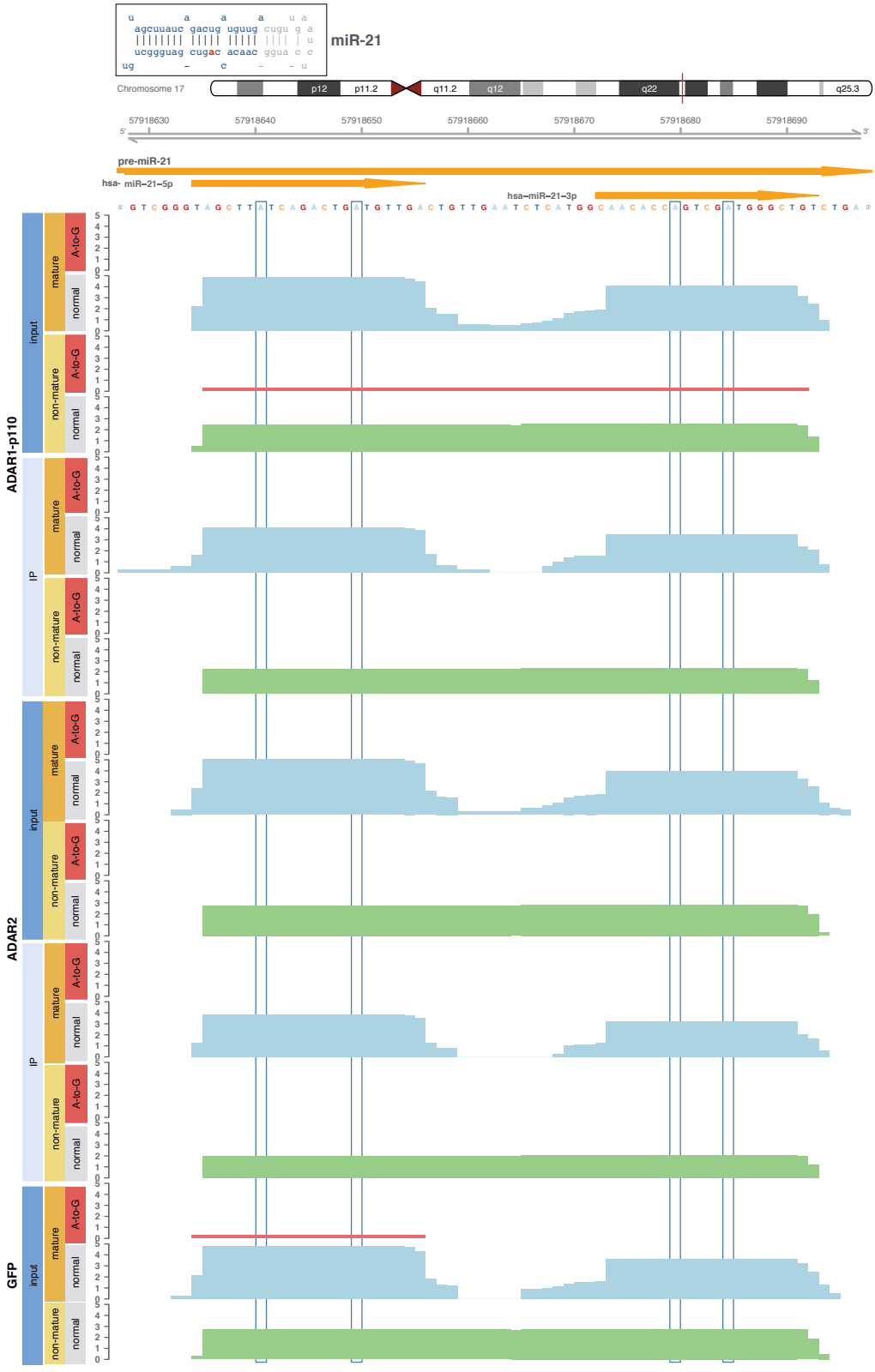


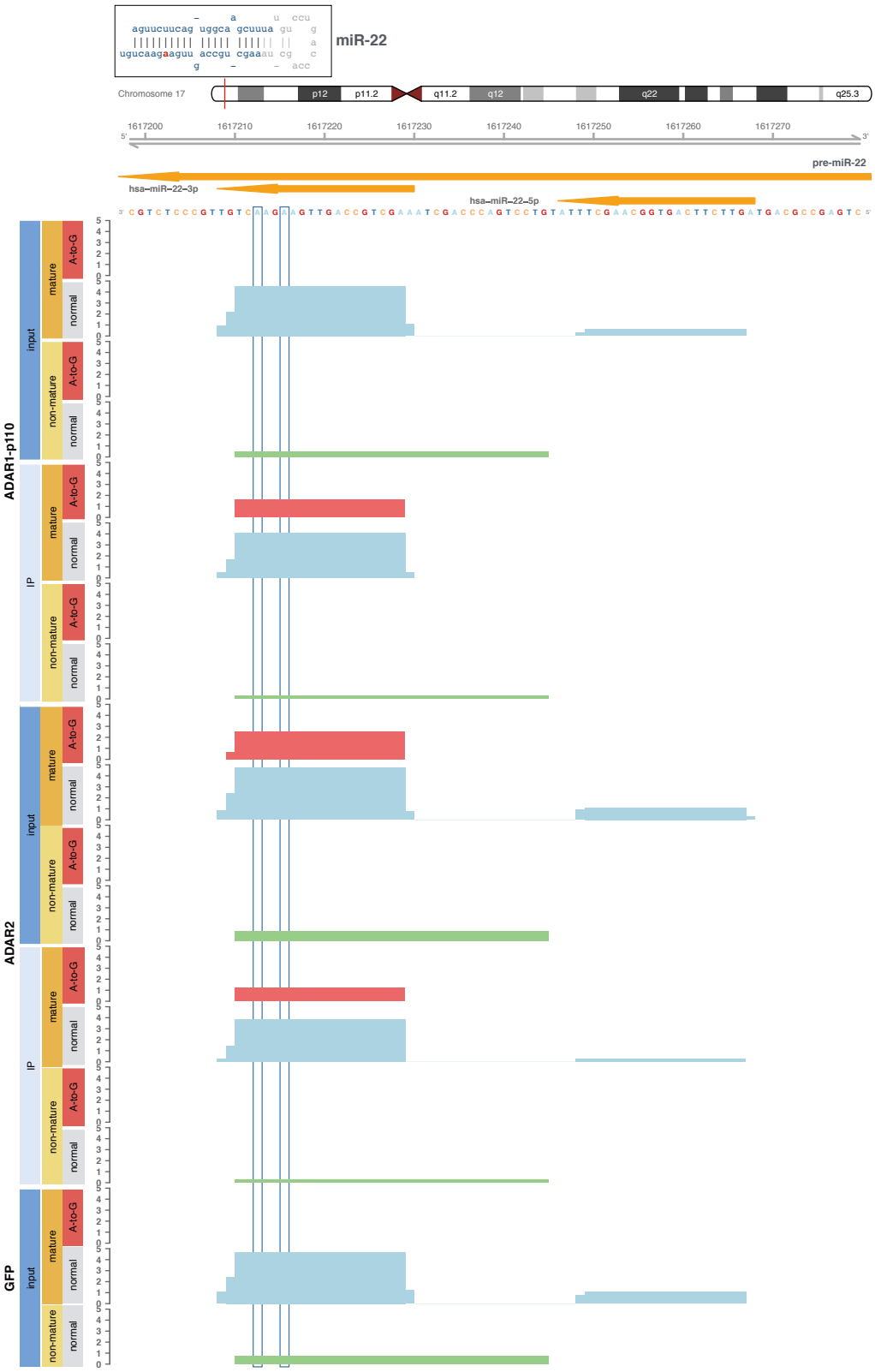
Figure S2. Genome browser view of all normal and collapse reads mapped to miRNA annotations. The editing positions detected in this study are shown by dark blue boxes for A-to-I edited miRNAs and pre-miRNAs. The secondary structures of the corresponding pre-miRNAs are shown in the upper left region as they appear in miRBase. The location of the miRNA on the chromosome is indicated in red. The genome coordinates are indicated on a gray axis under the chromosome. The mature and pre-miRNA annotations registered in miRBase are shown in orange, the arrows indicating direction of transcription. Mature miRNA reads with no mismatch are in light blue, non-mature miRNA reads with no mismatch in light green, and mature and non-mature miRNA reads with A-to-I edited reads in red. The y-axis of all tracks is in the log10 scale.

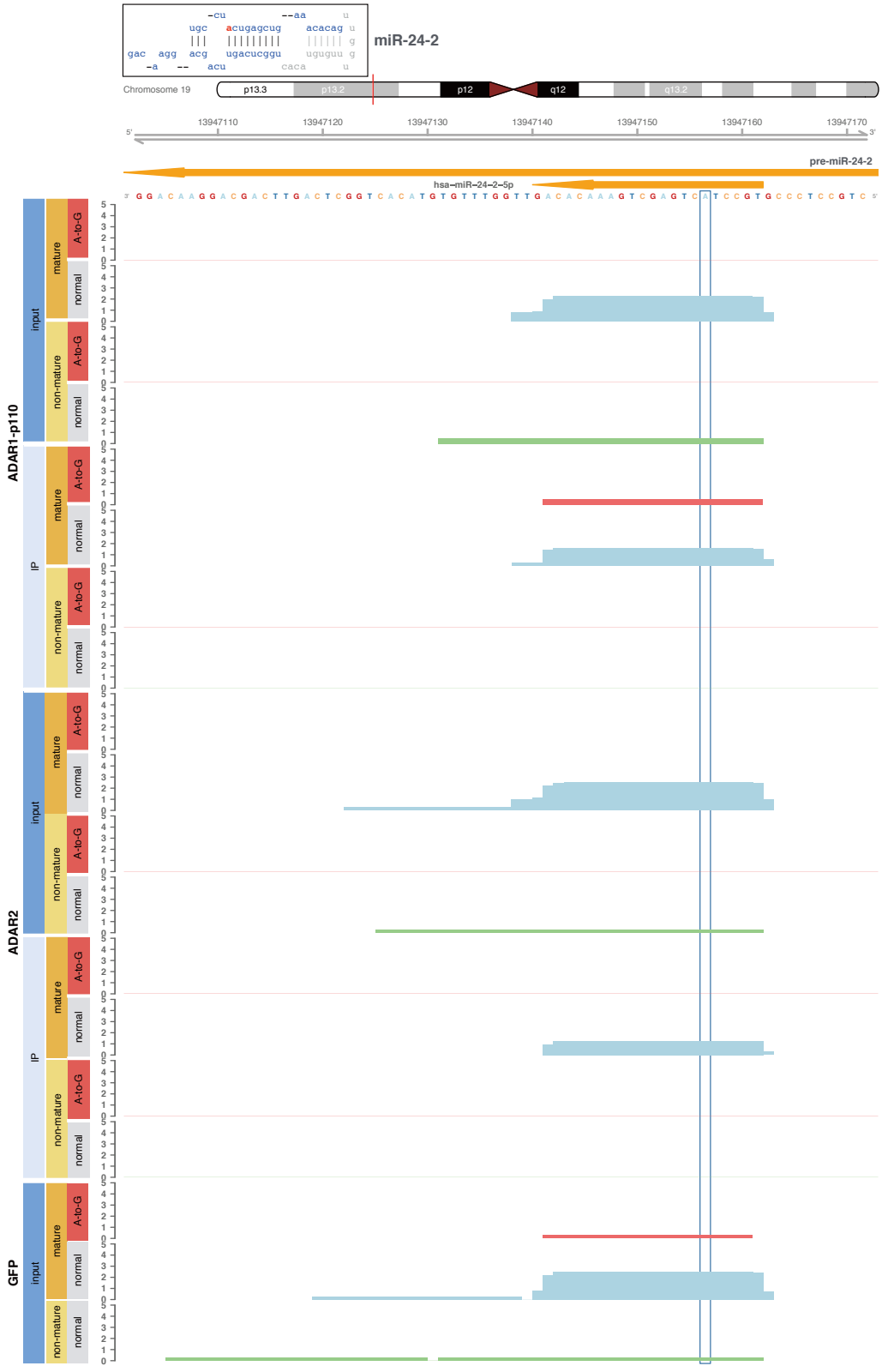

hsa-let-7g
Chromosome 3
52302300
52302310
52302320
52302330
52302340
52302350
52302360
52302370
hsa-let-7g-3p
hsa-let-7g-5p
pre-let-7g
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal


miR-19a
Chromosome 13
p12
q31.1
92003150
92003160
92003170
92003180
92003190
92003200
92003210
92003220
pre-miR-19a
hsa–miR–19a–5p
hsa–miR–19a–3p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

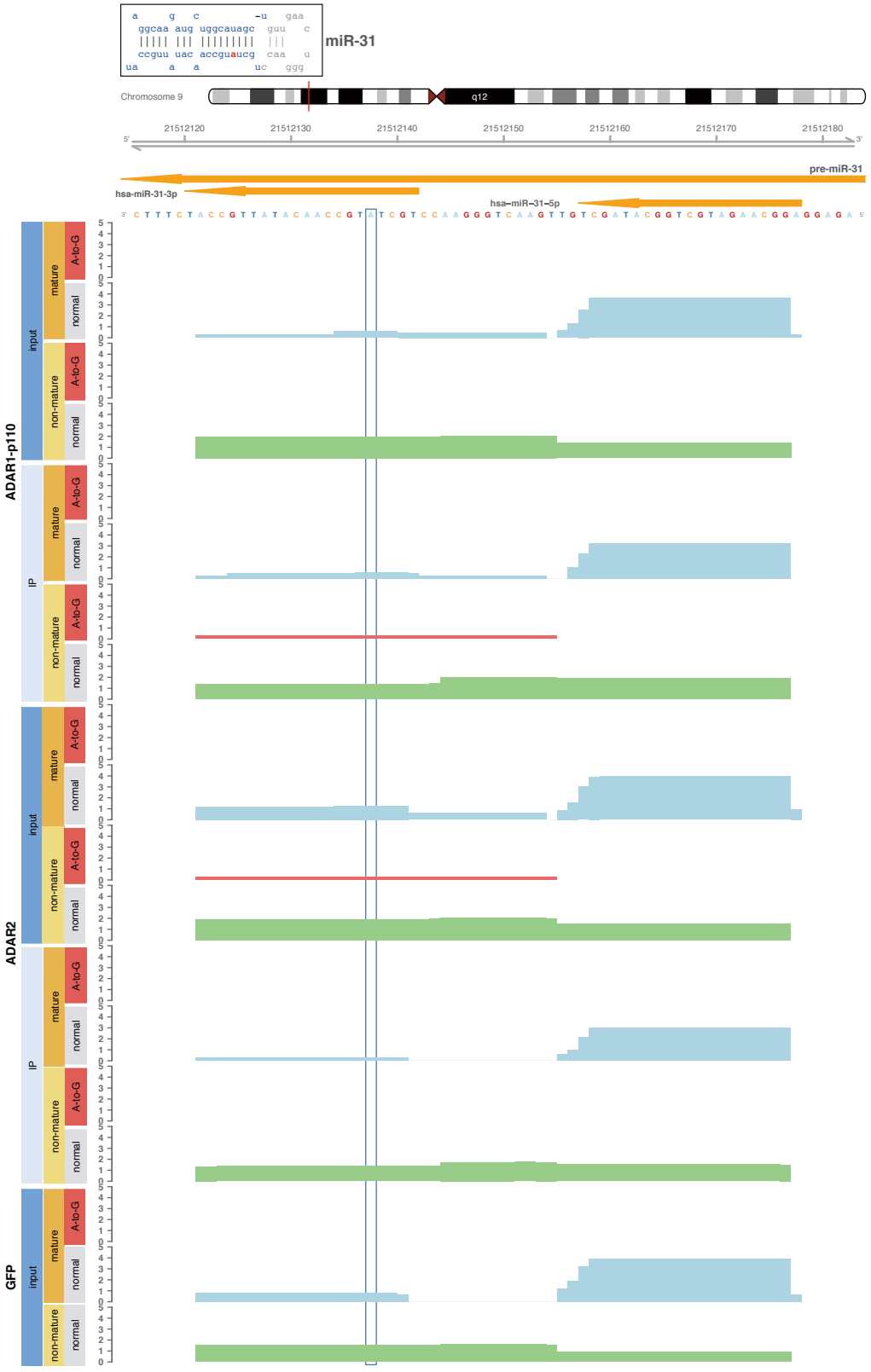

miR-20a
Chromosome 13
p12
q31.1
92003330
92003340
92003350
92003360
92003370
92003380
pre-miR-20a
hsa-miR-20a-5p
hsa-miR-20a-3p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal


miR-21
Chromosome 17
p12
p11.2
q11.2
q12
q22
q25.3
57918630
57918640
57918650
57918660
57918670
57918680
57918690
pre-miR-21
hsa-miR-21-5p
hsa-miR-21-3p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

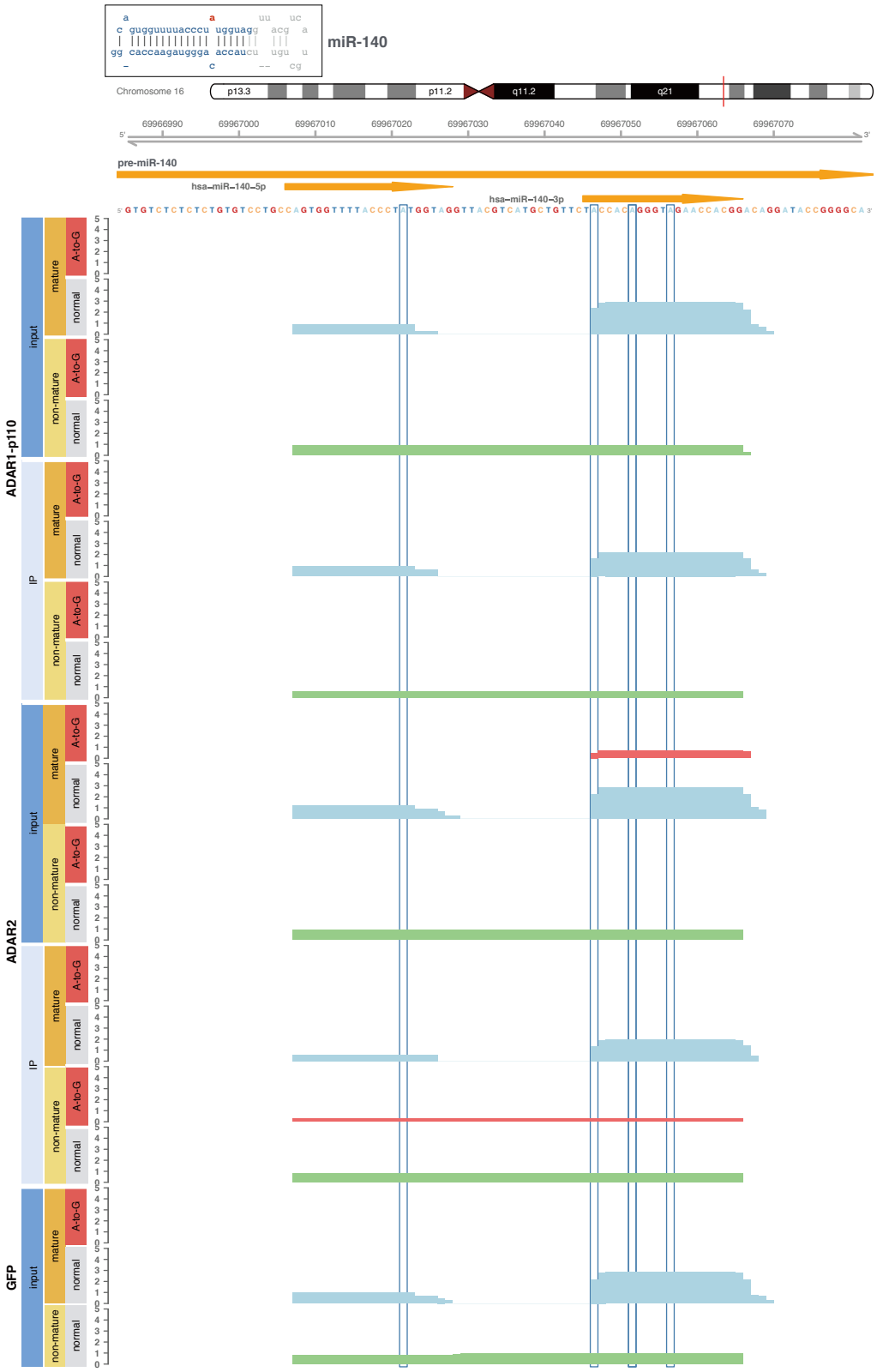

miR-22
Chromosome 17
p12
p11.2
q11.2
q12
q22
q25.3
1617200
1617210
1617220
1617230
1617240
1617250
1617260
1617270
pre-miR-22
hsa–miR–22–3p
hsa–miR–22–5p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

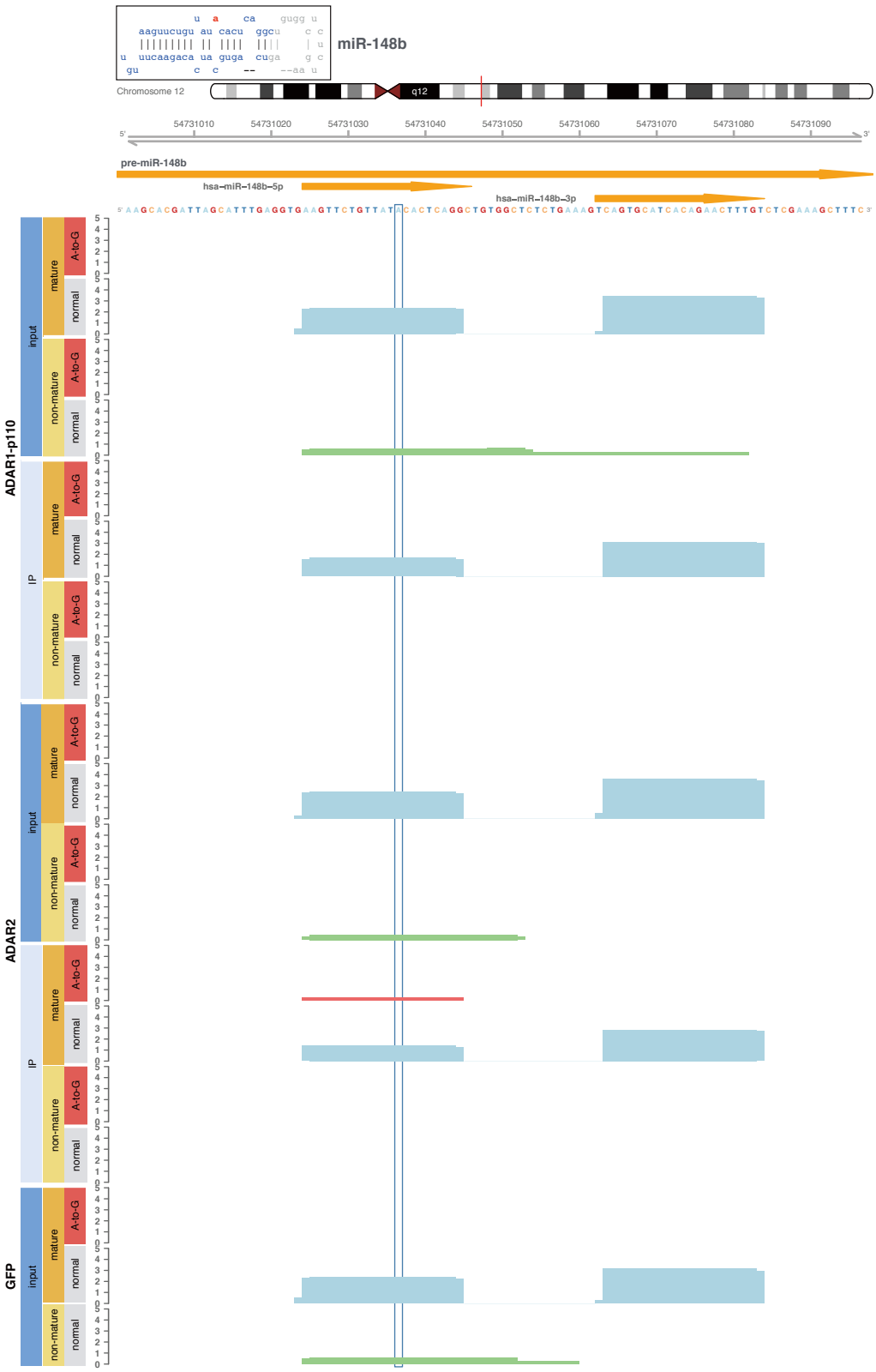

miR-24-2
Chromosome 19
p13.3
p13.2
p12
q12
q13.2
13947110
13947120
13947130
13947140
13947150
13947160
13947170
pre-miR-24-2
hsa-miR-24-2-5p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

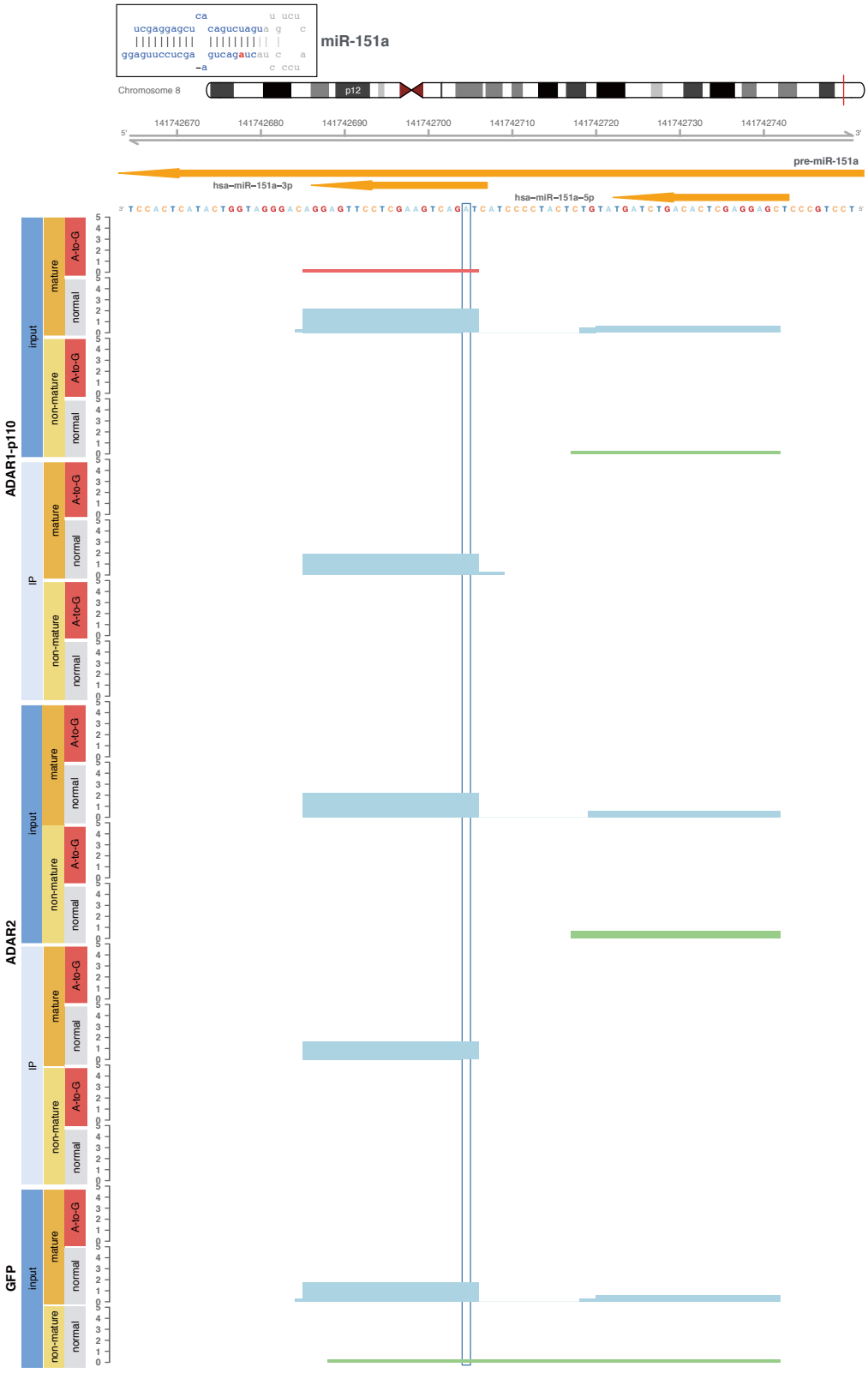

miR-27a
Chromosome 19
p13.3
p13.2
p12
q12
q13.2
13947260
13947270
13947280
13947290
13947300
13947310
13947320
pre-miR-27a
hsa-miR-27a-3p
hsa-miR-27a-5p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

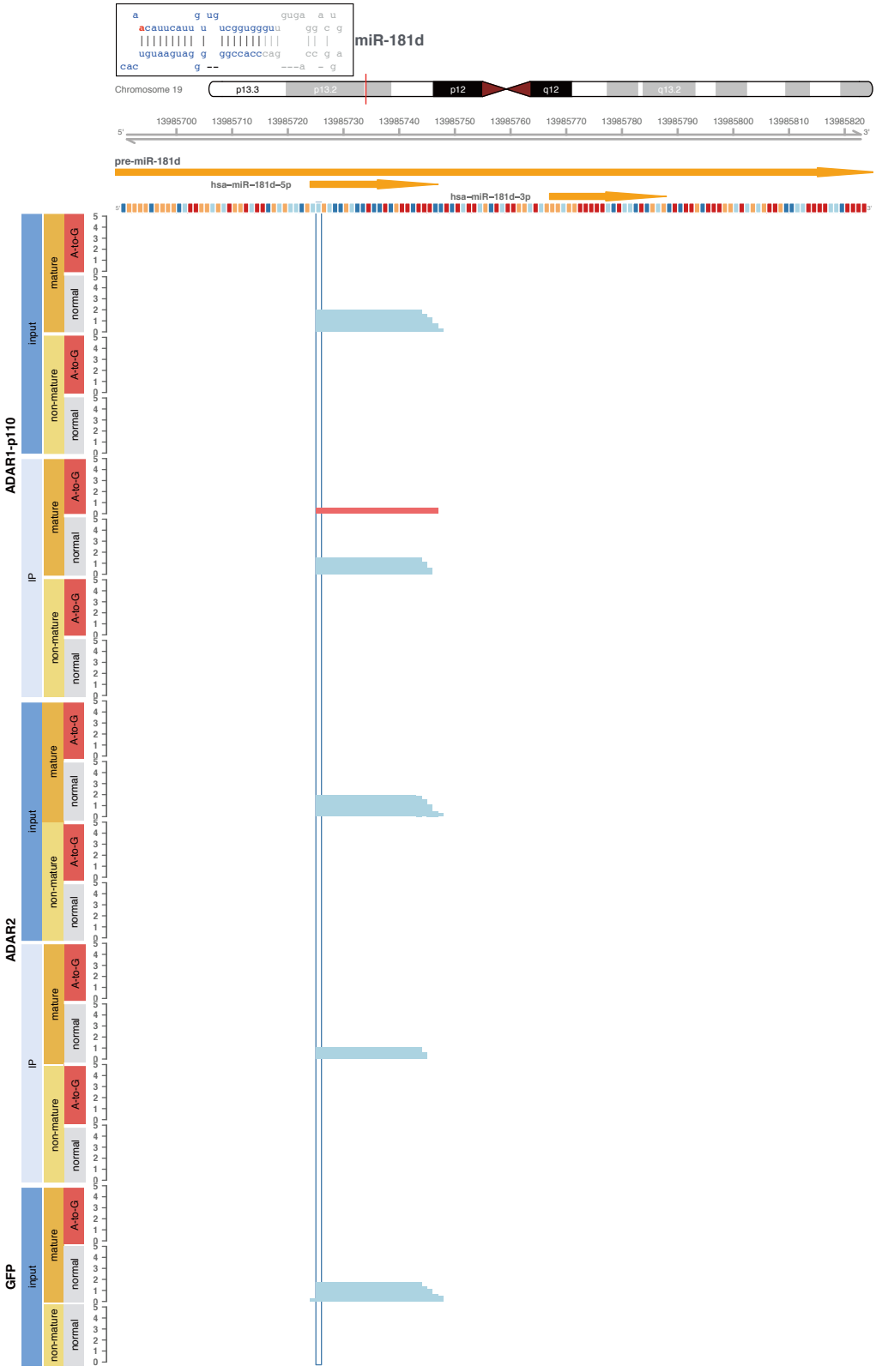

miR-31
Chromosome 9
q12
21512120
21512130
21512140
21512150
21512160
21512170
21512180
pre-miR-31
hsa-miR-31-3p
hsa-miR-31-5p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

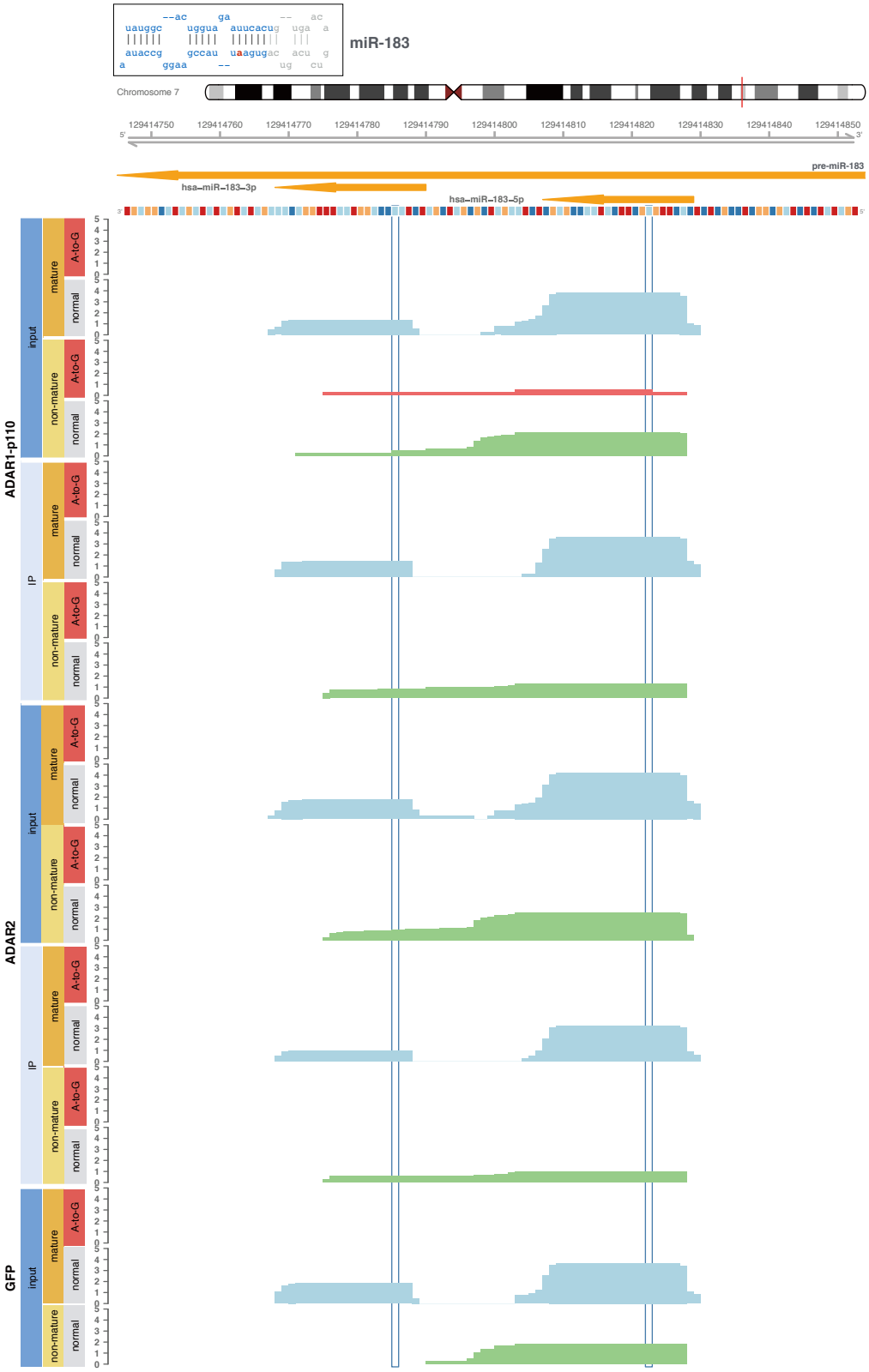

miR-99a
Chromosome 21
p13
p12
p11.2
q21.1
q21.3
q22.11
q22.3
17911420
17911430
17911440
17911450
17911460
17911470
17911480
pre-miR-99a
hsa-miR-99a-5p
hsa-miR-99a-3p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

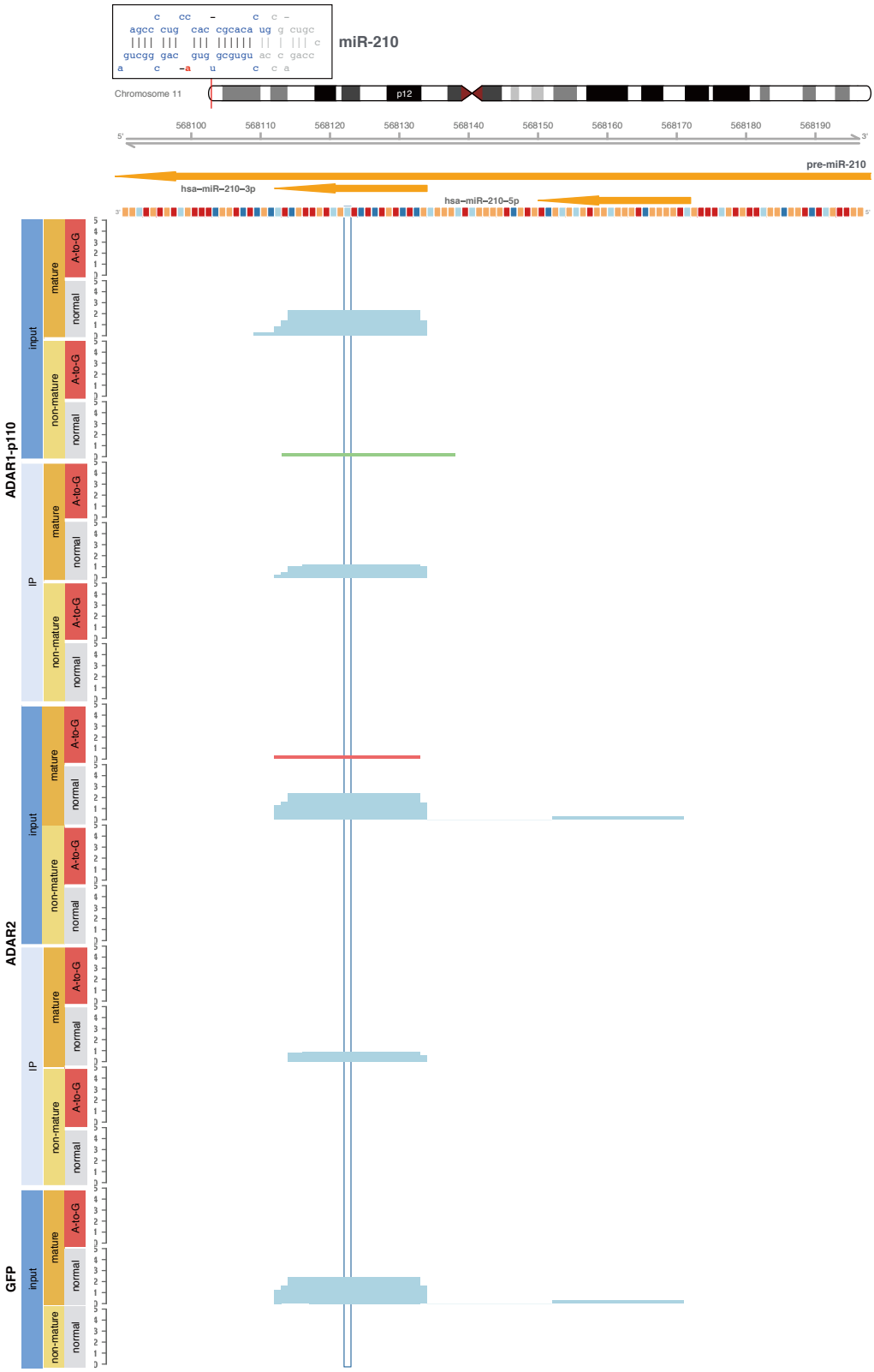

miR-140
Chromosome 16
p13.3
p11.2
q11.2
q21
69966990
69967000
69967010
69967020
69967030
69967040
69967050
69967060
69967070
pre-miR-140
hsa-miR-140-5p
hsa-miR-140-3p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

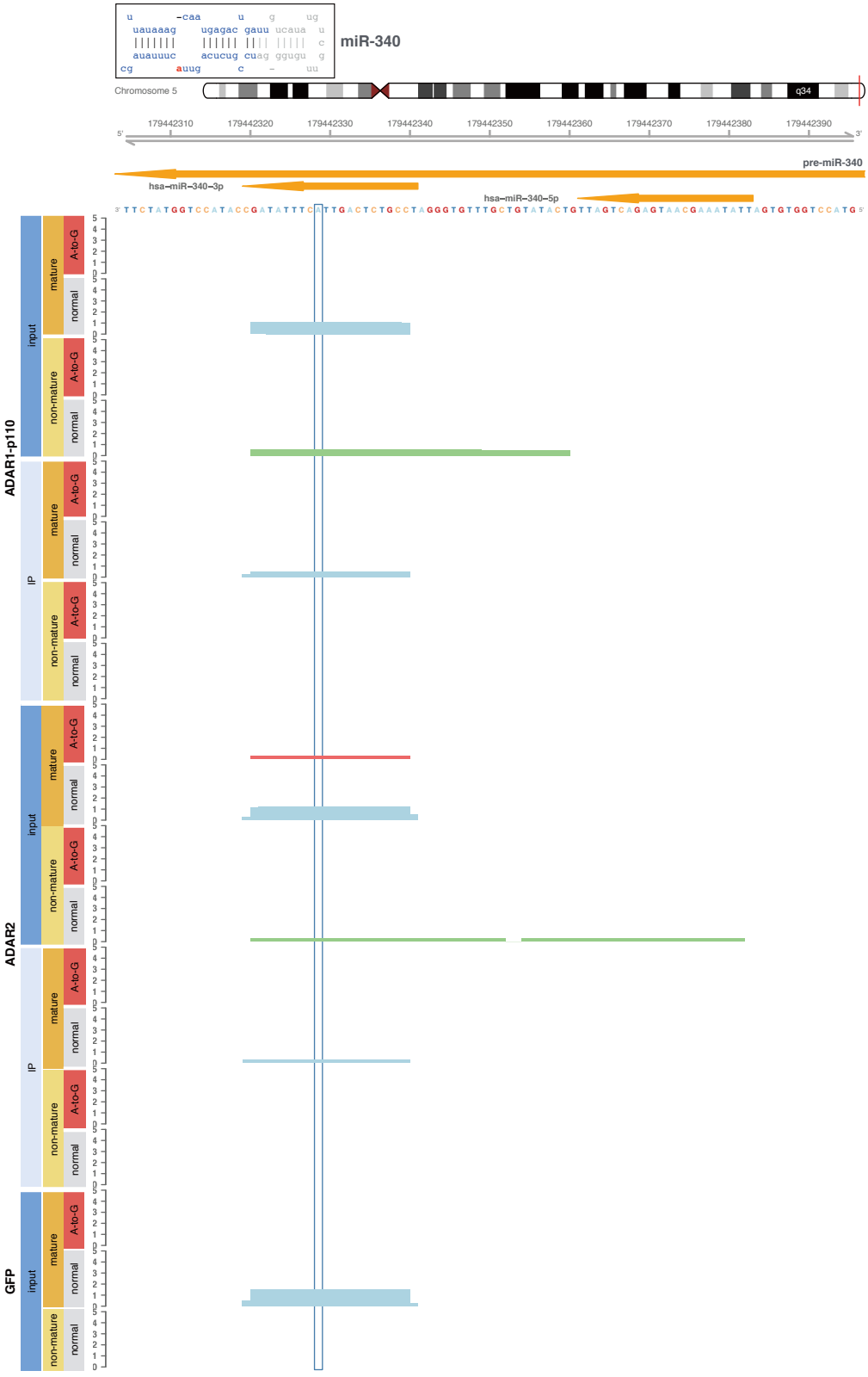

miR-148b
Chromosome 12
q12
54731010
54731020
54731030
54731040
54731050
54731060
54731070
54731080
54731090
pre-miR-148b
hsa-miR-148b-5p
hsa-miR-148b-3p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

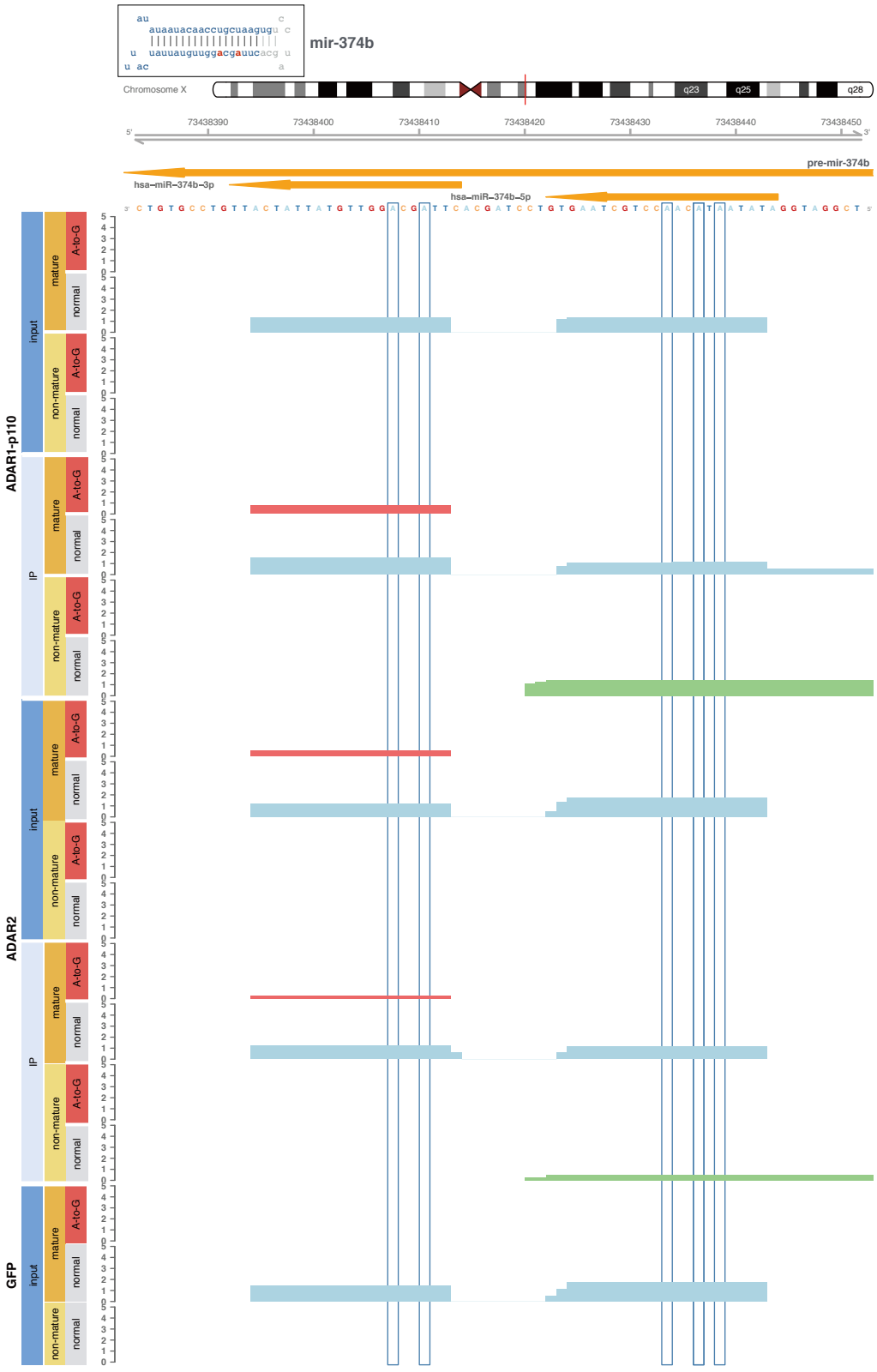

miR-151a
Chromosome 8
p12
141742670
141742680
141742690
141742700
141742710
141742720
141742730
141742740
pre-miR-151a
hsa-miR-151a-3p
hsa-miR-151a-5p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

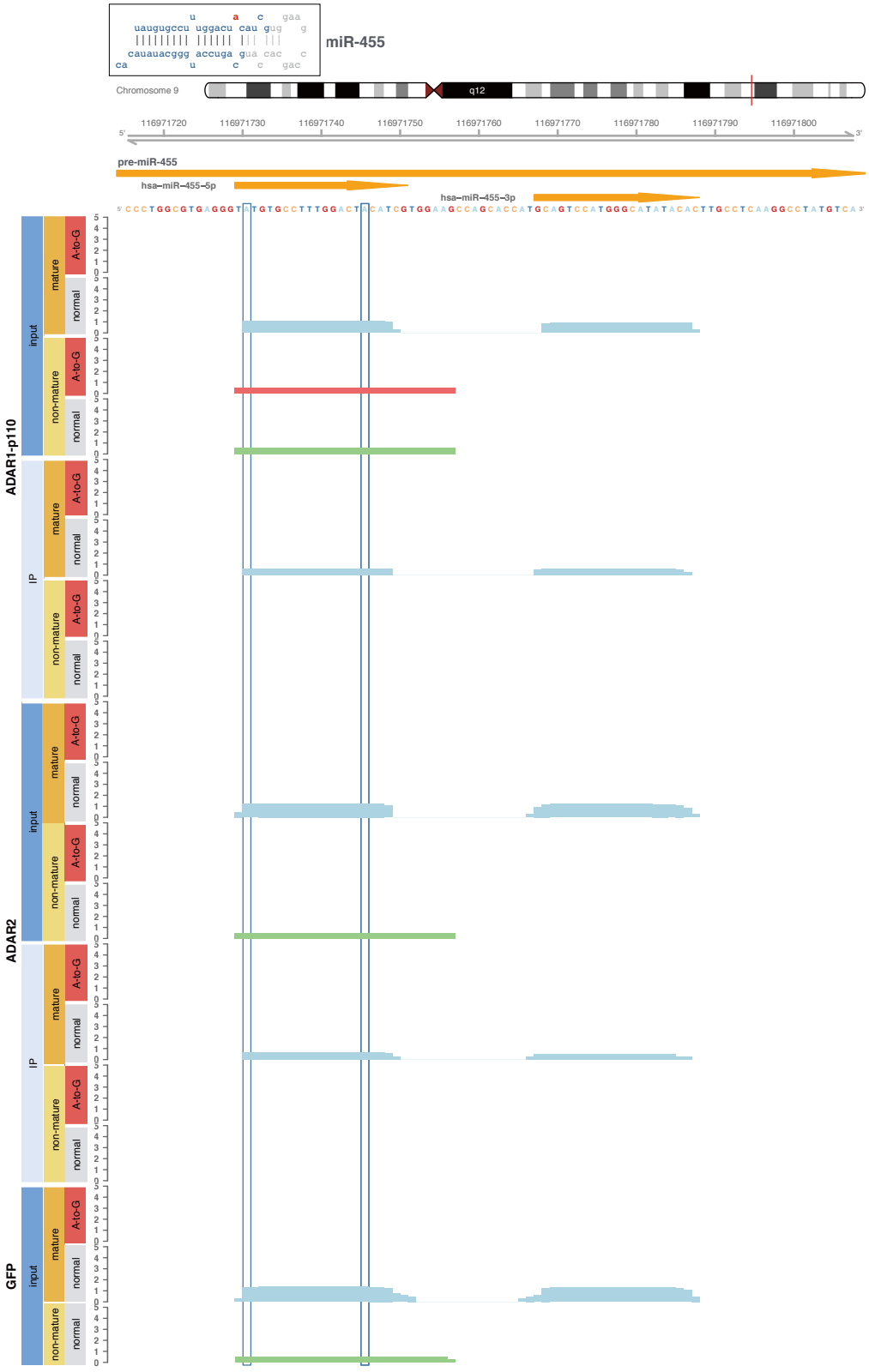

miR-181d
Chromosome 19
p13.3
p13.2
p12
q12
q13.2
13985700
13985710
13985720
13985730
13985740
13985750
13985760
13985770
13985780
13985790
13985800
13985810
13985820
pre-miR-181d
hsa-miR-181d-5p
hsa-miR-181d-3p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

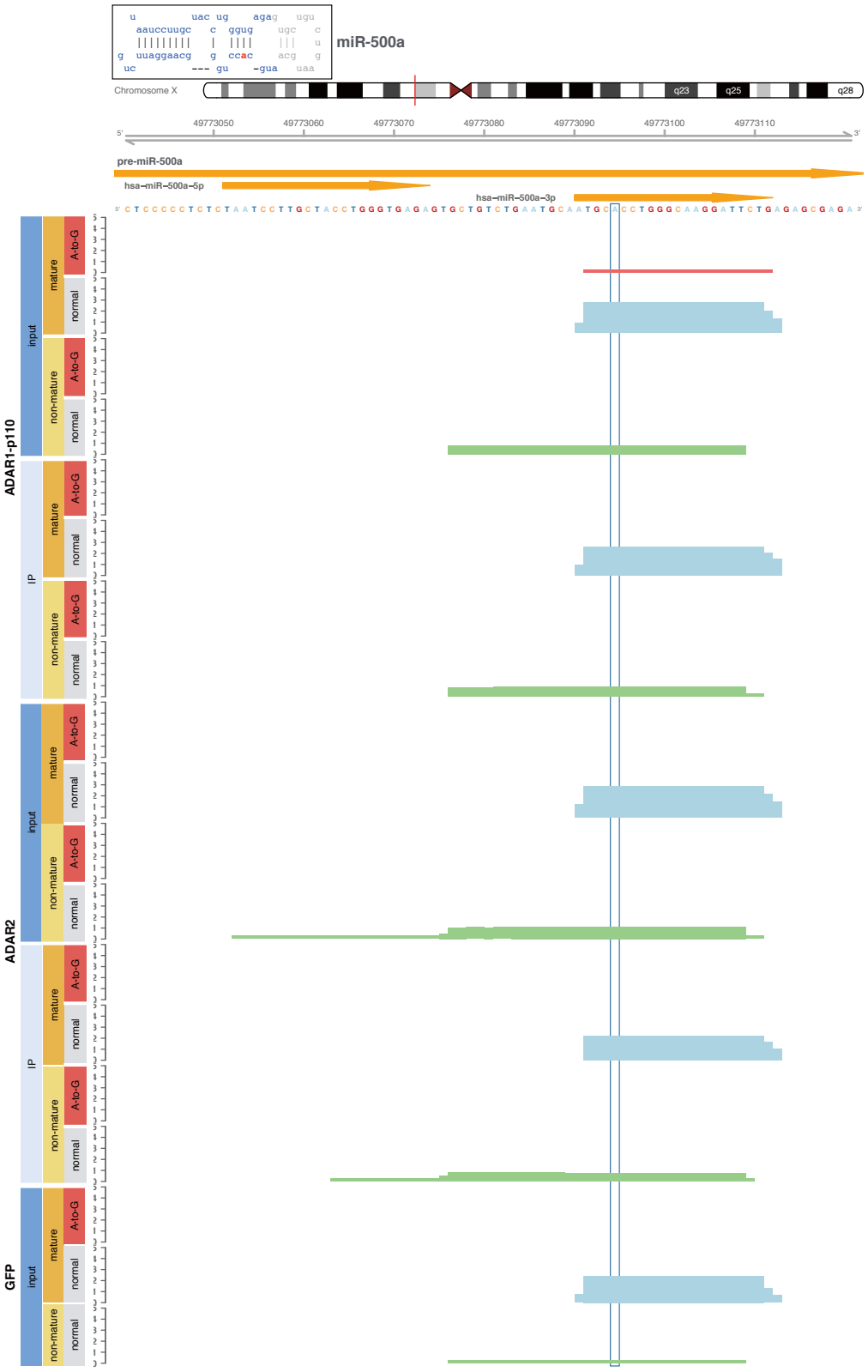

miR-183
Chromosome 7
129414750
129414760
129414770
129414780
129414790
129414800
129414810
129414820
129414830
129414840
129414850
pre-miR-183
hsa-miR-183-3p
hsa-miR-183-5p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

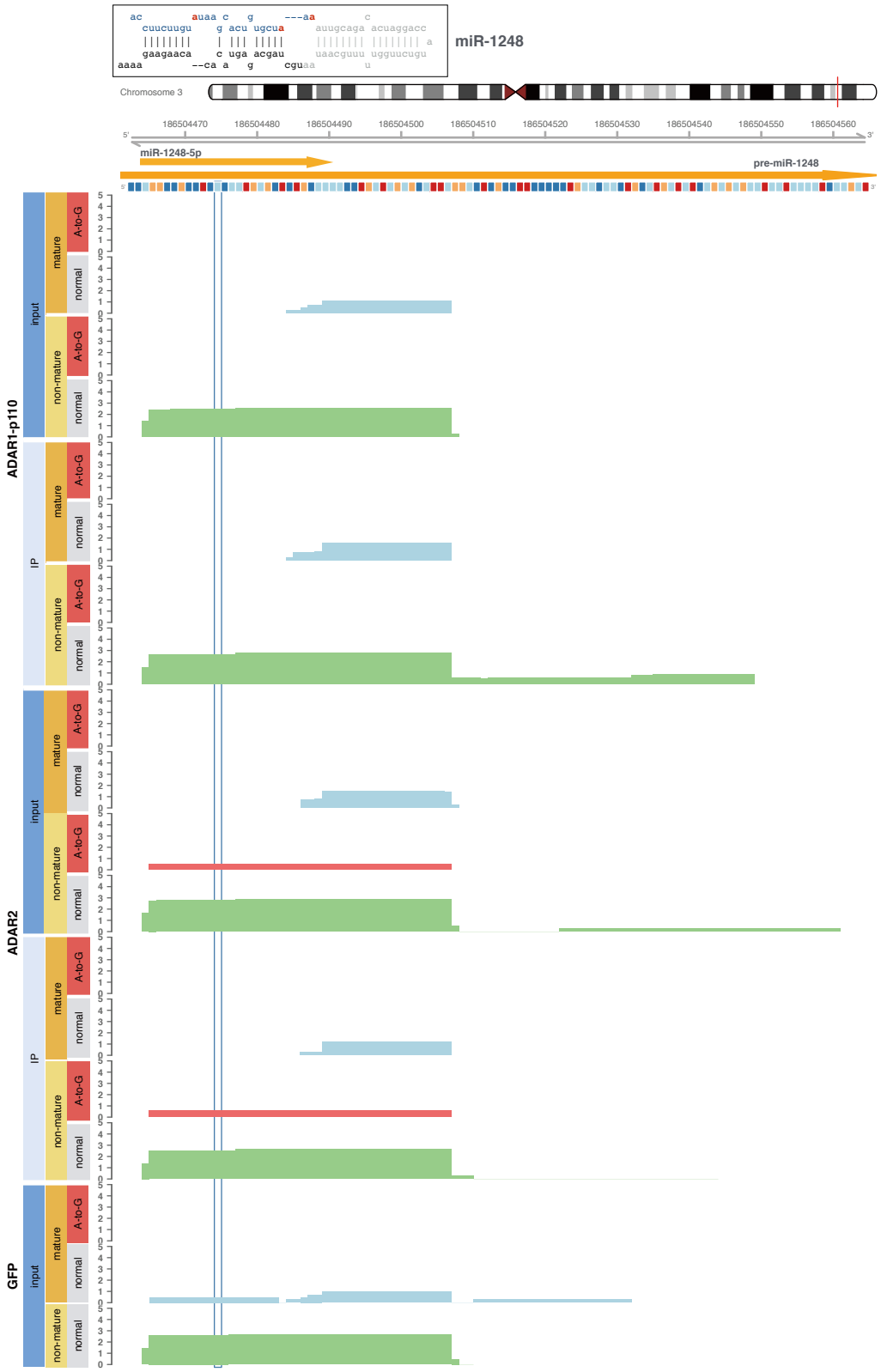

miR-210
Chromosome 11
p12
568100
568110
568120
568130
568140
568150
568160
568170
568180
568190
pre-miR-210
hsa–miR–210–3p
hsa–miR–210–5p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

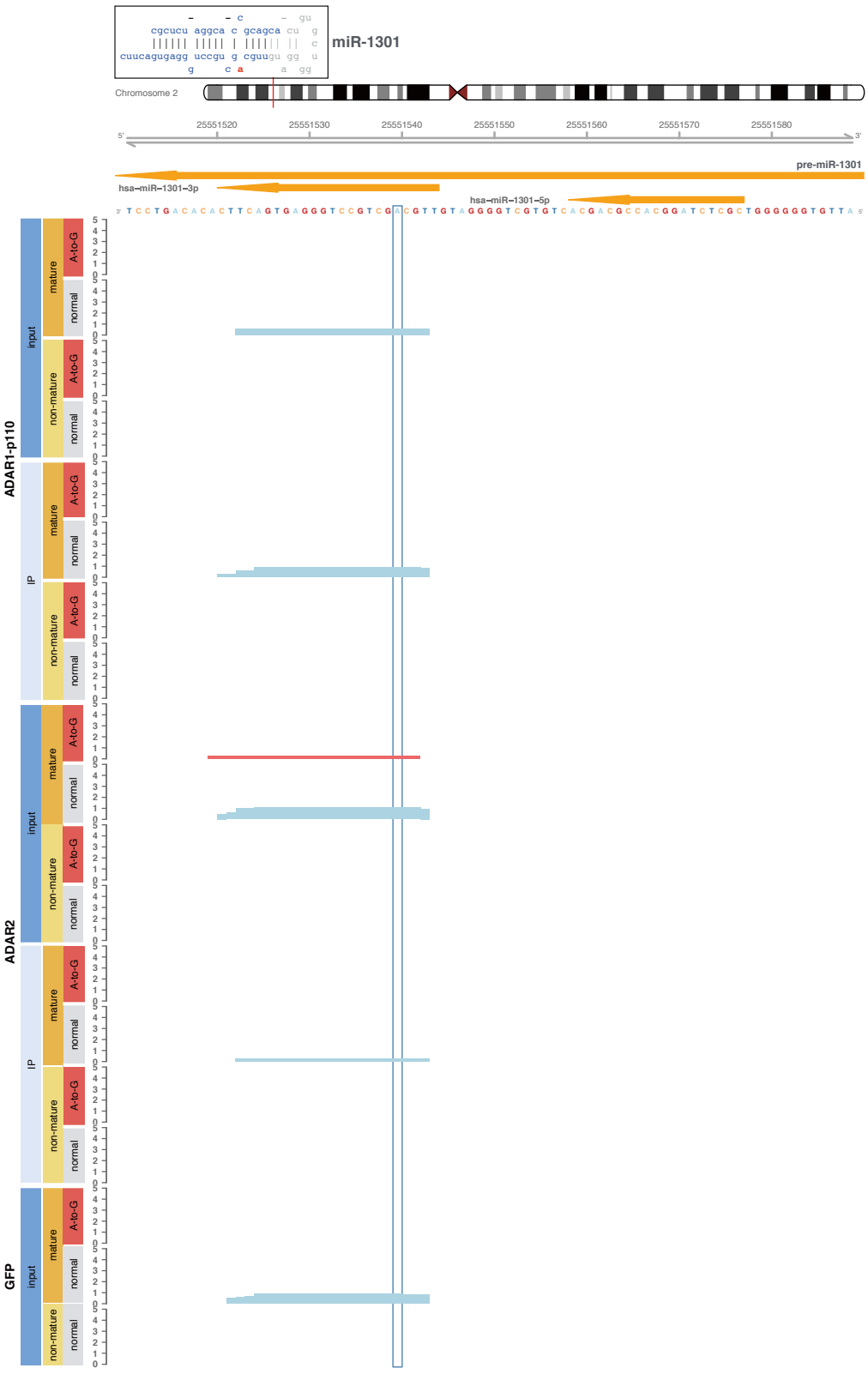

miR-340
Chromosome 5
q34
179442310
179442320
179442330
179442340
179442350
179442360
179442370
179442380
179442390
pre-miR-340
hsa-miR-340-3p
hsa-miR-340-5p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
normal
A-to-G

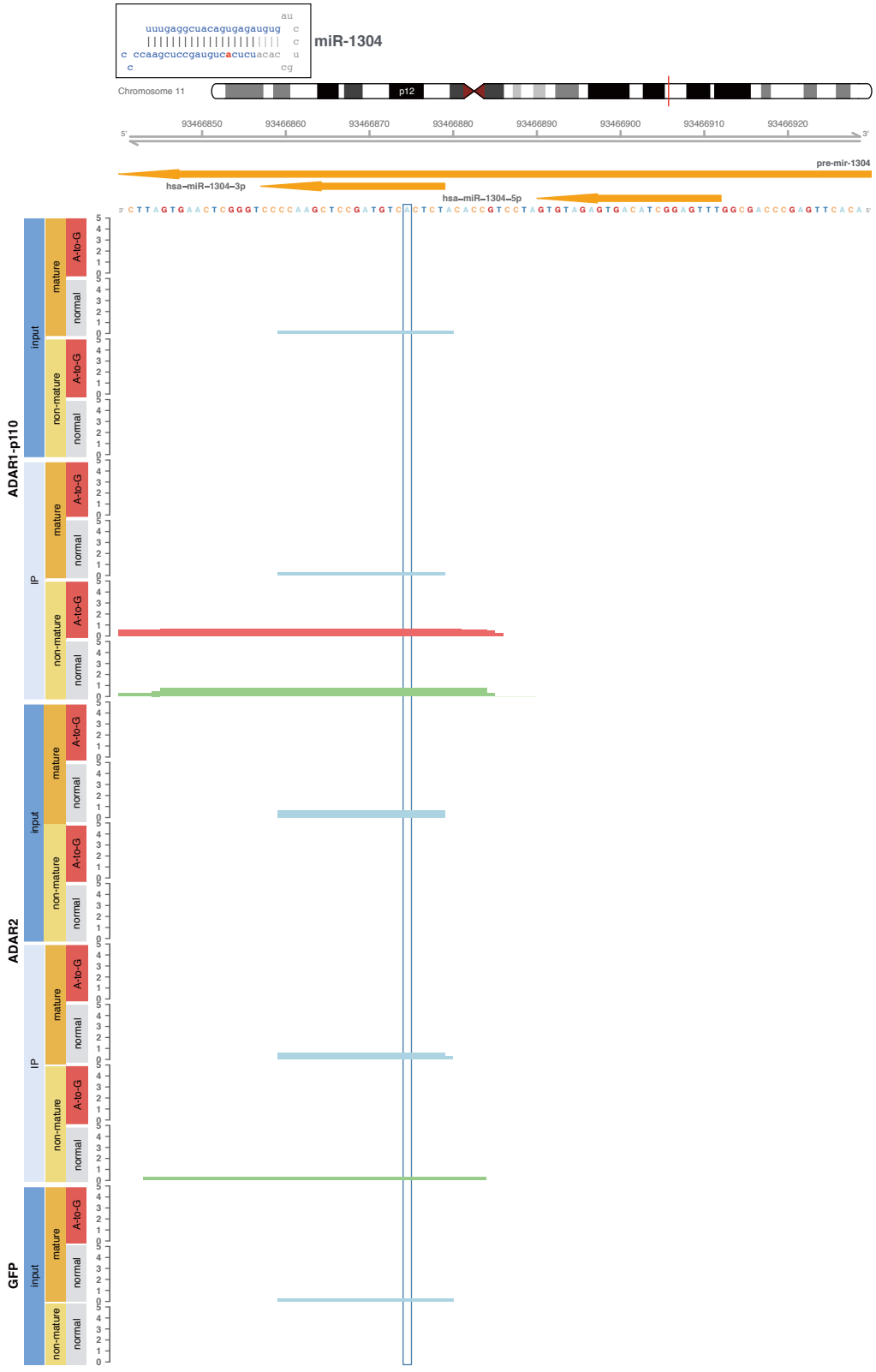

mir-374b
Chromosome X
q23
q25
q28
73438390
73438400
73438410
73438420
73438430
73438440
73438450
pre-mir-374b
hsa-miR-374b-3p
hsa-miR-374b-5p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

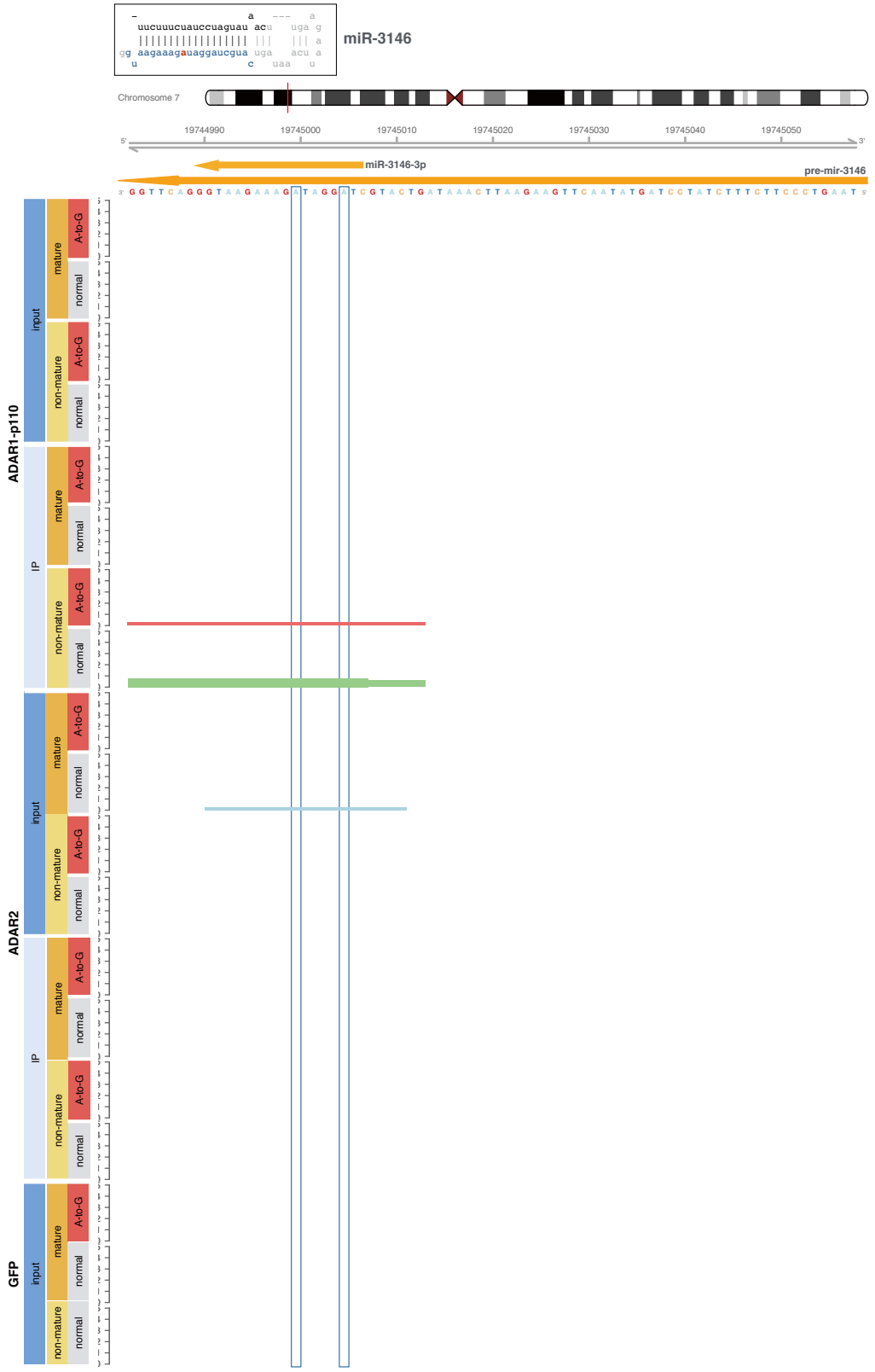

miR-455
Chromosome 9
q12
116971720
116971730
116971740
116971750
116971760
116971770
116971780
116971790
116971800
pre-miR-455
hsa-miR-455-5p
hsa-miR-455-3p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

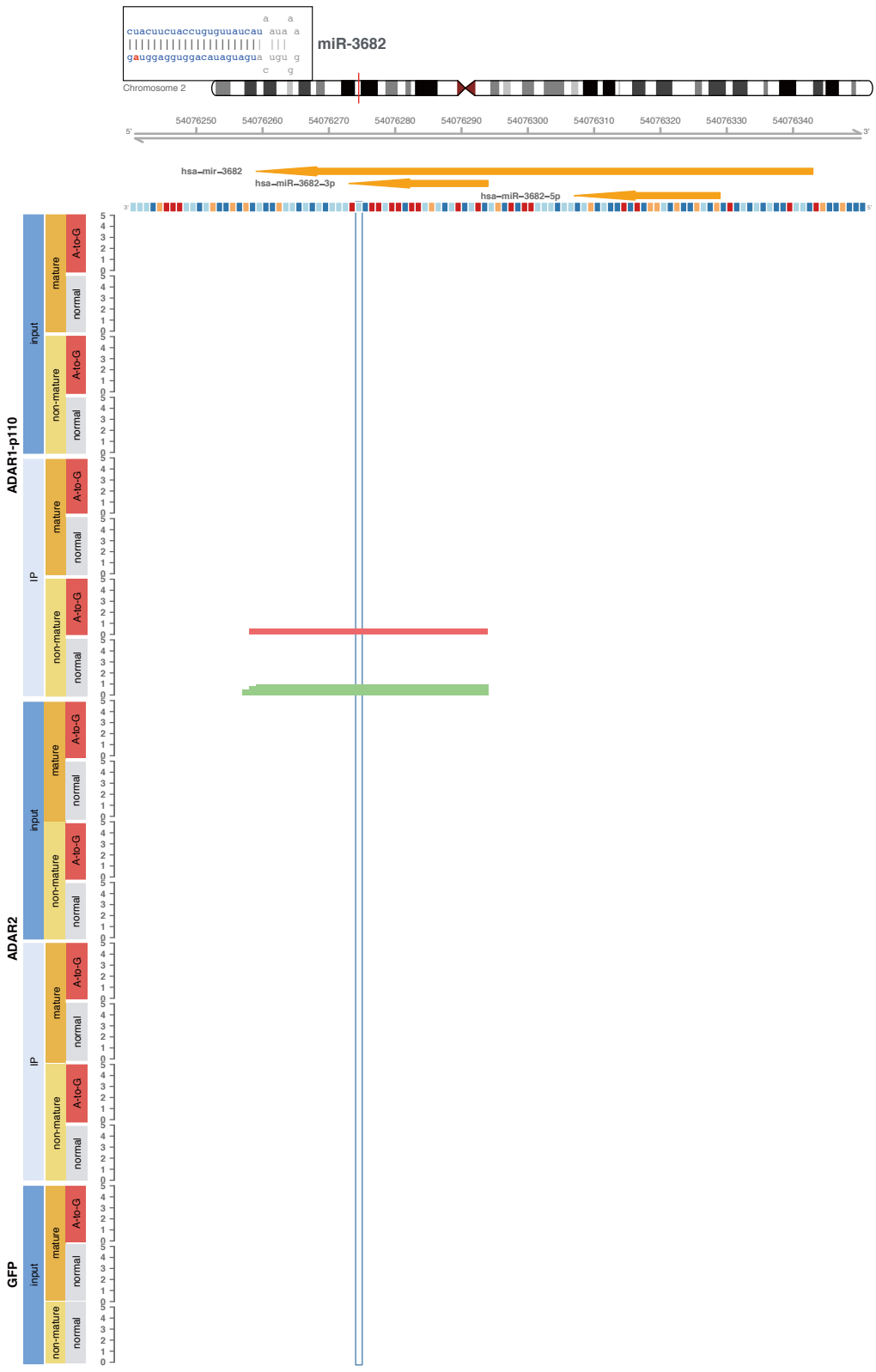

miR-500a
Chromosome X
q23
q25
q28
49773050
49773060
49773070
49773080
49773090
49773100
49773110
pre-miR-500a
hsa-miR-500a-5p
hsa-miR-500a-3p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

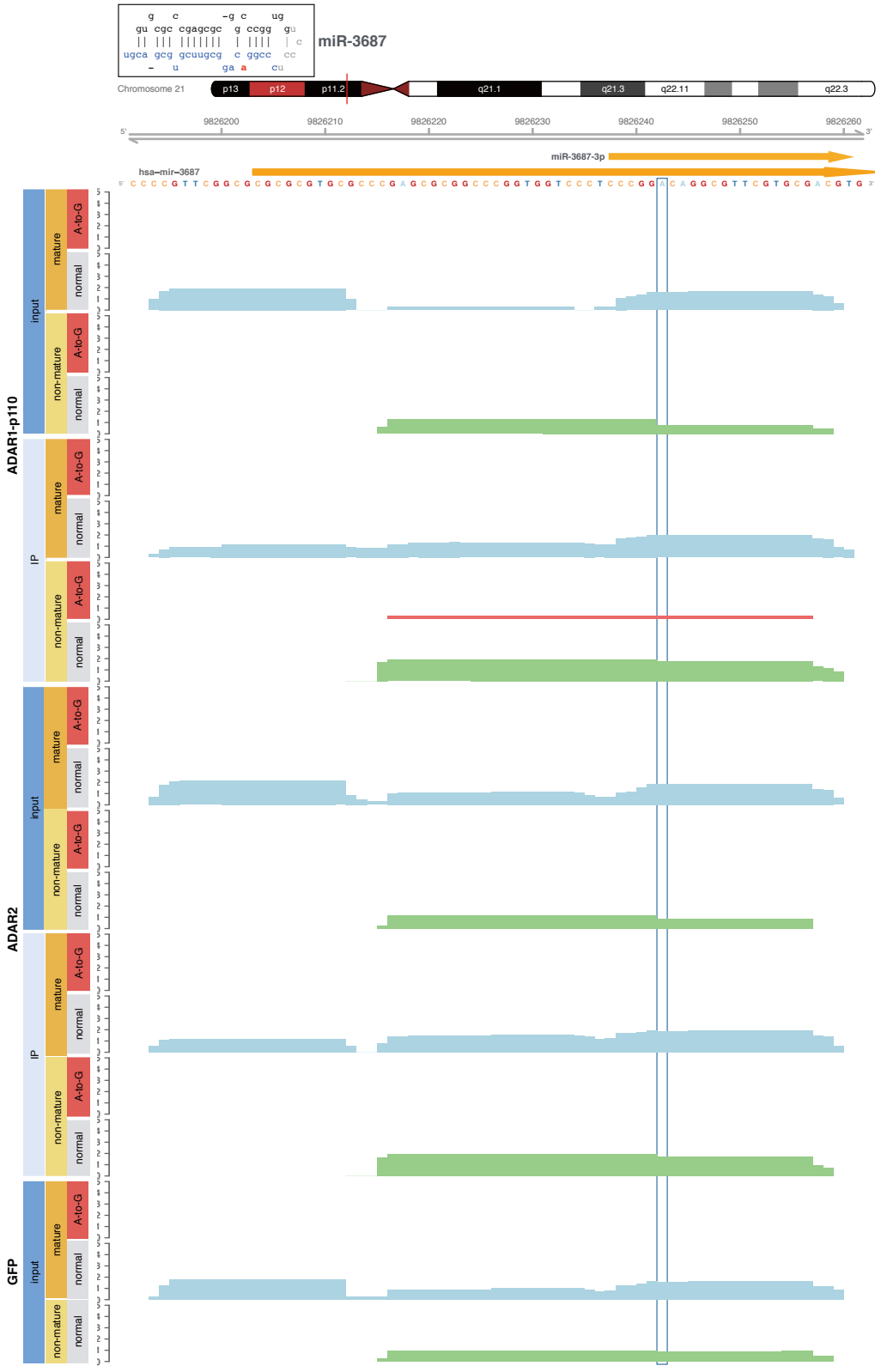

miR-1248
Chromosome 3
186504470
186504480
186504490
186504500
186504510
186504520
186504530
186504540
186504550
186504560
miR-1248-5p
pre-miR-1248
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

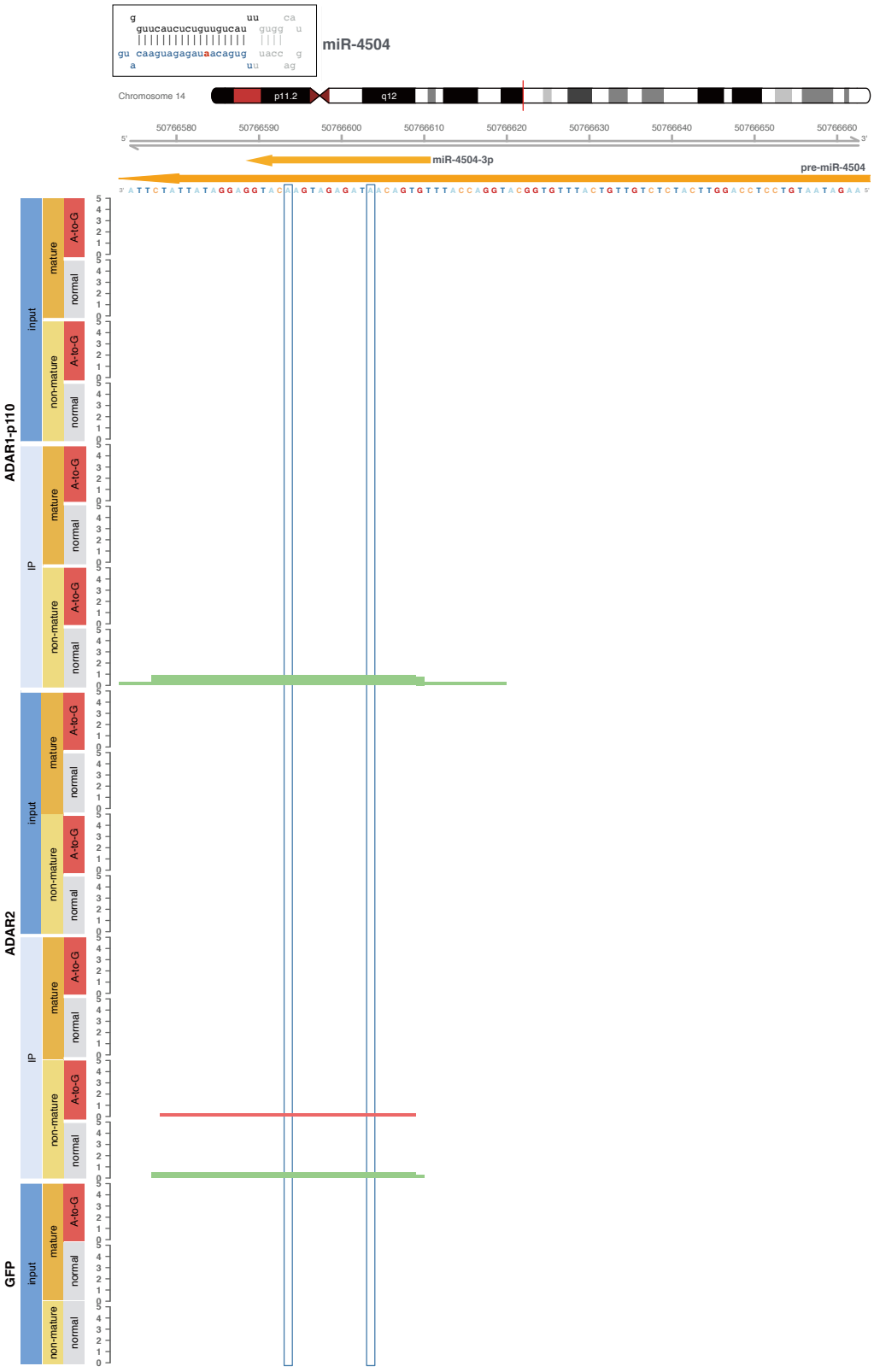

miR-1301
Chromosome 2
25551520
25551530
25551540
25551550
25551560
25551570
25551580
pre-miR-1301
hsa-miR-1301-3p
hsa-miR-1301-5p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
normal
A-to-G


miR-1304
Chromosome 11
p12
93466850
93466860
93466870
93466880
93466890
93466900
93466910
93466920
pre-mir-1304
hsa-miR-1304-3p
hsa-miR-1304-5p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal


miR-3146
Chromosome 7
19744990
19745000
19745010
19745020
19745030
19745040
19745050
miR-3146-3p
pre-mir-3146
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal


miR-3682
Chromosome 2
54076250
54076260
54076270
54076280
54076290
54076300
54076310
54076320
54076330
54076340
hsa-mir-3682
hsa-miR-3682-3p
hsa-miR-3682-5p
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
normal
A-to-G


miR-3687
Chromosome 21
p13
p12
p11.2
q21.1
q21.3
q22.11
q22.3
9826200
9826210
9826220
9826230
9826240
9826250
9826260
miR-3687-3p
hsa-mir-3687
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal


miR-4504
Chromosome 14
p11.2
q12
50766580
50766590
50766600
50766610
50766620
50766630
50766640
50766650
50766660
miR-4504-3p
pre-miR-4504
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal


mir-5583
Chromosome 18
q11.2
q12.1
q12.3
q21.2
q23
37256690
37256700
37256710
37256720
37256730
37256740
37256750
37256760
37256770
hsa-miR-5583-3p
hsa-miR-5583-5p
pre-mir-5583
ADAR1-p110
ADAR2
GFP
input
IP
mature
non-mature
A-to-G
normal

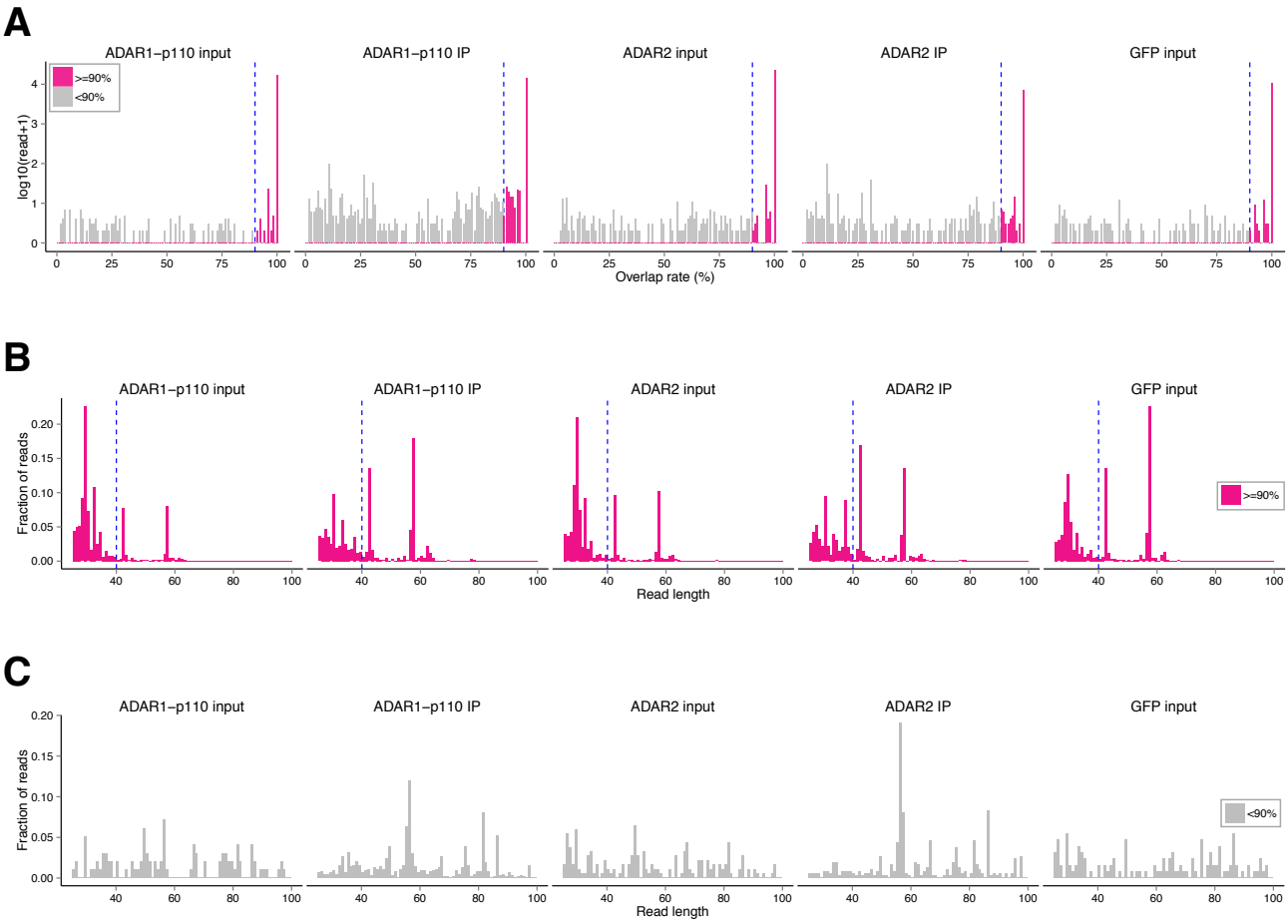


Figure S3. Length distribution of longer reads (≥ 25 nt) mapped to miRNA regions and their overlap with pre-miRNA annotation. (A) Percentage of overlap between each longer read (25 nt) and its pre-miRNA annotation in ADAR1-p110 input and IP, and ADAR2 input and IP samples, respectively. The horizontal axis indicates the percentage of overlap with the corresponding pre-miRNA annotation, and the vertical axis indicates log10 (raw read count+1). The dashed line indicates 90% overlap. The pink and gray lines indicate the reads overlapped with ≥ 90% and < 90% with pre-miRNA annotations, respectively. The length distribution of the reads overlapped ≥ 90% (B) and < 90% (C) with annotated miRNAs.

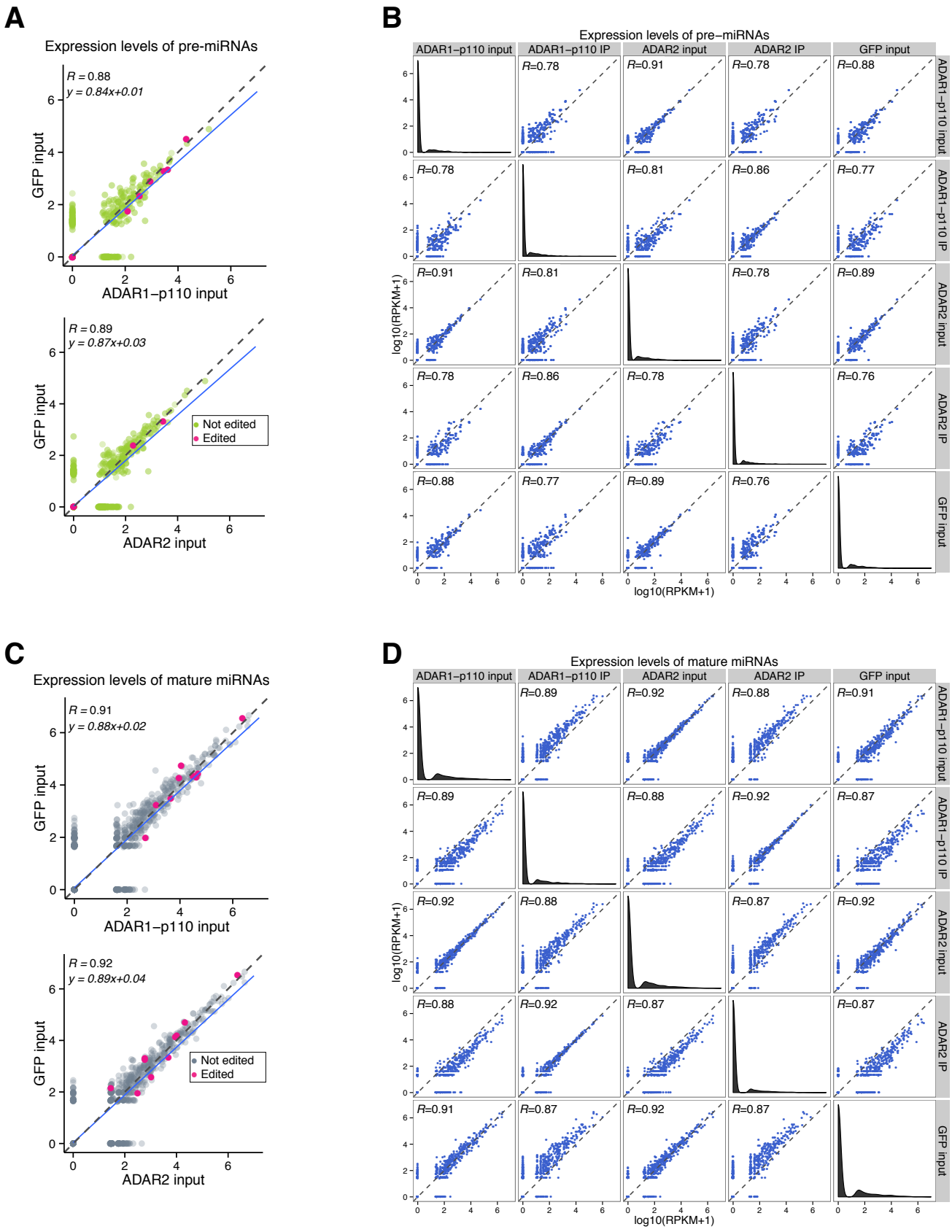


Figure S4. miRNA expression levels and binding preferences of each ADAR isoform to miRNAs.Correlation between the mean expression levels of pre-miRNAs (A) or mature miRNAs (C) in ADAR1-p110 (upper) or ADAR2 (lower) input sample and those of GFP control input. Round-robin comparison of expression levels of pre-miRNAs (B) and mature miRNAs (D) across all samples sequenced. X-axis and Y-axis are log10(RPKM+1). R indicates the Pearson correlation coefficient and y represents the regression line. The dashed gray line corresponds to y=x. Pink dots in (A, C): edited pre-miRNAs (A) and mature miRNAs (C). The histogram filled in black in (B, D) in the diagonally aligned panels shows the distribution of expression levels (RPKM) in each sample.

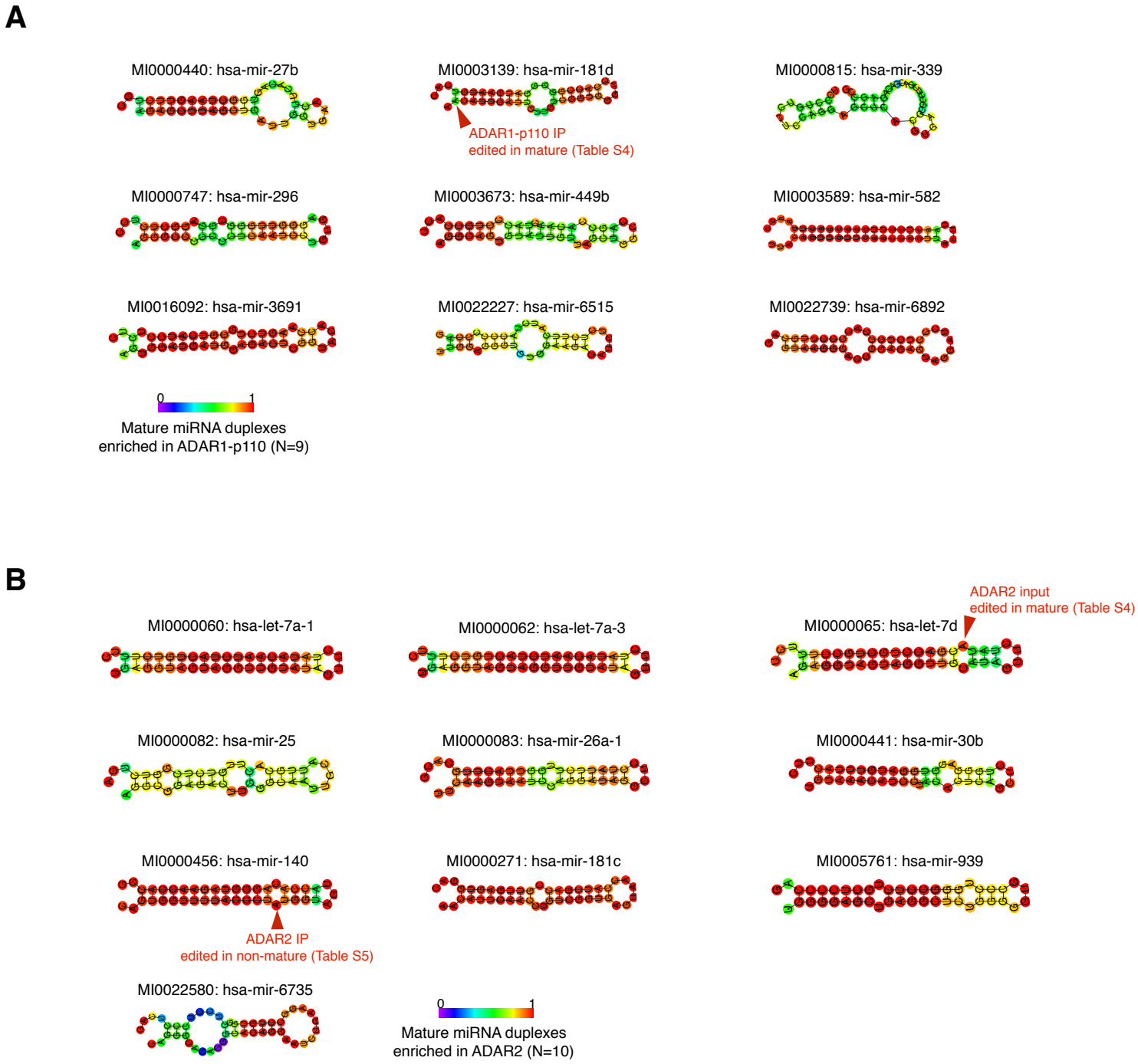


Figure S5. Predicted secondary structures of pre-miRNA that were enriched in the IP samples of ADAR1-p110 (A) versus ADAR2 (B) as shown in Figure 3C and D. Individual secondary structures were predicted by CentroidFold with default parameters. Gradient color corresponds to base-pairing and loop probability for each base.

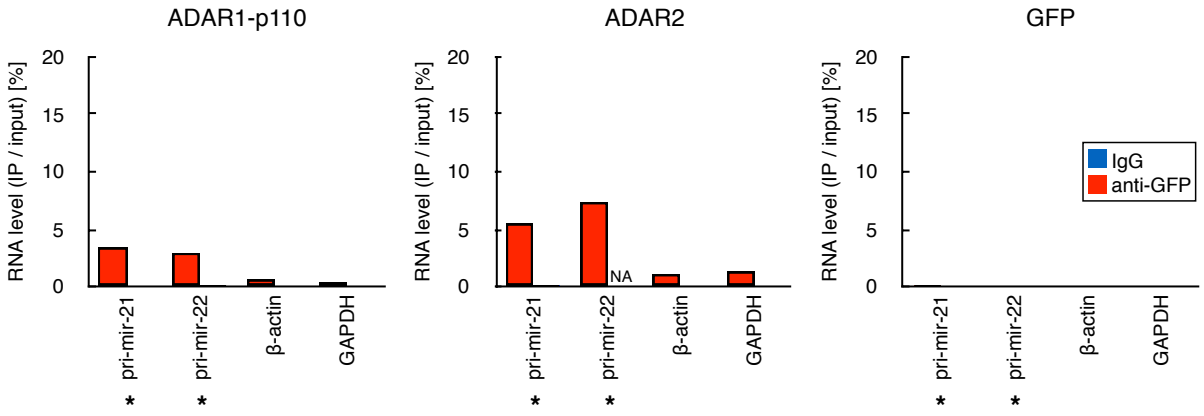
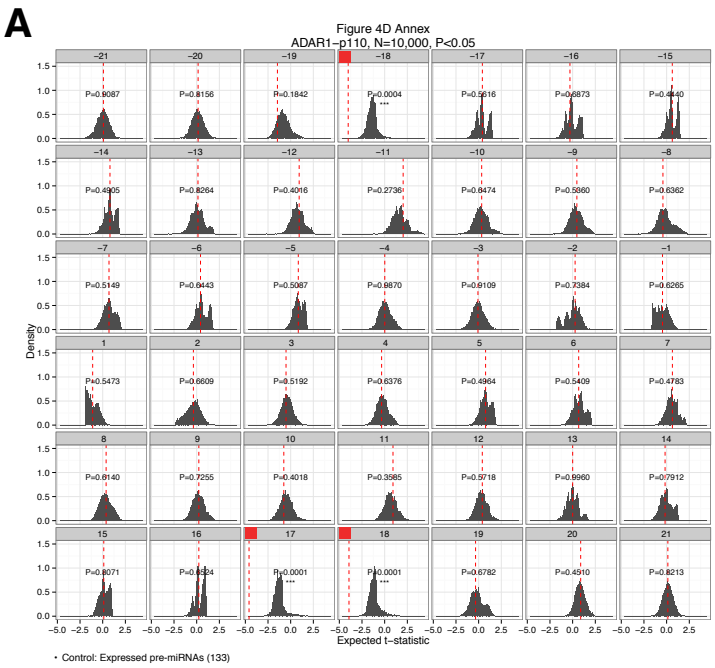
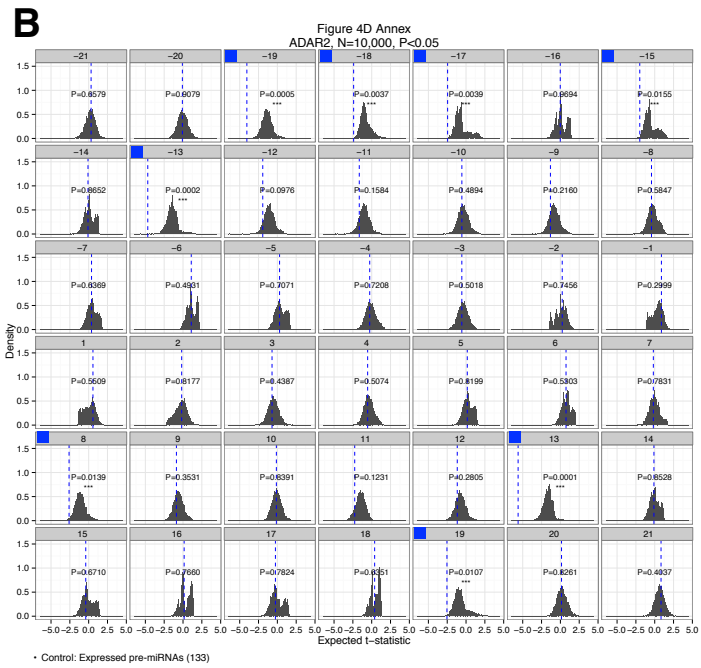
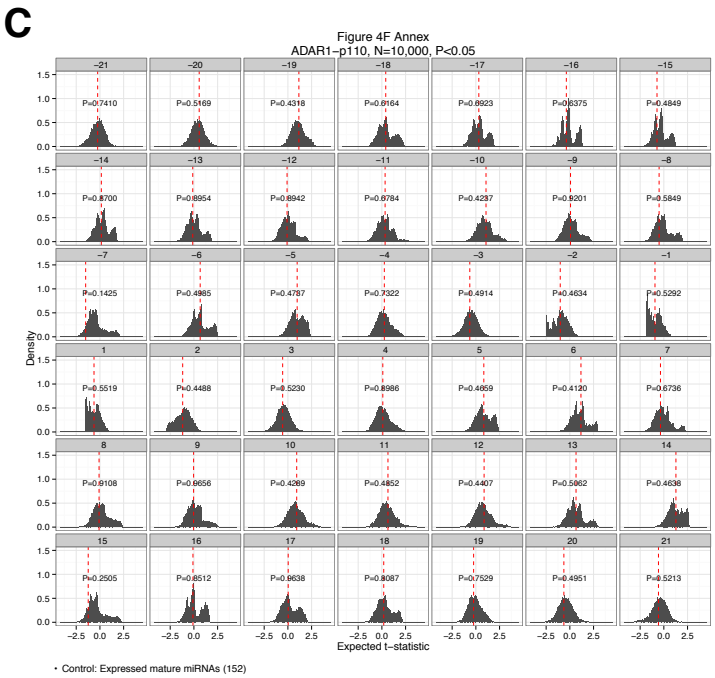
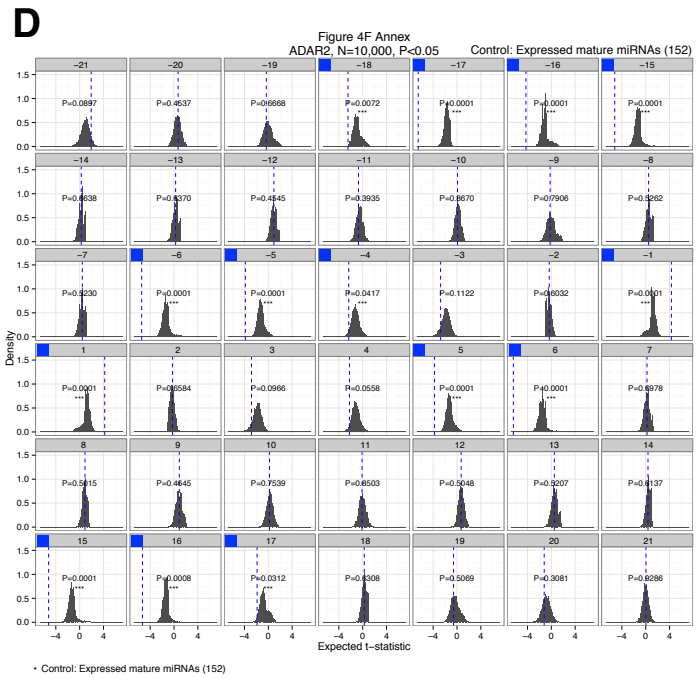


Figure S6. Results of real-time PCR for pri-miRNAs using purified total RNA from IP and input samples of cells overexpressing mGFP-ADAR1-p110 (left), mGFP-ADAR2 (center), and mGFP (right). The results for anti-GFP and control IgG IP samples antibody are shown in red and blue bars, respectively. The ratio of IP to input indicates relative RNA expression levels. Stars (pri-mir-21 and pri-mir-22) show endogenous pri-miRNAs. β-actin and GAPDH are used as controls.









(Legend on next page)

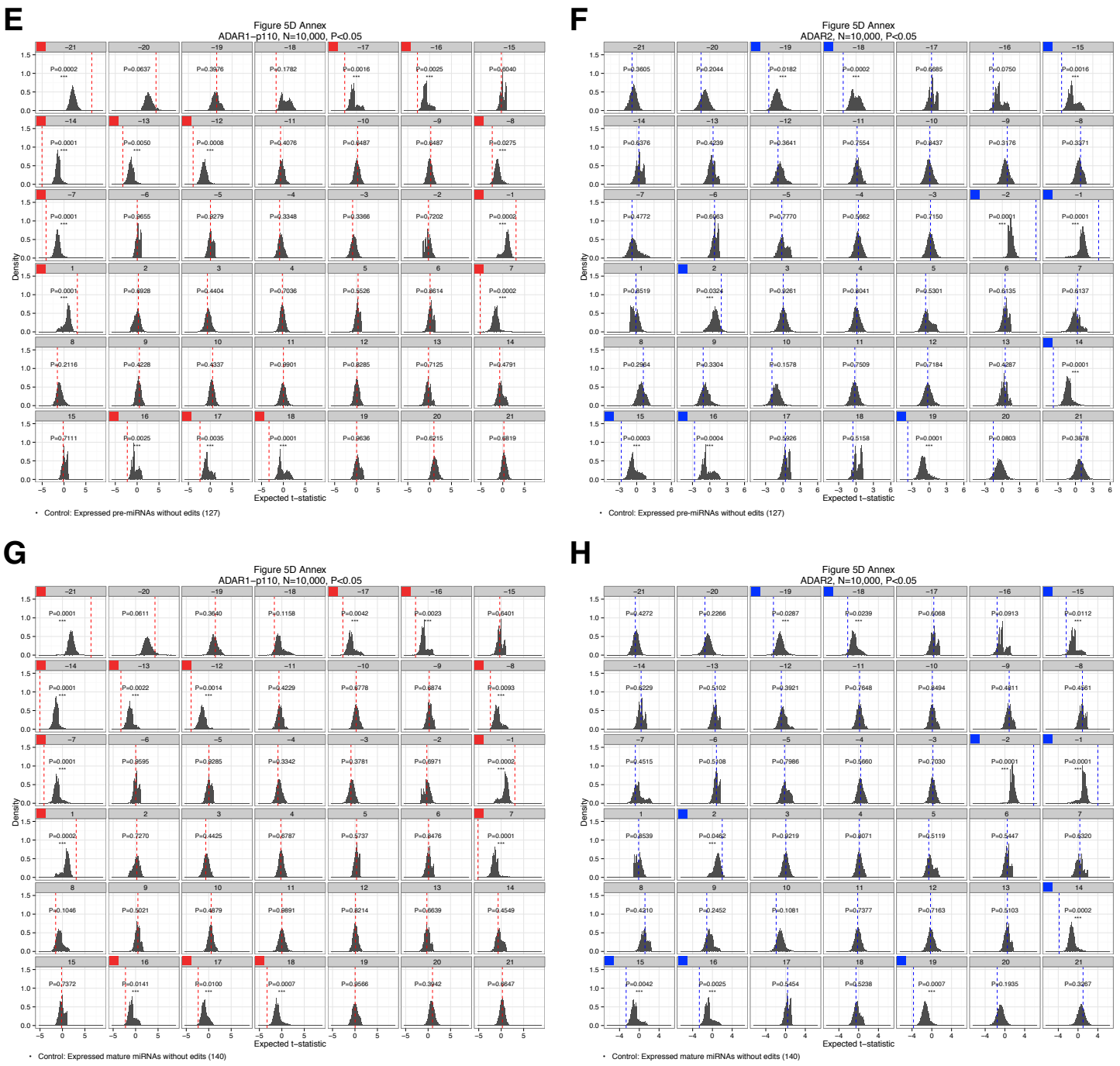


Figure S7. Bootstrap t-statistical test. Distribution of bootstrap t-statistic at each position on pre-miRNAs enriched in ADAR1-p110 IP (A) or ADAR2 IP (B) compared to all expressed pre-miRNAs, or candidate mature miRNA duplexes enriched in ADAR1-p110 IP (C) or ADAR2 IP (D) compared to all candidate mature miRNA duplexes expressed in both input samples, and their levels of significance for each position. Distribution of bootstrap t-statistic at each position in pre-miRNAs enriched in ADAR1-p110 IP (E,F) or ADAR2 IP (G,H) compared to the pre-miRNAs (E,G) or candidate mature miRNA duplexes expressed in both input samples (F,H), and their level of significance for each position. Vertical axis: density; horizontal axis: expected t-statistic. Red or blue dashed lines: observed t-statistic; red or blue squares: positions for which the difference was significant.

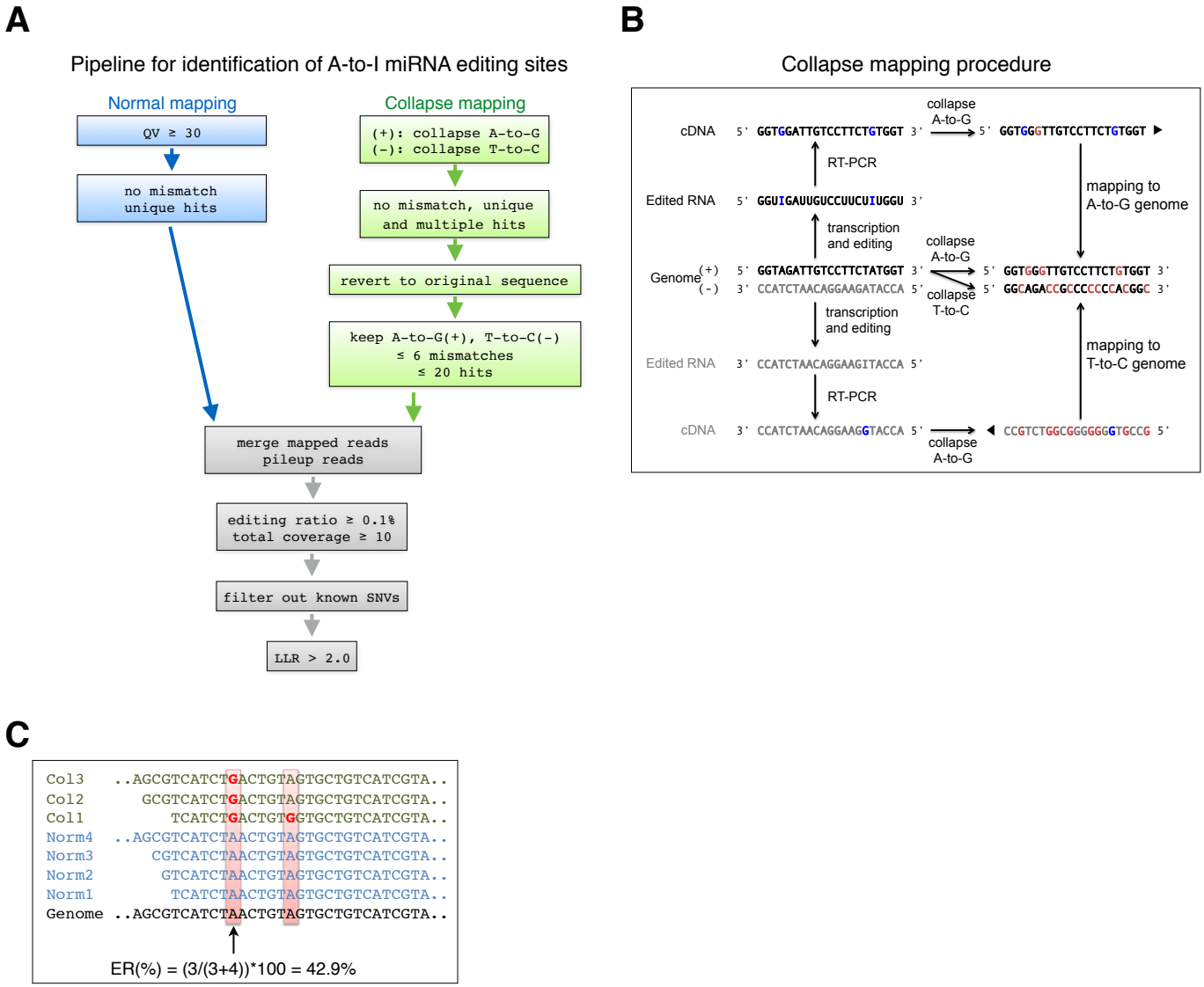


Figure S8. (A) Computational pipeline for identification of A-to-I miRNA editing sites by normal and collapse mapping procedures. Blue (normal mapping): reads which had at least one nucleotide with a quality value (QV) of less than 30 were discarded; mapping was performed with the output of "unique hits with no mismatches". Green (collapse mapping): mapping was performed with the output of "multiple hits with no mismatches", the remainder (no hits) and the reference genome were collapsed (A-to-G for forward reads and (+) strand genome; T-to-C for reverse reads and (-) strand genome); the collapsed reads were mapped again to output "unique and multiple hits with no mismatches"; collapsed sequences were reverted back to their original sequence to identify mismatches against the reference genome; reads with mismatches other than A-to-G for (+) reads and T-to-C for (-) reads were discarded; reads with more than 6 mismatches were discarded; collapsed reads which had mapped to more than 20 locations in the genome were discarded. Gray (identification of editing sites): merge "normal" and "collapse" uniquely mapped reads; output per-nucleotide coverage; calculate A/(A+G) editing ratio (ER) for each A-to-G site; sites with ER ≥ 0.1% and with a coverage of 5 or more were validated as editing sites. (B) Procedure for A-to-G collapse mapping. After RT-PCR, the positions corresponding to an A-to-I editing site carry a "G". All "A"s in (+) reads were collapsed to "G"s. All "A"s in (-) reads were collapsed to "G"s. All "A"s in the (+) strand of the genome were collapsed to "G"s. All "T"s on the (+) strand of the genome were collapsed to "C"s. After alignment of the collapsed (+) and (-) reads to the A-to-G and T-to-C collapsed genome sequences, respectively, the reads and the genome were reverted back to their original sequence and compared. See Materials and Methods section for further details. (C) Illustration of the visual output of computational analysis for normal and collapse reads aligned against a reference genome. Non-edited reads (blue) are obtained by normal mapping as described in (A), and potentially edited reads containing A-to-G mismatches (green) are selected by collapse mapping, as described in (A).

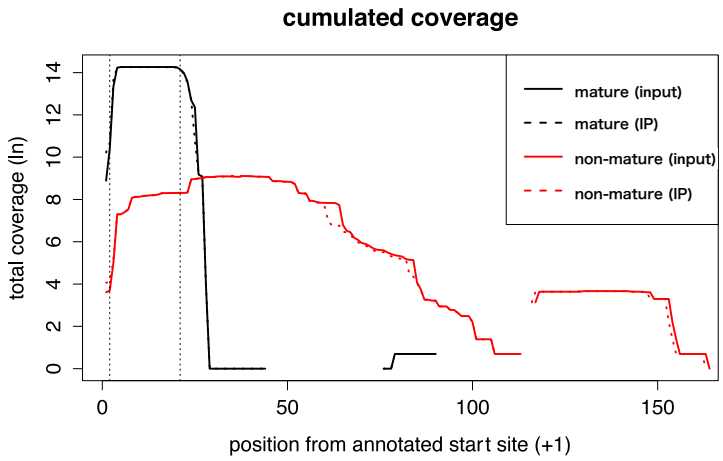


Figure S9. Cumulated coverage for all annotated mature (black) or non-mature miRNAs (red) given by the sum of all input (full lines) and IP (dotted lines) samples. X-axis: nucleotide position from the 5' end of miRNA or pre-miRNA annotation (starting from +1); y-axis: log2 scale of the per-nucleotide sum of coverages.

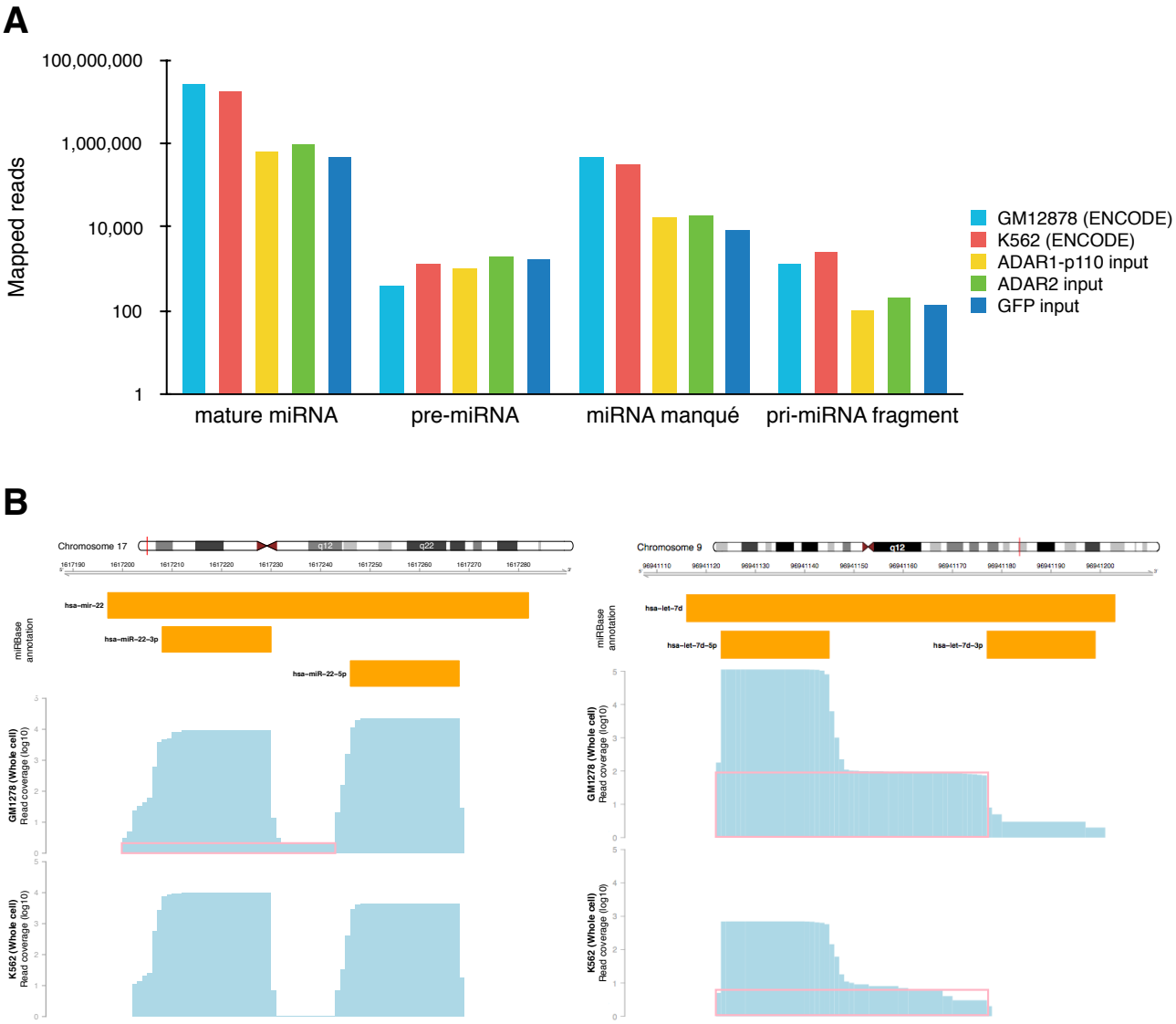


Figure S10. ENCODE small RNA-seq data analysis confirming the expression of miRNA manqué. (A) Total mapped reads of our ADAR sequencing data alongside ENCODE small RNA-seq data of GM12878 (light blue) and K562 cells (red). X-axis shows the total mapped reads for each type of miRNA species. (B) Example of miRNA manqué transcripts found in miR-22 and let-7d loci from different two cell lines, GM1278 and K562, were highlighted (light pink).

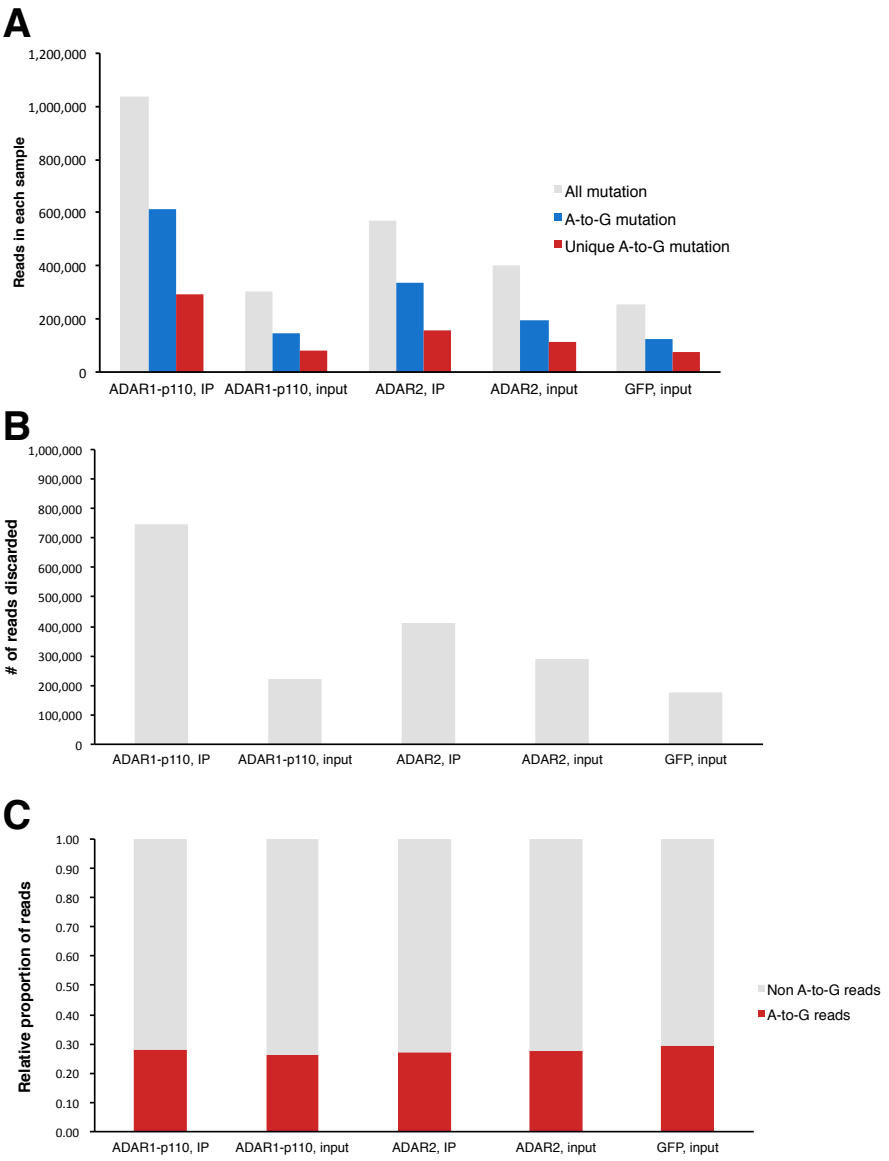


Figure S11. The numbers and percentages of A-to-G mismatched reads among all types of mismatches. (A) The number of reads of all types of mismatches containing A-to-G, A-to-C, A-to-U, G-to-C, G-to-U, G-to-A, C-to-U, C-to-A, C-to-G, U-to-A, U-to-G, and U-to-C in ADAR1-p110 IP and input, and ADAR2 IP and input, and GFP input samples, respectively. Gray bar indicates the number of reads with all types of mismatches. Blue bar, the number of reads with A-to-G mismatches in single reads. Red bar, the number of reads with unique A-to-G mismatches in paired reads.The numbers of reads that have non A-to-G mismatches in each library, which are the subtracted numbers of blue bar from gray bar, are ADAR1-p110 IP: 427,316 reads, ADAR1-p110 input: 155,175 reads, ADAR2 IP: 232,075 reads, ADAR2 input: 207,307 reads, GFP input control: 125,611 reads, respectively. (B) The numbers of discarded reads which have non A-to-G mismatches in each sample. (C) The percentages of discarded reads in each sample.