Targeted RNA editing: novel tools to study post-transcriptional regulation

TJC 11-01-22

Molecular Cell

CellPress

Technology review

Targeted RNA editing: novel tools to study post-transcriptional regulation

Weijin Xu,[1,3,4] Jeetayu Biswas,[2,3,5] Robert H. Singer,[2] and Michael Rosbash[1,*]

[1]Howard Hughes Medical Institute, Department of Biology, Brandeis University, Waltham, MA 02451, USA

[2]Department of Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, NY 10461, USA

[3]These authors contributed equally

[4]Present address: CD Capital, Shanghai 200041, People's Republic of China

[5]Present address: Department of Medicine, Weill Cornell Medicine, New York Presbyterian Hospital, New York, NY 10021, USA

*Correspondence: rosbash@brandeis.edu

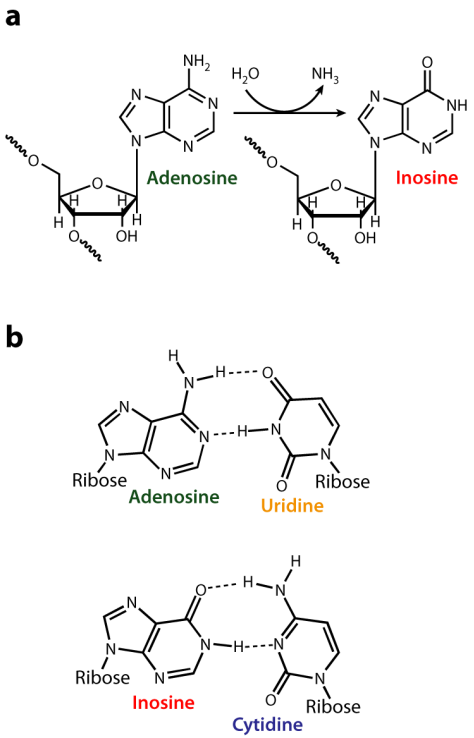
https://doi.org/10.1016/j.molcel.2021.10.010

Introduction to RNA editing

- Adenosine-to-inosine (A-to-I) editing
- Cytosine-to-uracil (C-to-U) editing

Adenosine-to-inosine (A-to-I) editing in mRNA

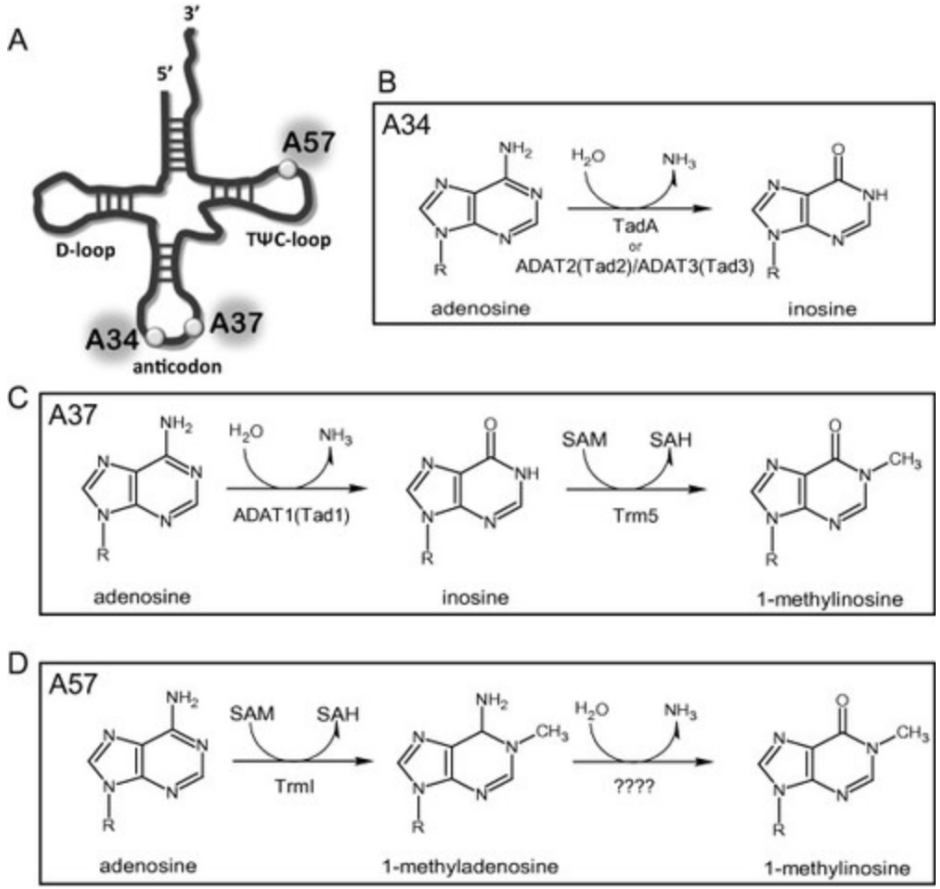
- Hydrolytic deamination of adenine base to inosine (a) by **adenosine deaminase acting on RNA** (*ADAR*) enzymes
- Adenosine pairs with Uridine, but Inosine pairs with Cytidine leading to a A>G or C>T mismatch on translation (b)



Nishikura K. 2010.
Annu. Rev. Biochem. 79:321–49

Adenosine-to-inosine (A-to-I) editing in tRNA

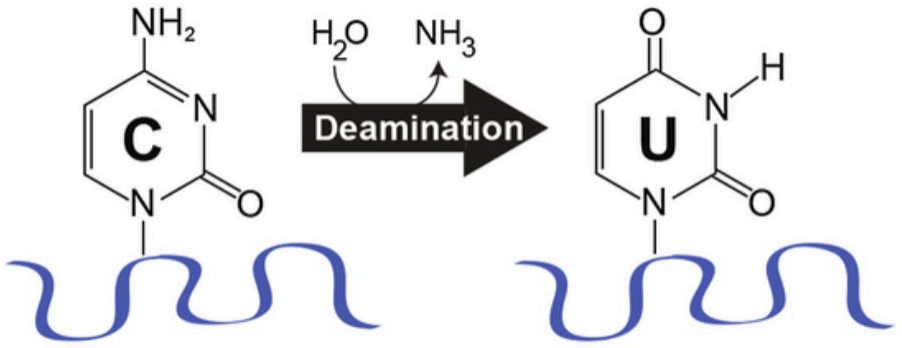
- Deamination of ADAT mechanism similar to ADAR
- Two inosines exist as 1-methyliosines
- Inosine can recognize A, C, U, expanding the number of recognizable codons



Torres AG et al., FEBS Letters 2014

Cytosine-to-uracil (C-to-U) editing

- The APOBEC (apolipoprotein B mRNA editing catalytic polypeptide-like) family of proteins regulate C-to-U editing
- Minor functions in mammals under physiological conditions, but important in pathogen defense



Refsland & Harris, Curr. Topics Microb. Immu. 2014

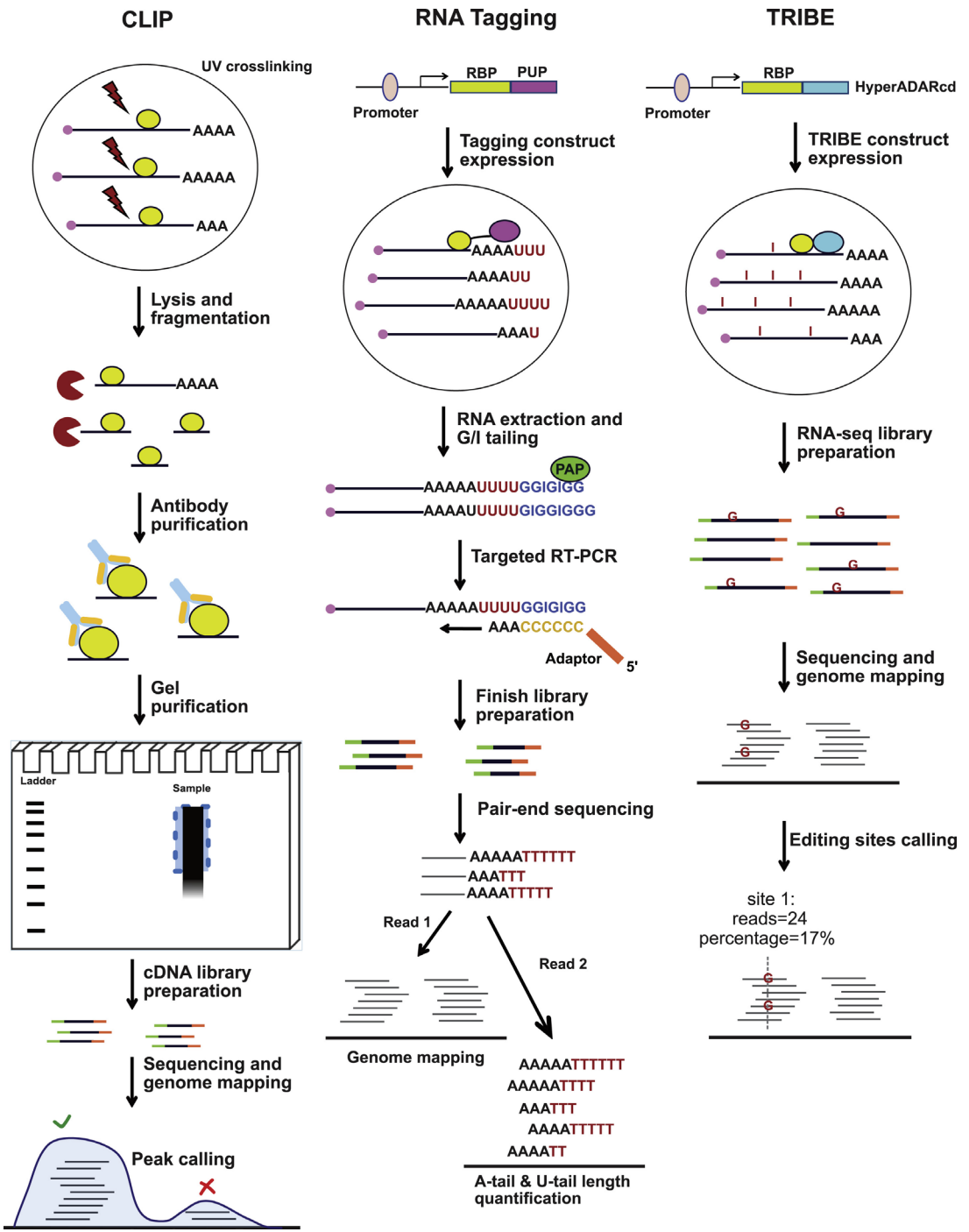
RNA editing applications – **(1) identification of RNA-binding protein (RBP) transcripts**

- RBPs are fundamental to all aspects of RNA processing, including alternative splicing, polyadenylation, nuclear export, intracellular localization, degradation, and translation, such as:
  - TDP-43 in ALS/FTLD
  - FMRP in Fragile X syndrome
  - SMN1 in SMA
- RBP-target identification is important for understanding disease physiology

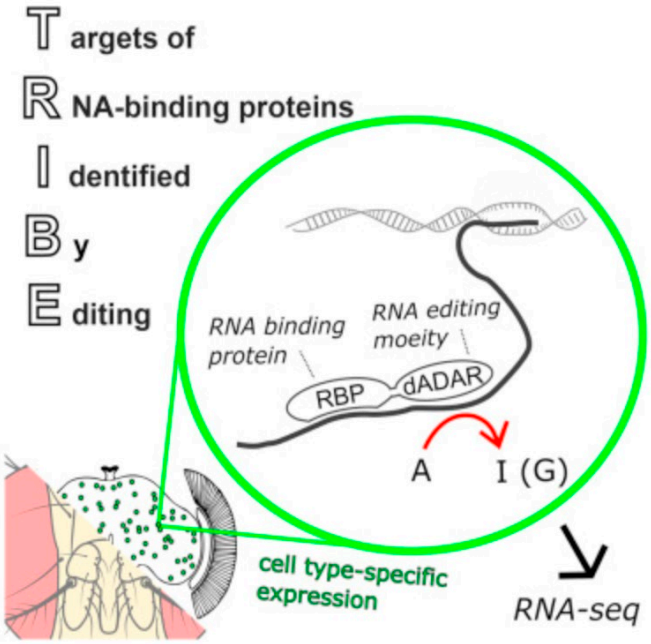
Using ADAR catalytic domain to discover RBP-binding targets

- Standard techniques for profiling RBP targets, such as crosslinking immunoprecipitation (CLIP) and its variants, are limited or suboptimal in some situations, e.g. when compatible antibodies are not available and when dealing with small cell populations such as neuronal subtypes and primary stem cells.
- Coupling of ADAR catalytic domain (ADARcd) to RBPs could aid transcripts
- RNAs are tagged in vivo, transient binding can also be picked up

3 identification techniques for targets of RBPs



TRIBE/hyperTRIBE: Targets of RNA-binding Proteins Identified by Editing

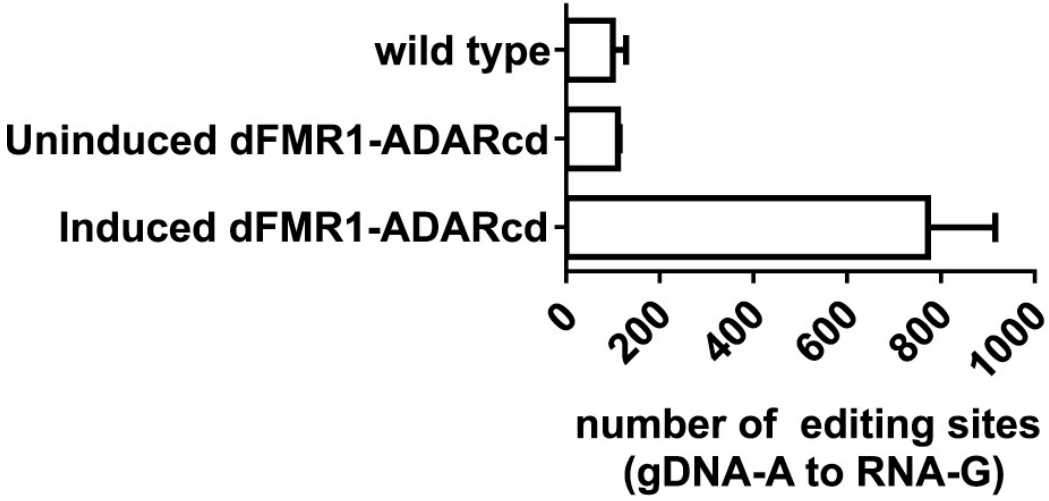


- Double-stranded RNA-binding deficient ADARcd is fused to RBP, making it specific for RNA editing (APOBEC also edit ssDNA)

McMahon et al., Cell 2016; Xu et al., RNA 2018

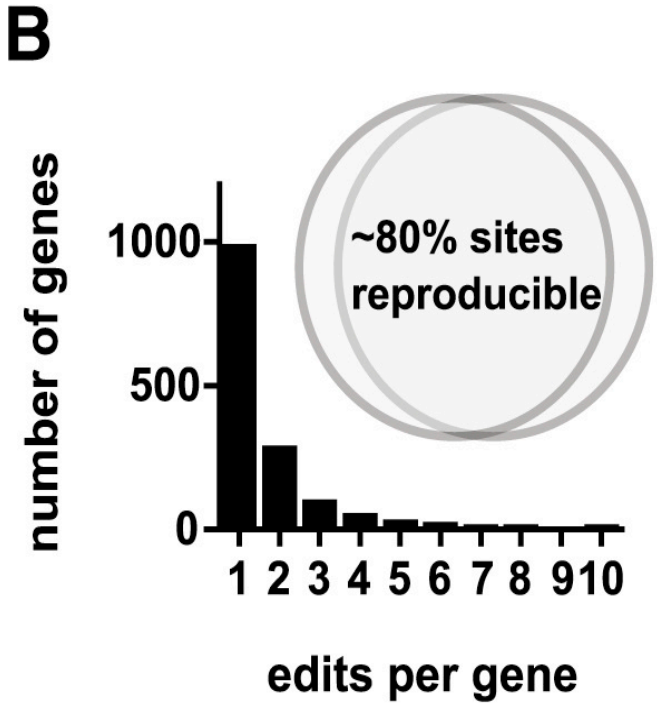
Hrp48-TRIBE Fusion Protein

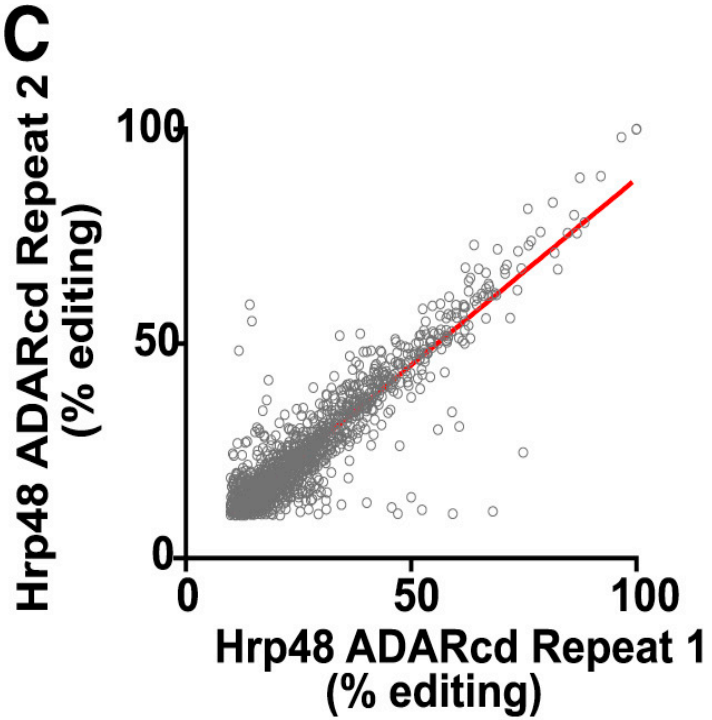
- Drosophila Hrp48 is a homolog of the mammalian hnRNP A/B family, which is well-characterized and avails of established antibodies
- Establishing inducible Gal4/UAS expression system



McMahon et al., Cell 2016; Xu et al., RNA 2018

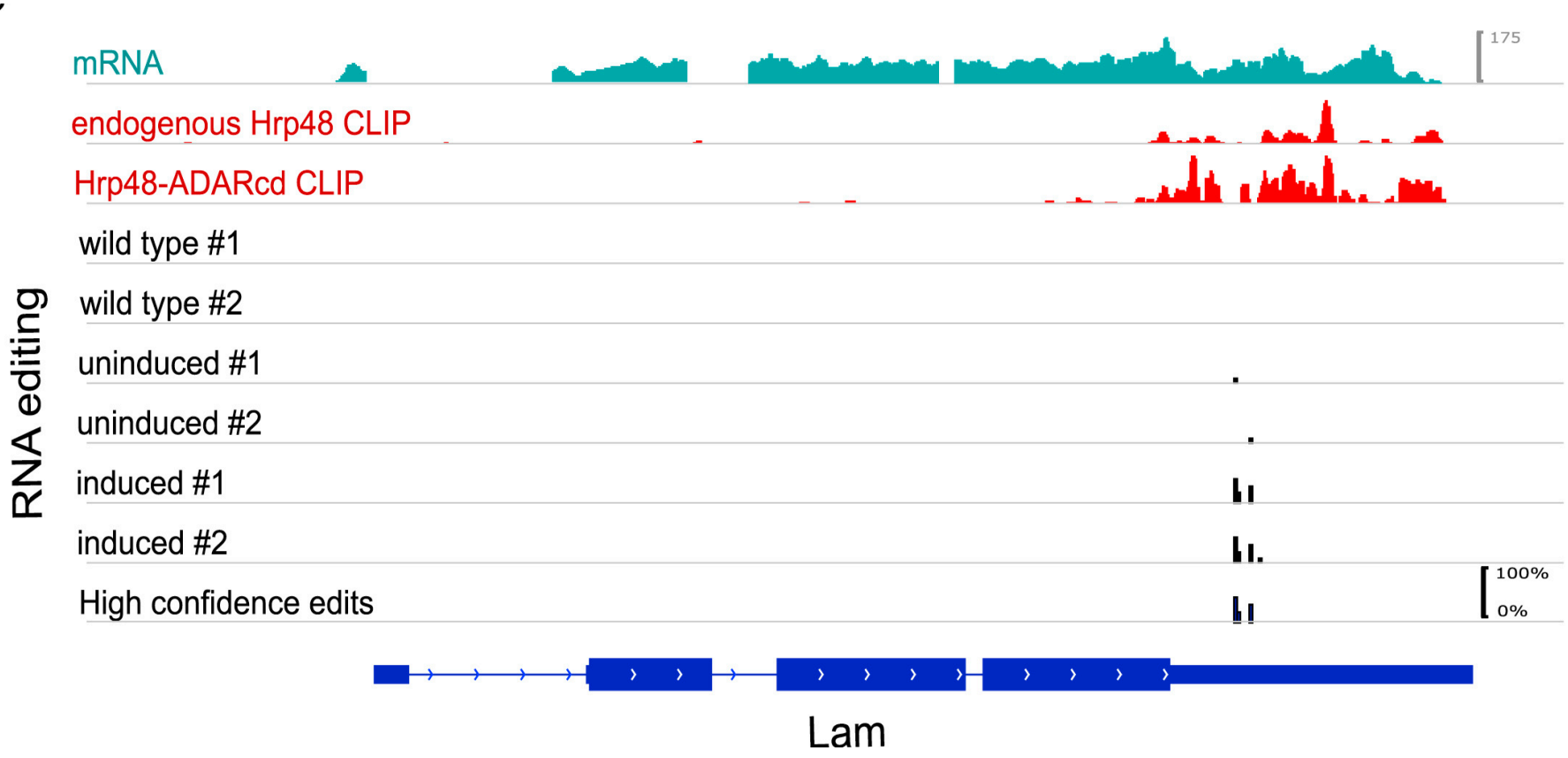
Hrp48-TRIBE induces mostly 1, highly reproducible editing event





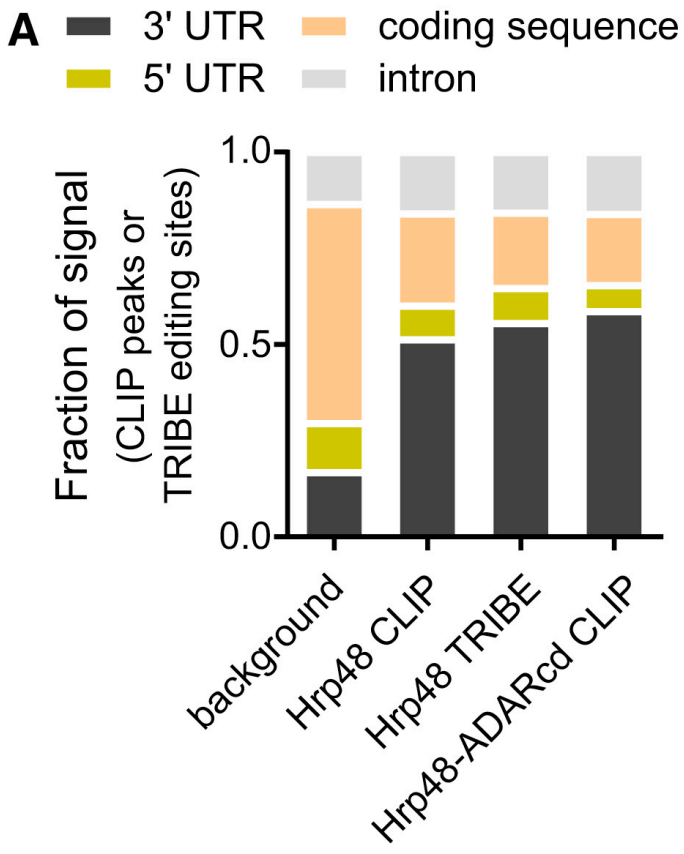
McMahon et al., Cell 2016; Xu et al., RNA 2018

CLIP and TRIBE are strongly enriched in the 3’ UTR



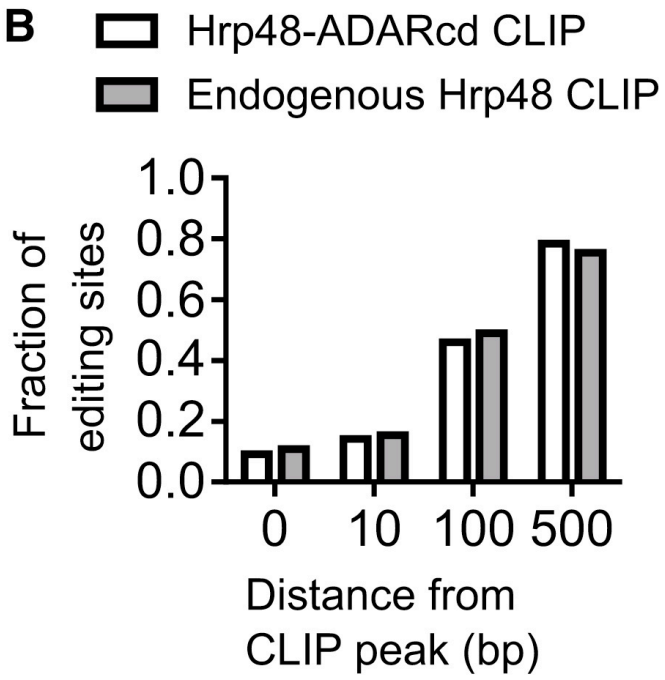
3‘ UTR binding does not affect function of RBP

- Immunohistochemistry showed regular cytoplasmic localization of Hrp48



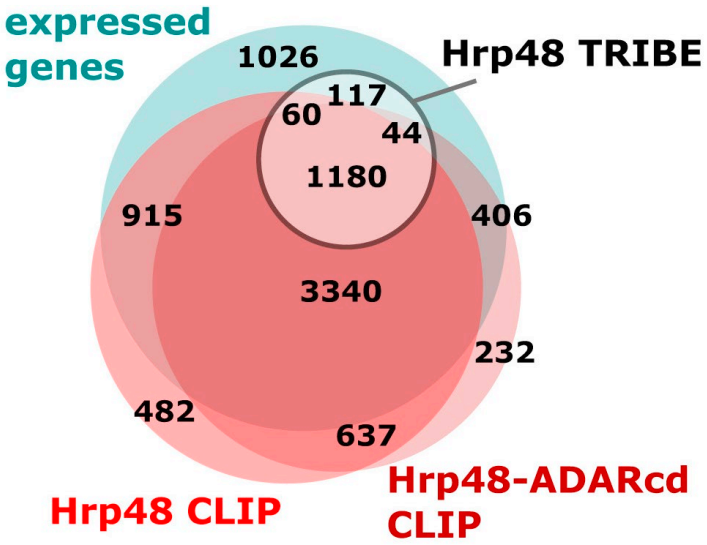
Characterization of editing peaks

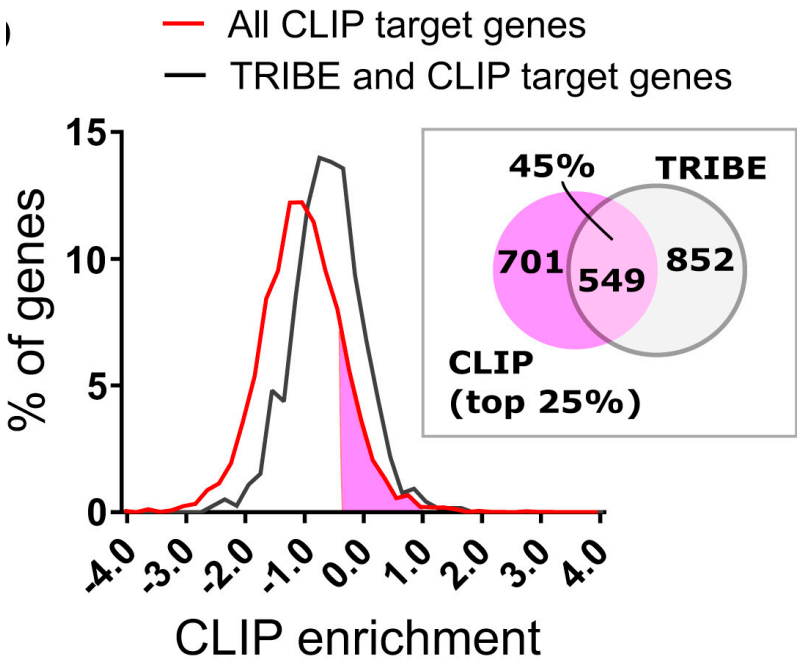
- ~ 80% of editing occurs < 500 bp from CLIP peak near RBP binding site



TRIBE vs. CLIP

- Most CLIP hits are also found by TRIBE

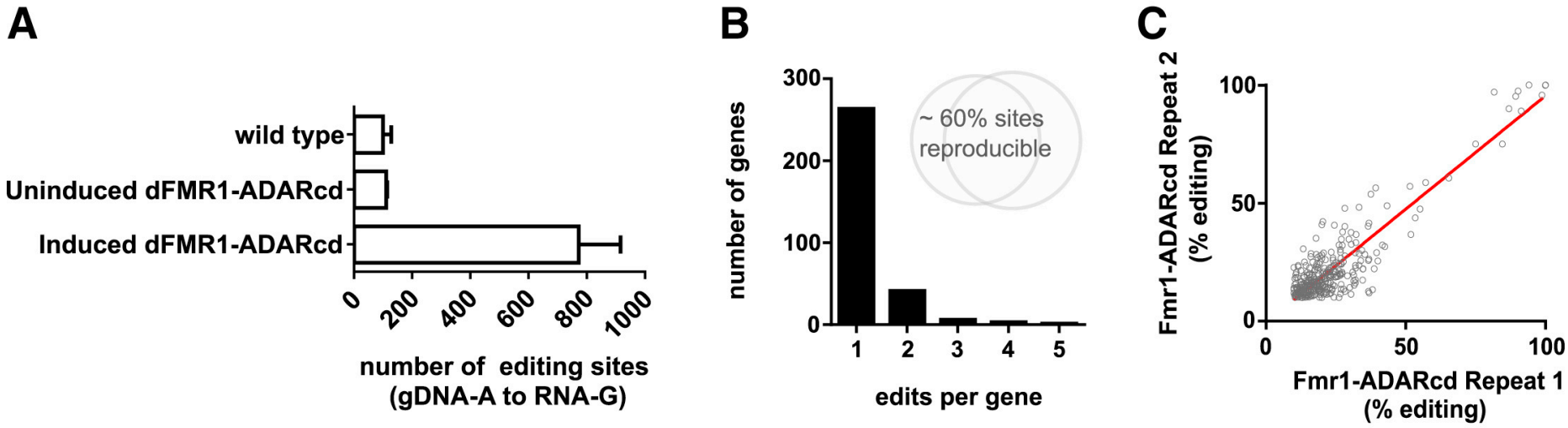




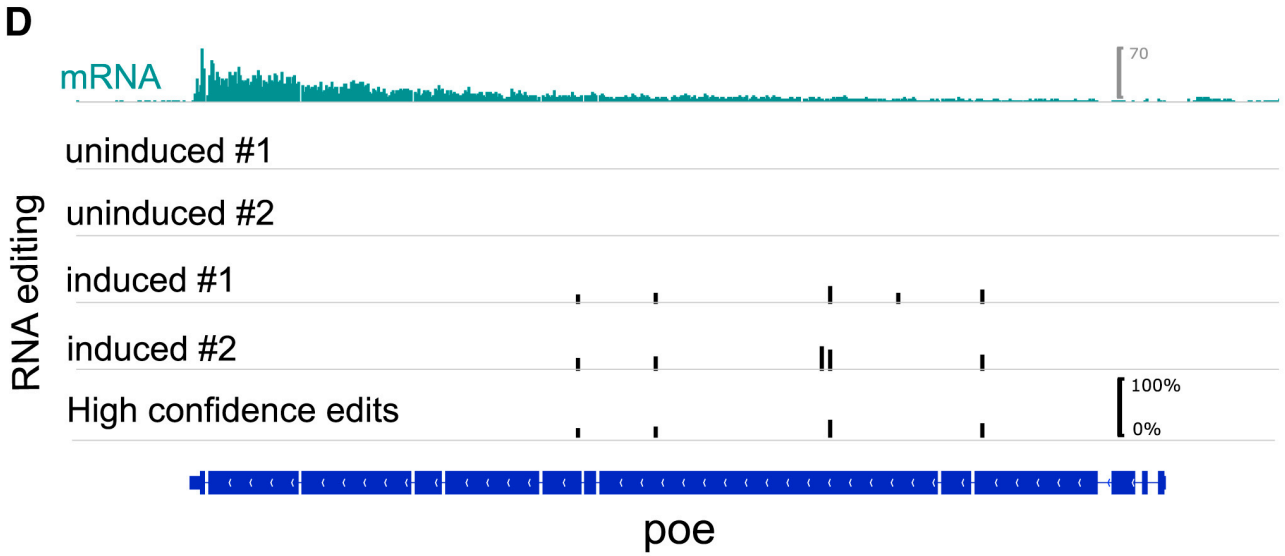
- TRIBE tends to detect freq. occurring CLIP hits

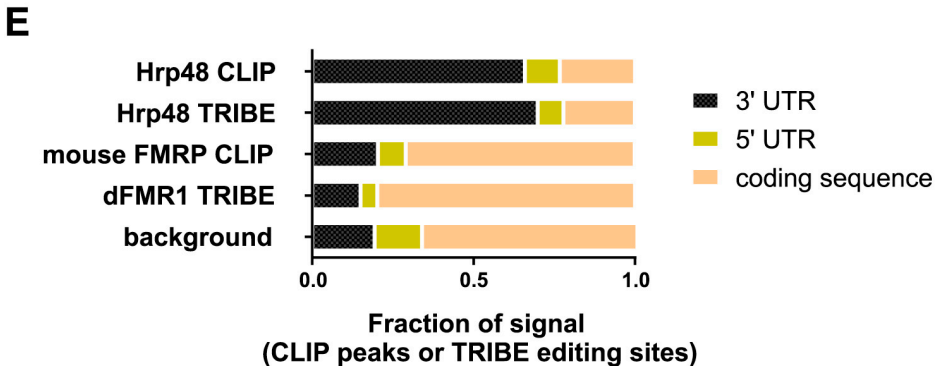
Generation of dFMR1-TRIBE

- dFMR1 is the Drosophila ortholog of FMRP
- Similarly to Hrp48, dFMR1 was fused to ADARcd



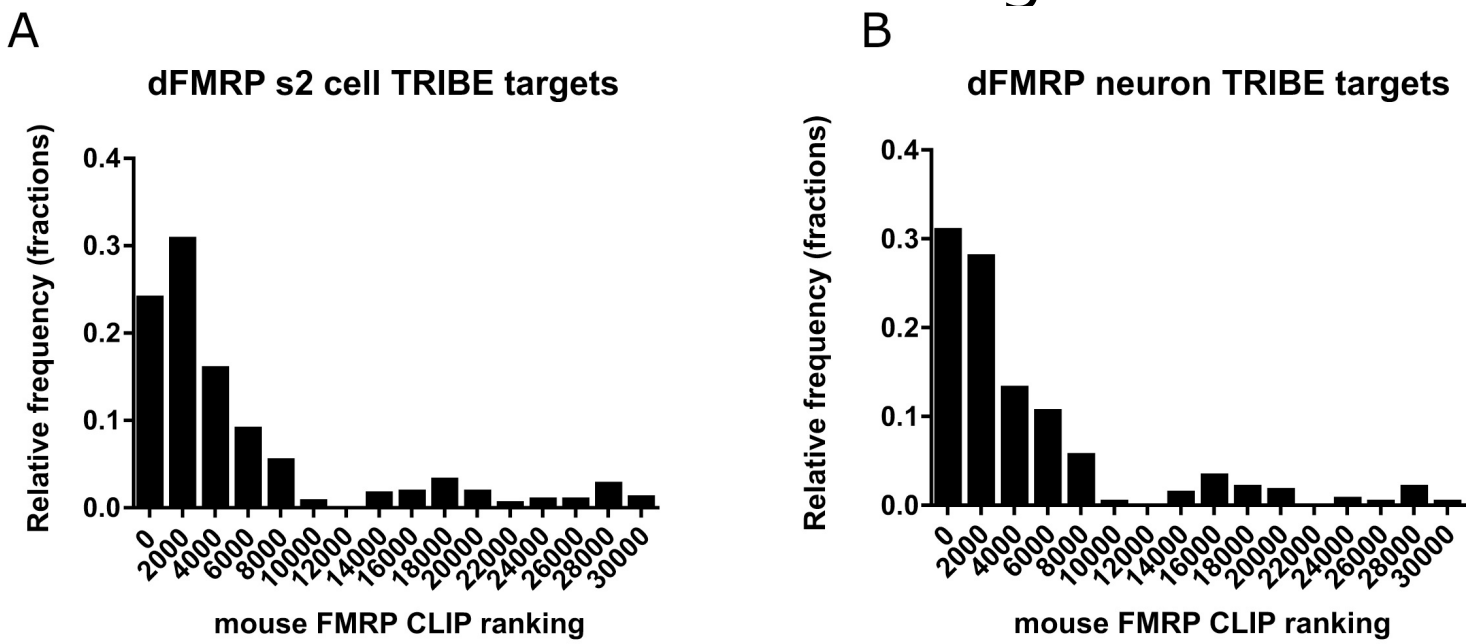
dFMR1-TRIBE Preferentially Edits Coding Sequence, Reflecting Prior CLIP data





Comparison to mammalian FMRP

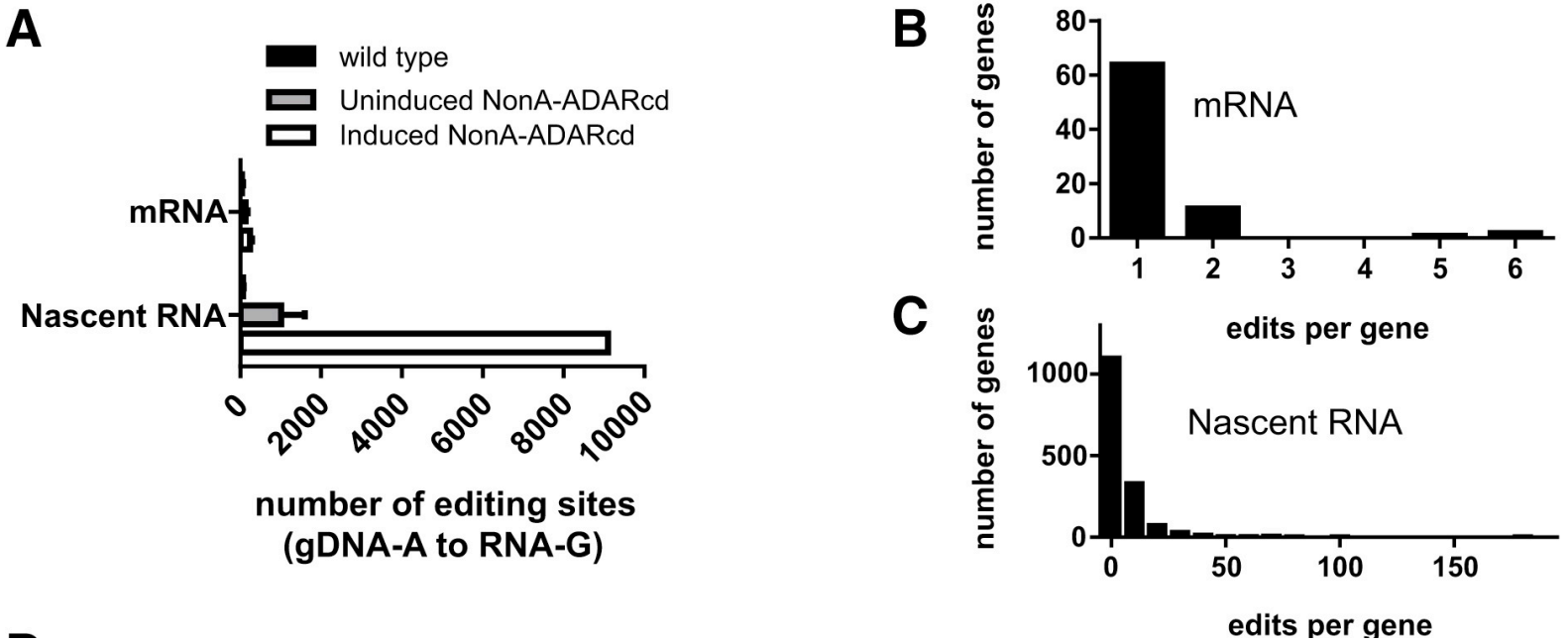
- Examination of mouse brain FMRP CLIP targets revealed that the mouse homologs of the dFMR1-TRIBE target genes were biased toward higher CLIP rankings, suggesting that dFMR1-TRIBE identifies conserved targets of FMRP



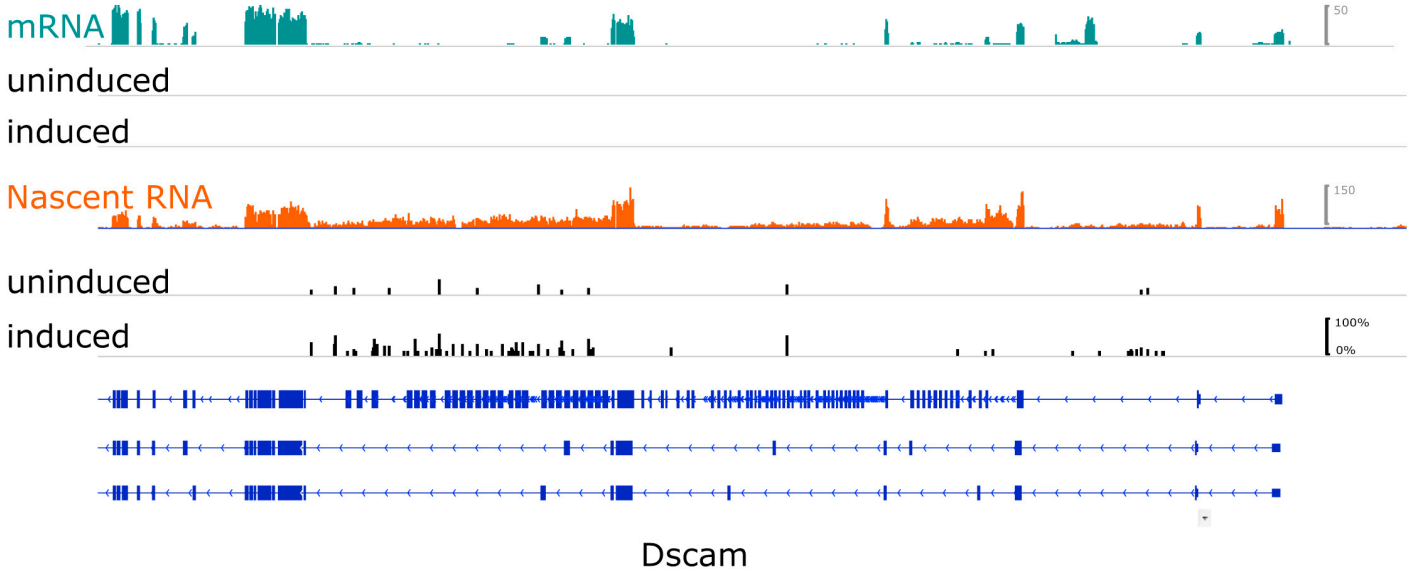
Probing NonA-TRIBE

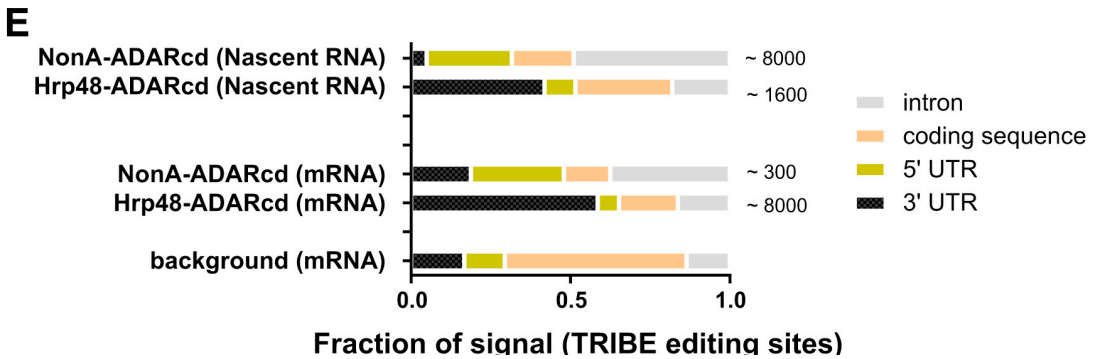
- NonA is ortholog of mammlian NonO
- NonO is a multifunctional protein involved in the function of nuclear paraspeckles as well as other nuclear events like splicing, mRNA export and the regulation of transcription
- NonA is involved in circadian biology
- NonA-TRIBE led only to moderate increase of mRNA editing

NonA-TRIBE preferentially edits nascent RNA at high rates, consistent with its nuclear function



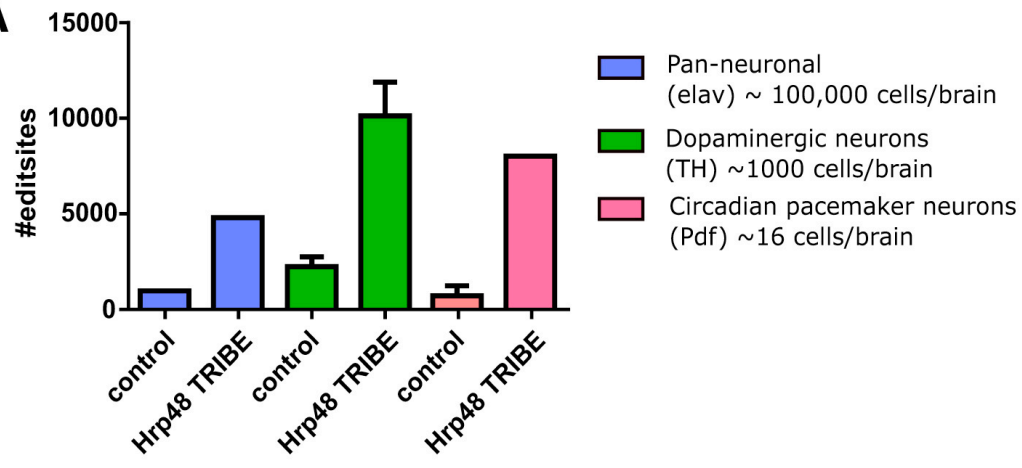
NonA-TRIBE editing events are largely intronic, Hrp48-TRIBE does not show preferential intronic editing in nascent RNA



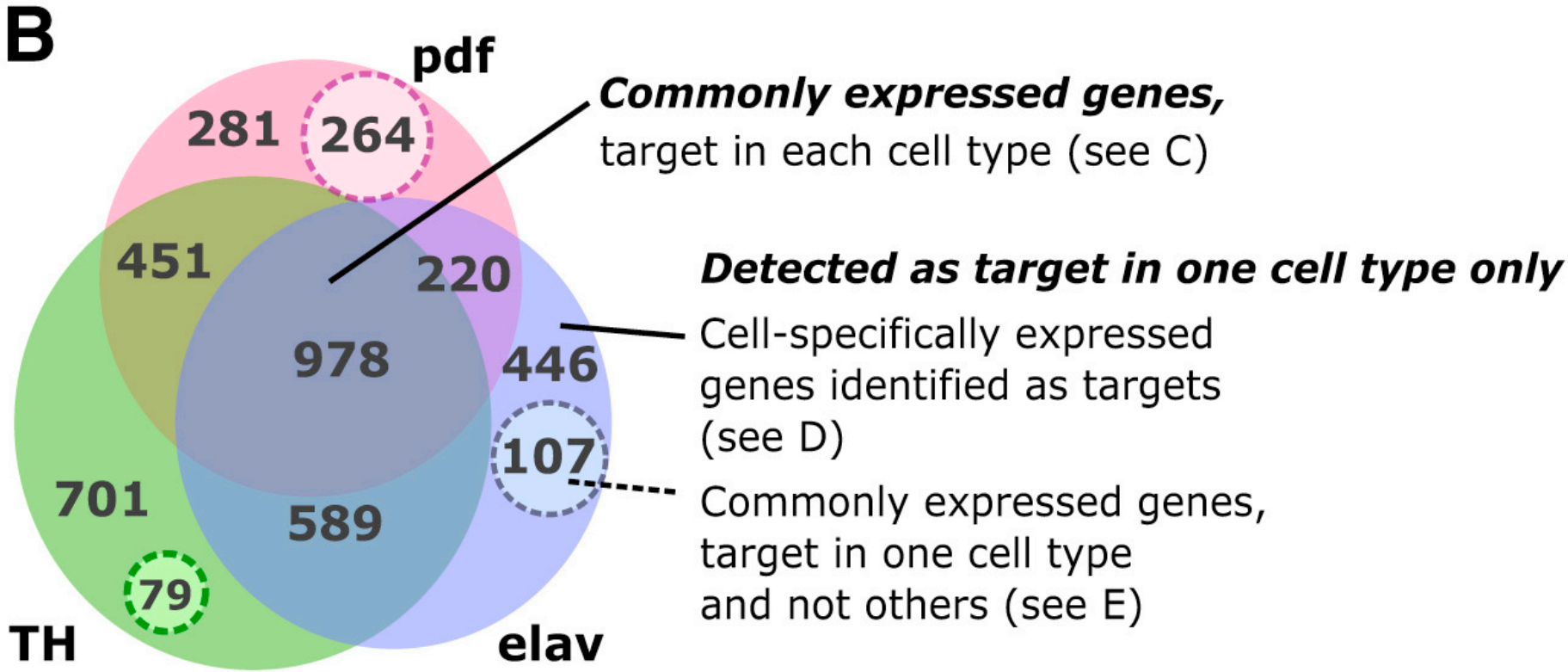


Cell-specific expression of Hrp48-TRIBE in fly neuronal subsets

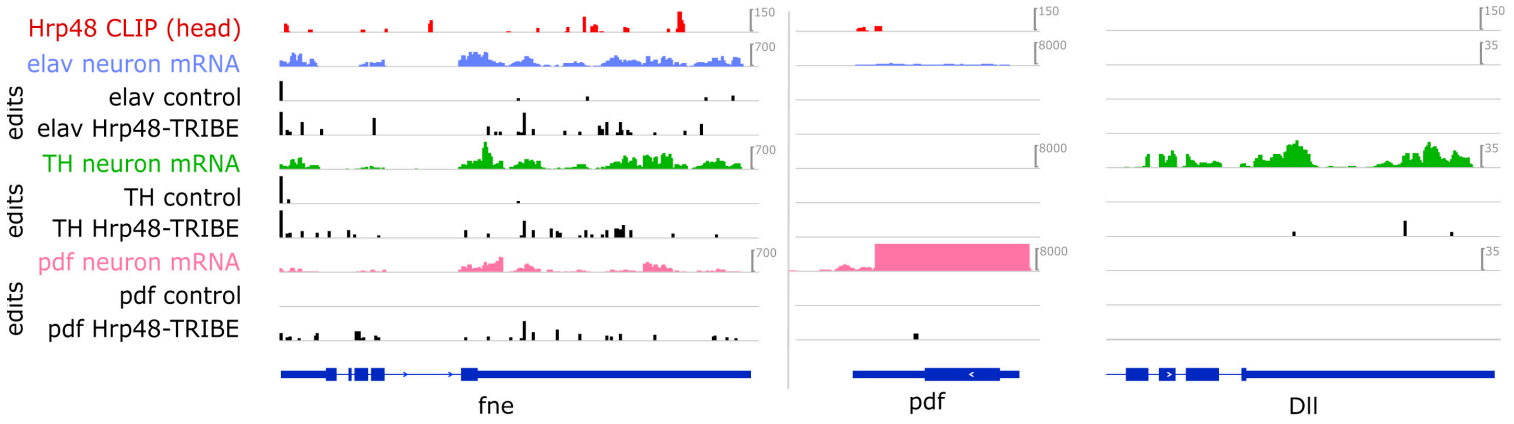
- UAS in combination with Gal4 driver lines, additional UAS-eGFP for cell sorting in:
  - core circadian PDF neuropeptide expressing cells (*pdf*-Gal4, ~16 cells/brain)
  - dopaminergic neurons (Tyrosine hydroxylase, *TH*-Gal4, ~1,000 cells/brain)
  - all neurons (pan-neuronal driver, *elav*-Gal4, ~100,000 cells/brain)

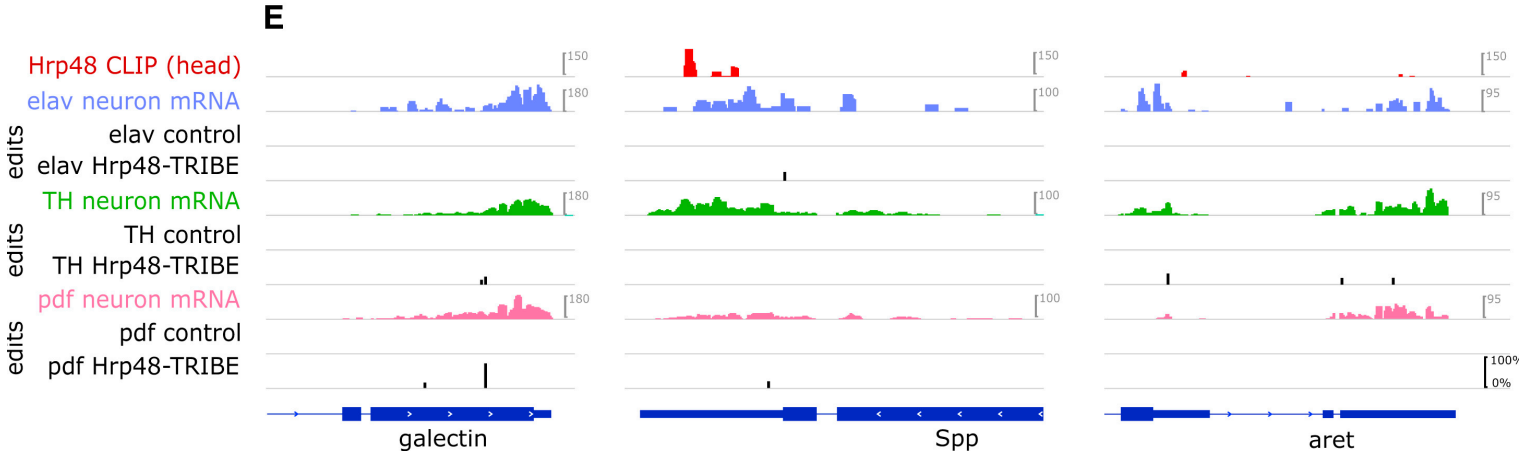


Target gene sets were similar in numbers and overlapping, but not identical



Identification of cell-specific genes

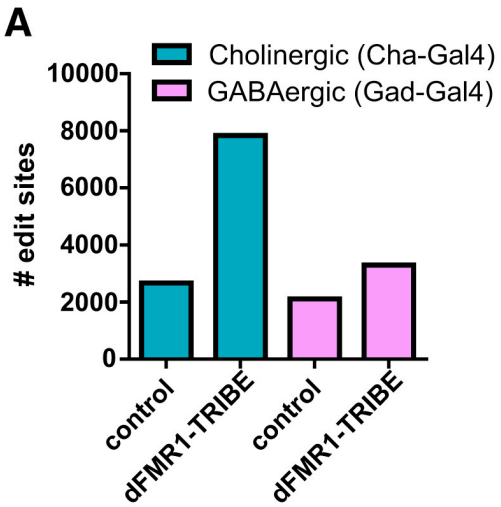


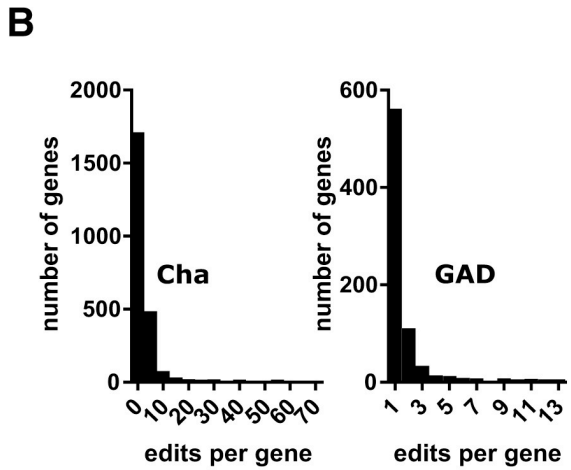


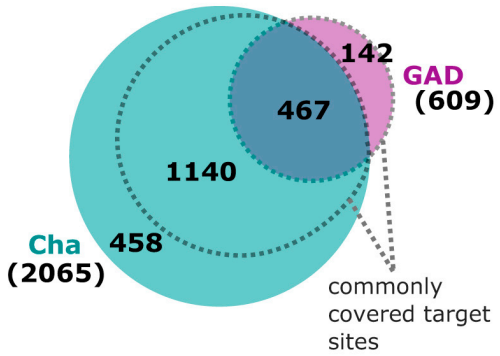
fne : pan-neuronal gene; pdf : almost only in pdf neurons expressed + edited;
Dll : only expressed + edited in TH neurons; galectin+Spp+aret expressed in all neurons but edited in specific subsets

dFMR1-TRIBE Can Identify RBP Targets in Specific Cells

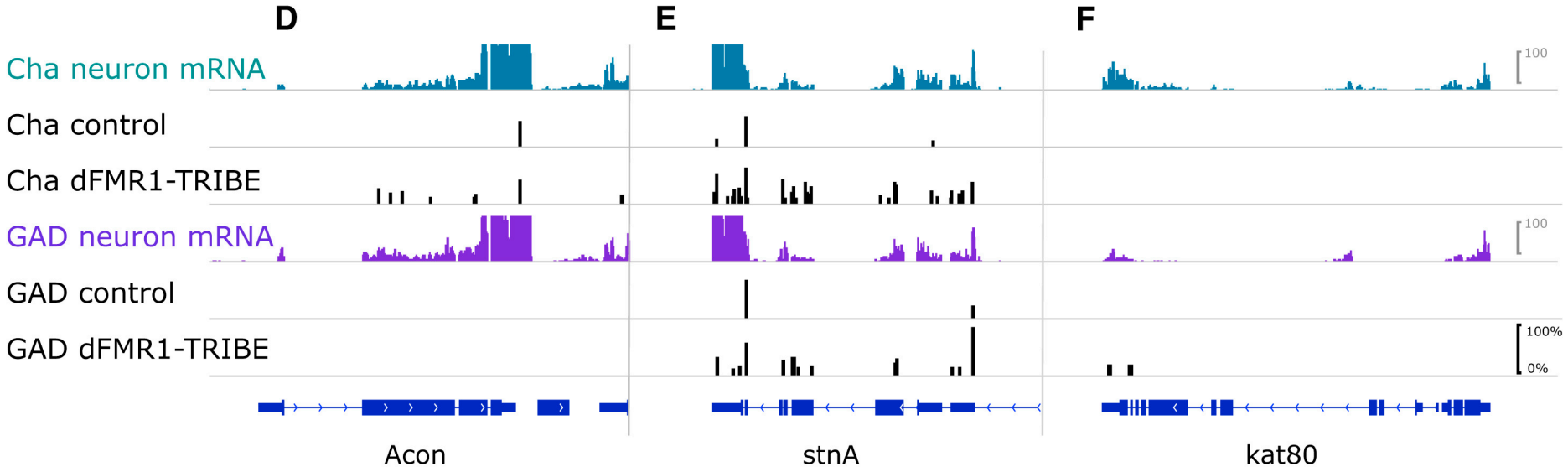
- Examination of dFMR1 targets in excitatory (cholinergic; Cha-Gal4) and inhibitory (GABAergic; GAD-Gal4) cells, corresponding to affected cells in mammalian pathology
- Much more excitatory, than inhibitory targets, maybe du to lower GAD-Gal4 TRIBE expression







Cell-specific RBP targets



- 45% of robust mouse brain CLIP targets that have clear fly homologs were also dFMR1-TRIBE targets in Cha-neurons

Conclusion

- New method for sensitive detection of editing events (16 neurons/brain, minimal input for sequencing 150 cells)
- Doesn't need antibodies, crosslinking or RNases, which can have a bias by themselves
- TRIBE can mark RNA over longer time periods and not a snapshot

- Systemic bias of CLIP cannot be assessed because no other high-resolution methods are available to compare to. However, CLIP targets are biased towards highly-expressed and long genes.
- ADAR needs adenosine to edit, may be bias towards false negatives

HyperTRIBE – a more active version of TRIBE from a mutagenesis screen

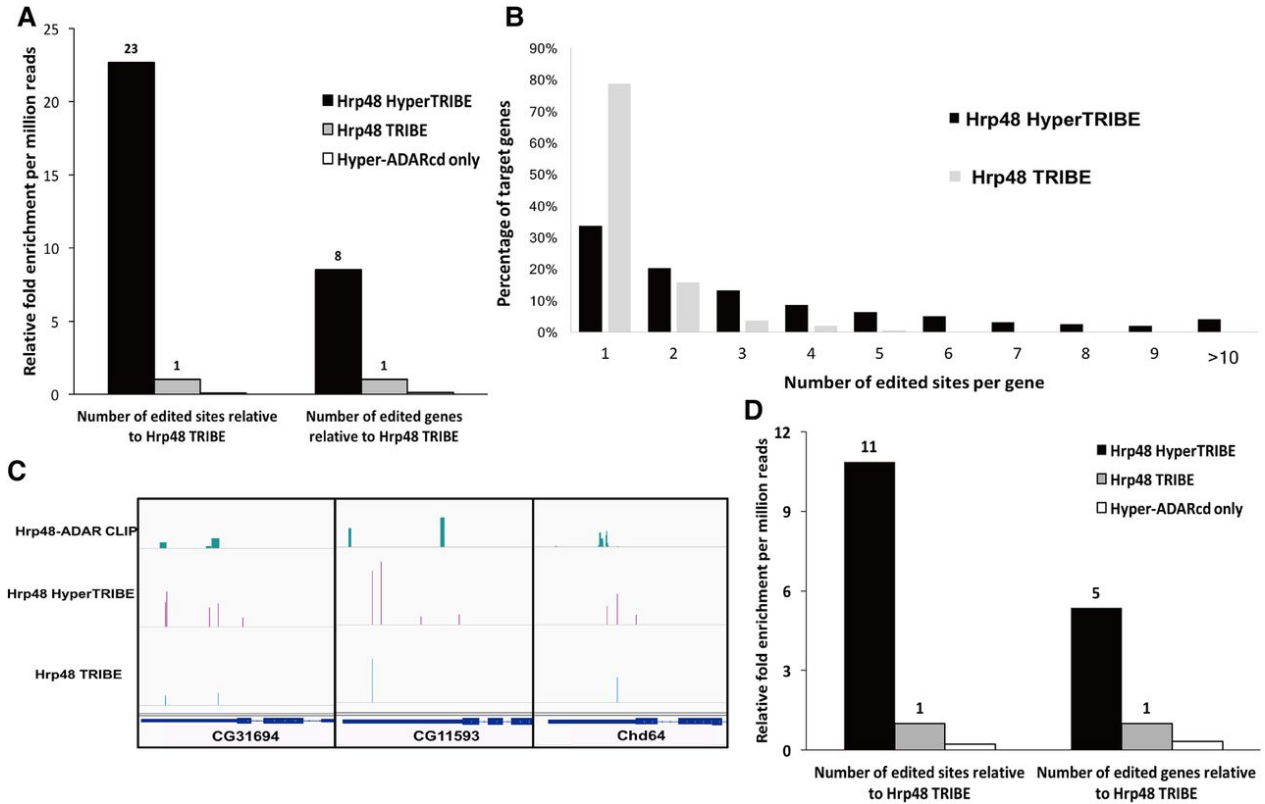
- ADARcd preferentially edits As after Us and before Gs (UAG), also As around double stranded regions are preferred
- Human mutagenesis found „hyperactive“ E488Q mutation in ADAR2
- E488 is highly conserved between human and drosophila
- Introduction of E488Q in ADARcd

Weijin Xu et al. RNA 2018;24:173-182

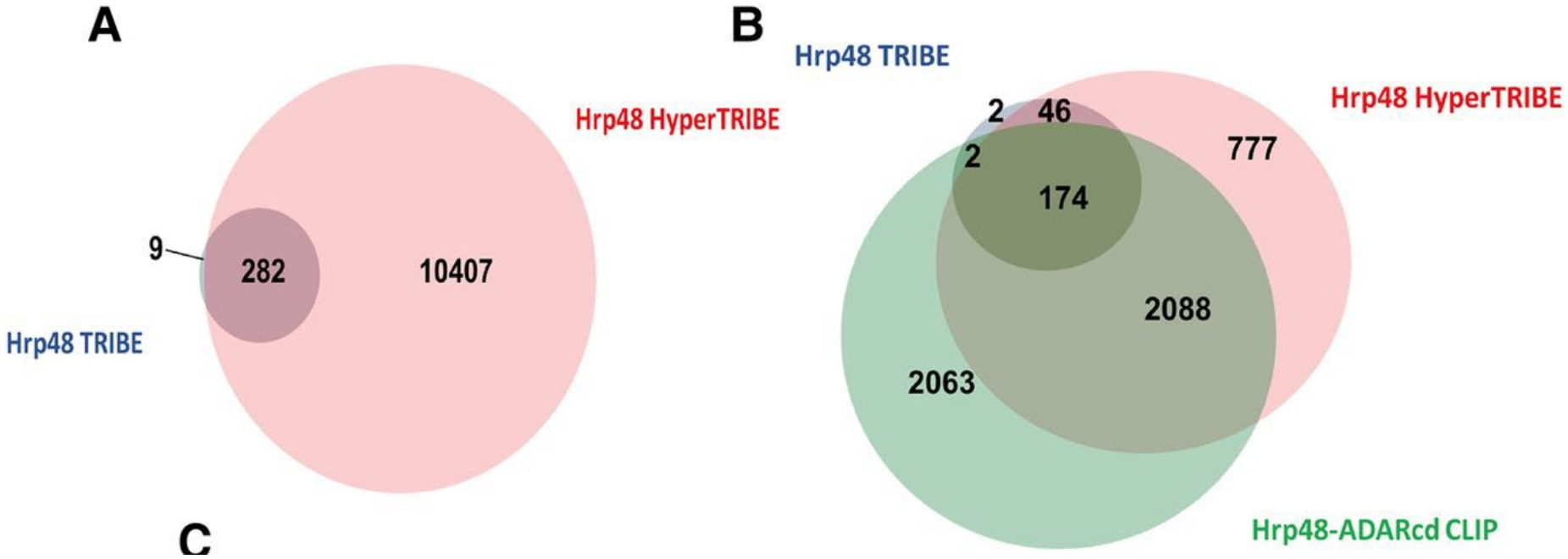
RNA

HyperTRIBE vs. TRIBE

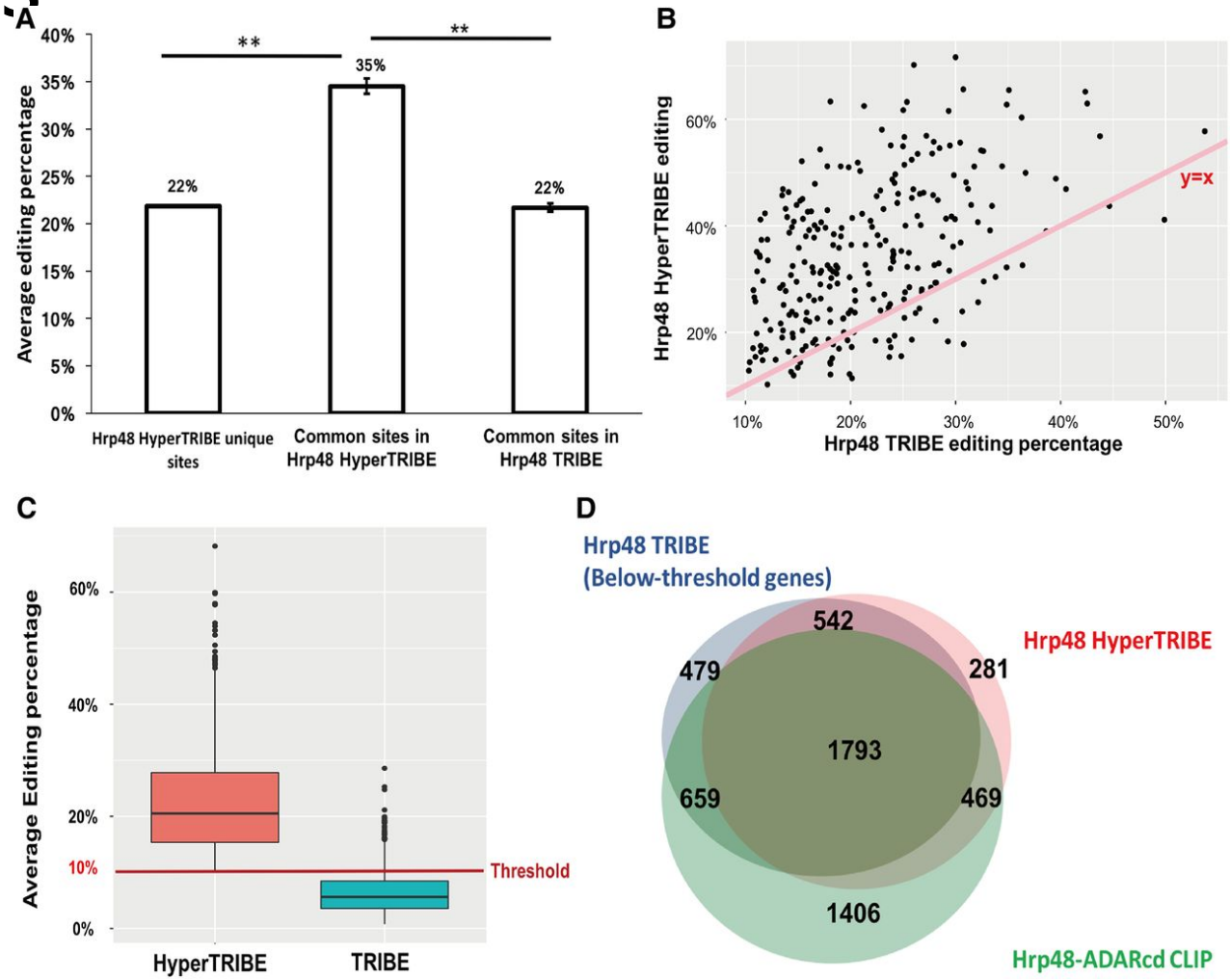
- HyperTRIBE results in ca. 20x higher editing activity compared to TRIBE, despite comparable expression levels



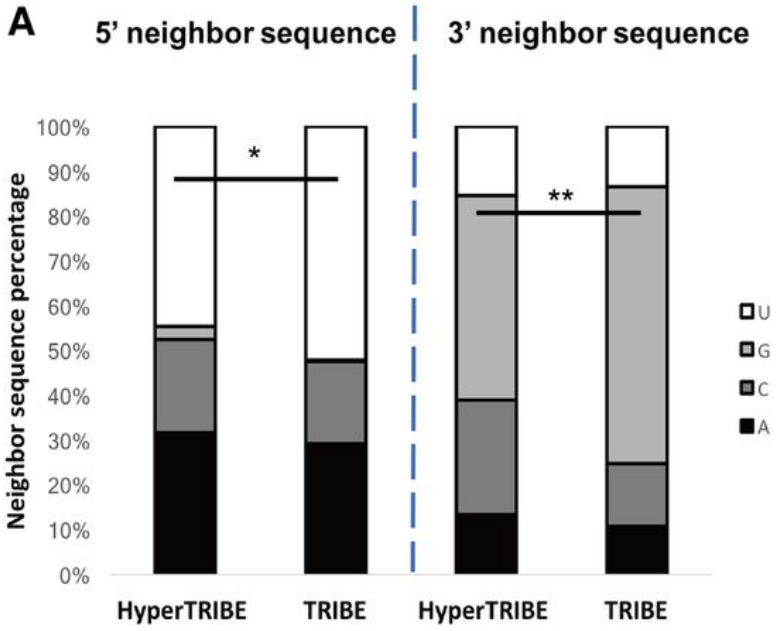
HyperTRIBE data faithfully reflect Hrp48 binding specificity with higher sensitivity than TRIBE

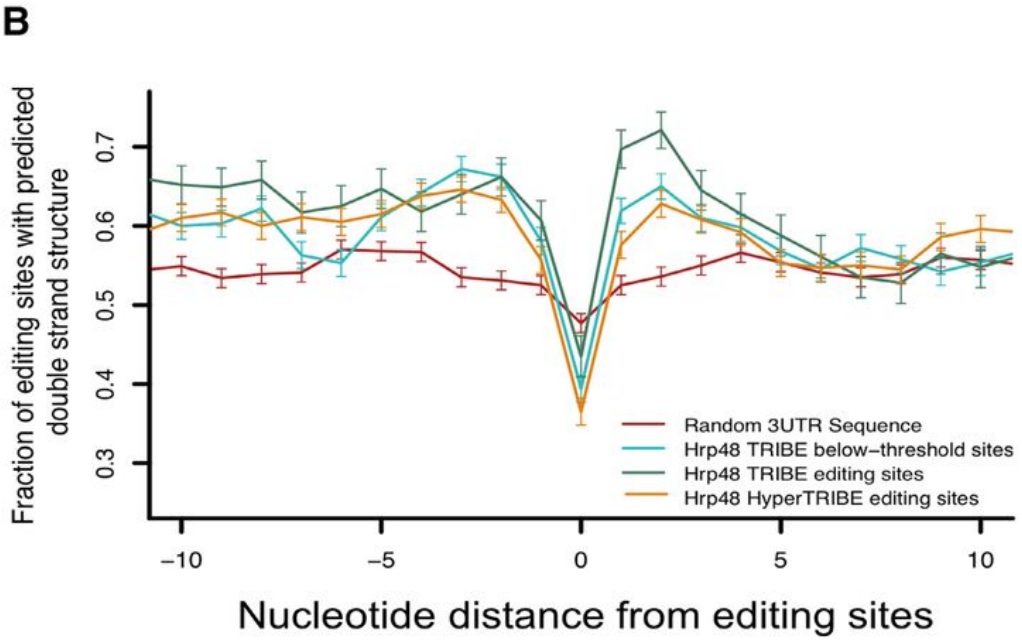


HyperTRIBE allows the detection of many below-threshold TRIBE sites

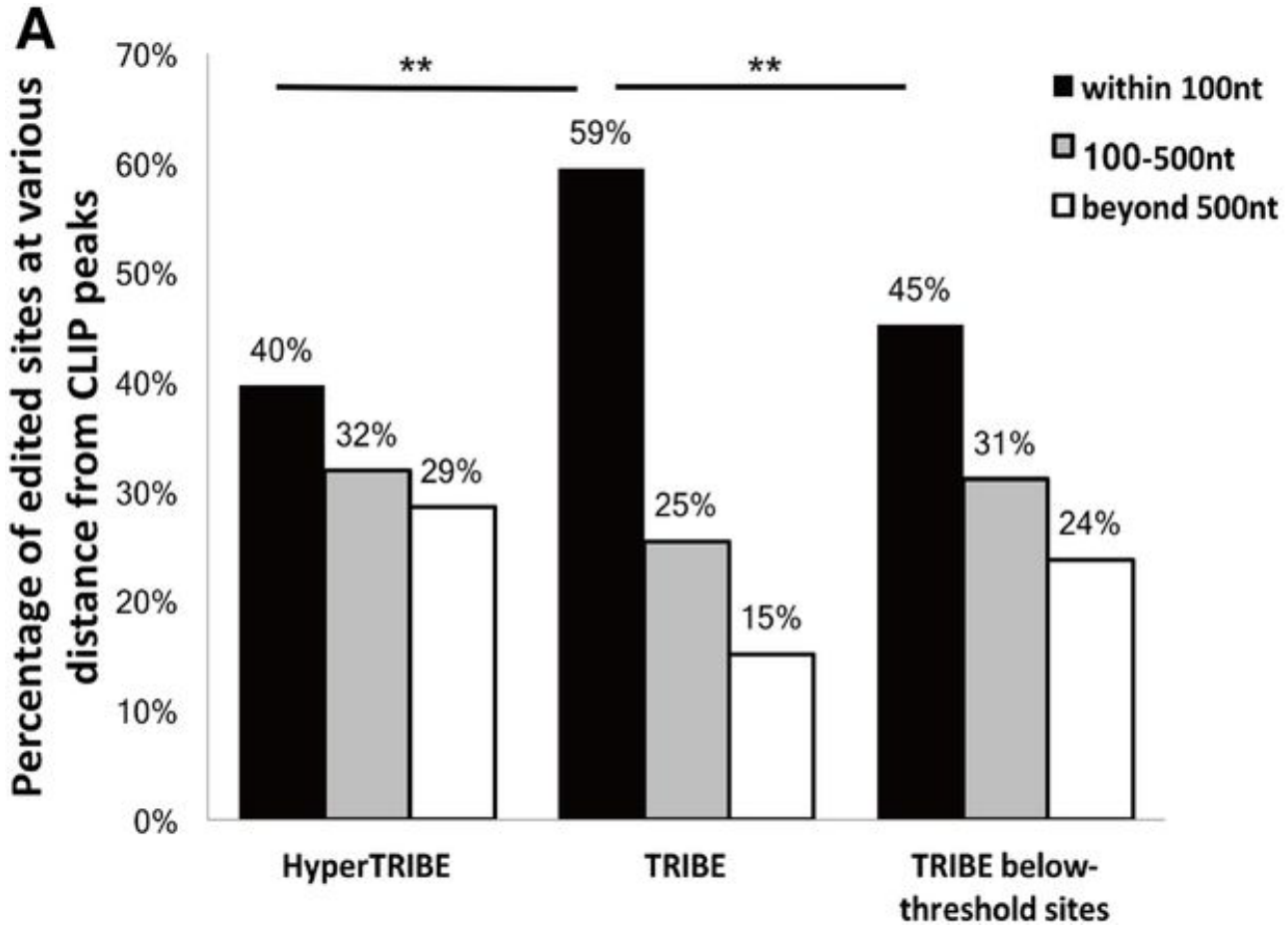


HyperTRIBE has less nearest neighbor sequence and double-stranded structure requirements than TRIBE.





HyperTRIBE edits distant adenosines more efficiently than TRIBE.



Conclusions

(i) HyperTRIBE edits more efficiently than TRIBE

(ii) even very low editing frequencies identify true TRIBE editing sites if they are identified in replicate experiments

(iii) there are many more bona fide TRIBE editing events than are identified with conservative thresholds

Nonetheless, the fraction of edited adenosines is low even for HyperTRIBE, presumably still reflecting the sequence and structural requirements of the ADARcd

DART-seq - deamination adjacent to RNA modification targets sequencing

Article | Published: 23 September 2019

DART-seq: an antibody-free method for global m^{6}A detection

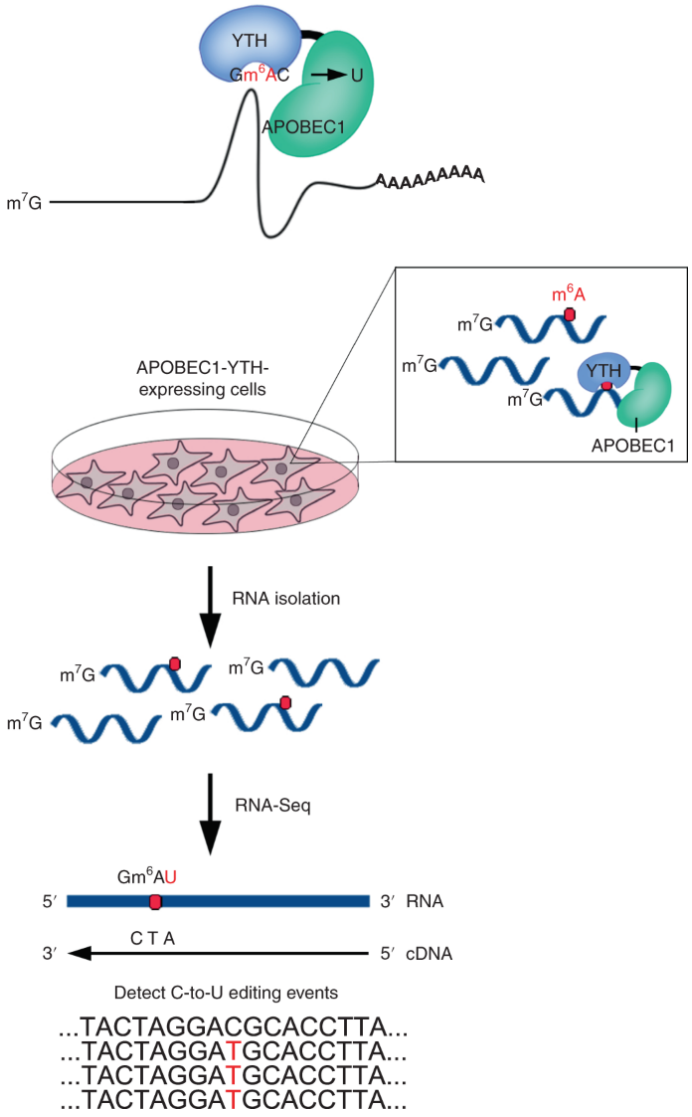
Kate D. Meyer

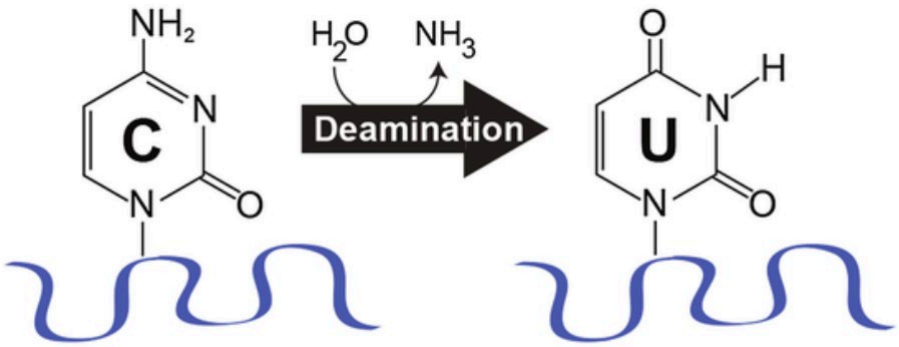
Nature Methods **16**, 1275–1280 (2019) | Cite this article

Rationale of the study

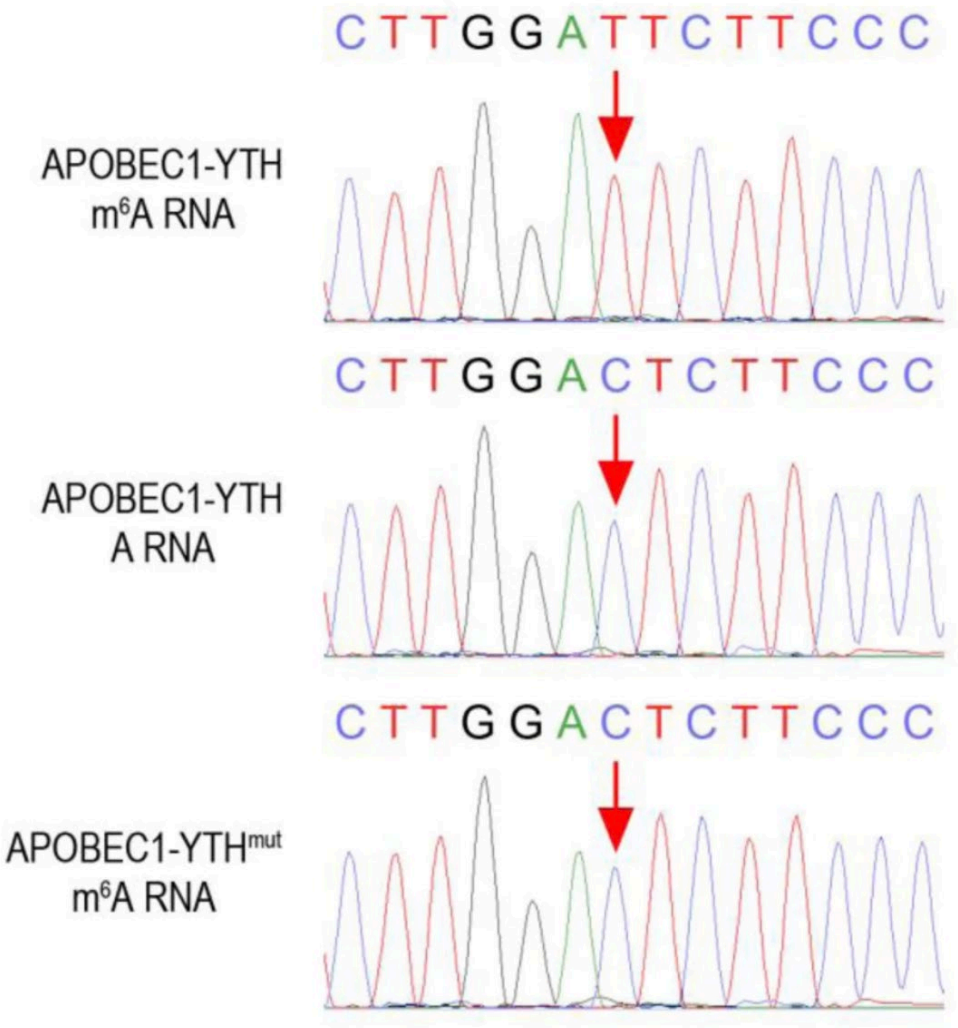
- N^6-methyladenosine (m^6A), plays diverse roles in RNA regulation
- m^6A is an important regulator of physiological processes such as post-transscriptional gene regulation
- Previously, antibody-based purification of m^6A was used (MeRIP-Seq or m^6A-seq), however, either these methods needed a lot of input RNA (150 ng) or they did not distinguish between m^6A and m^6Am
- Lastly, antibody-based approaches are more expensive than direct methods

Development of APOBEC1-YTH

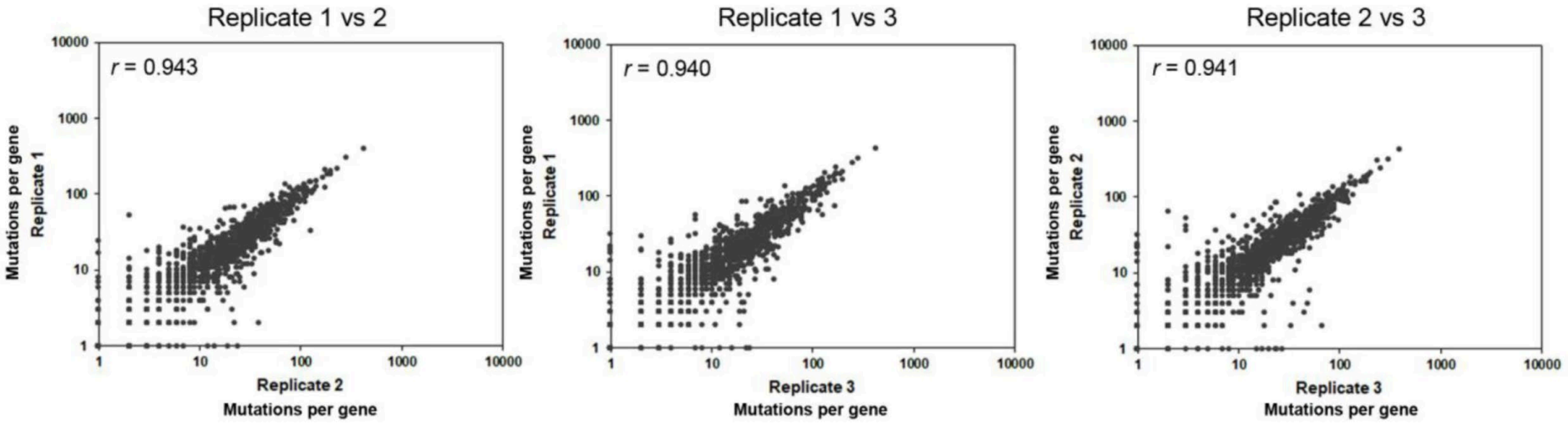




APOBEC1-YTH^{mut} serves as additional negative control

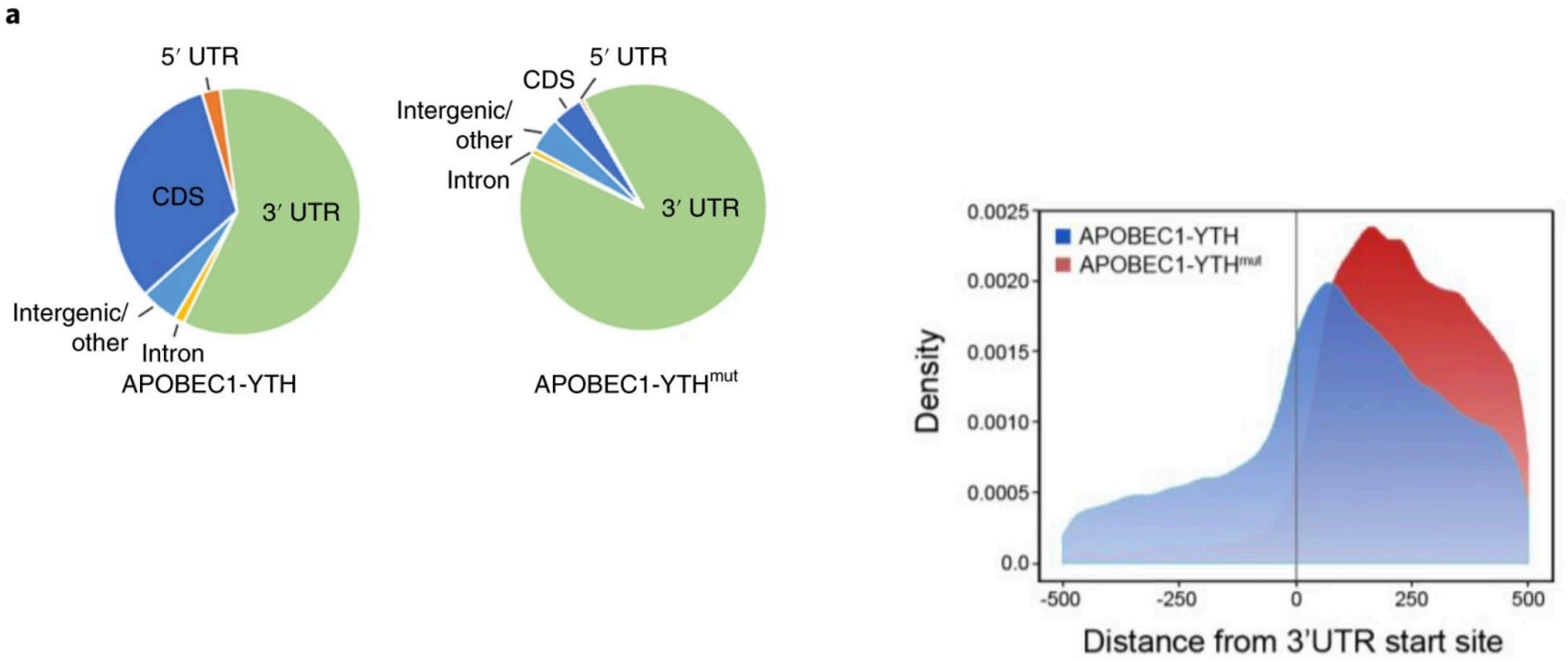


High degree of overlap between individual replicates



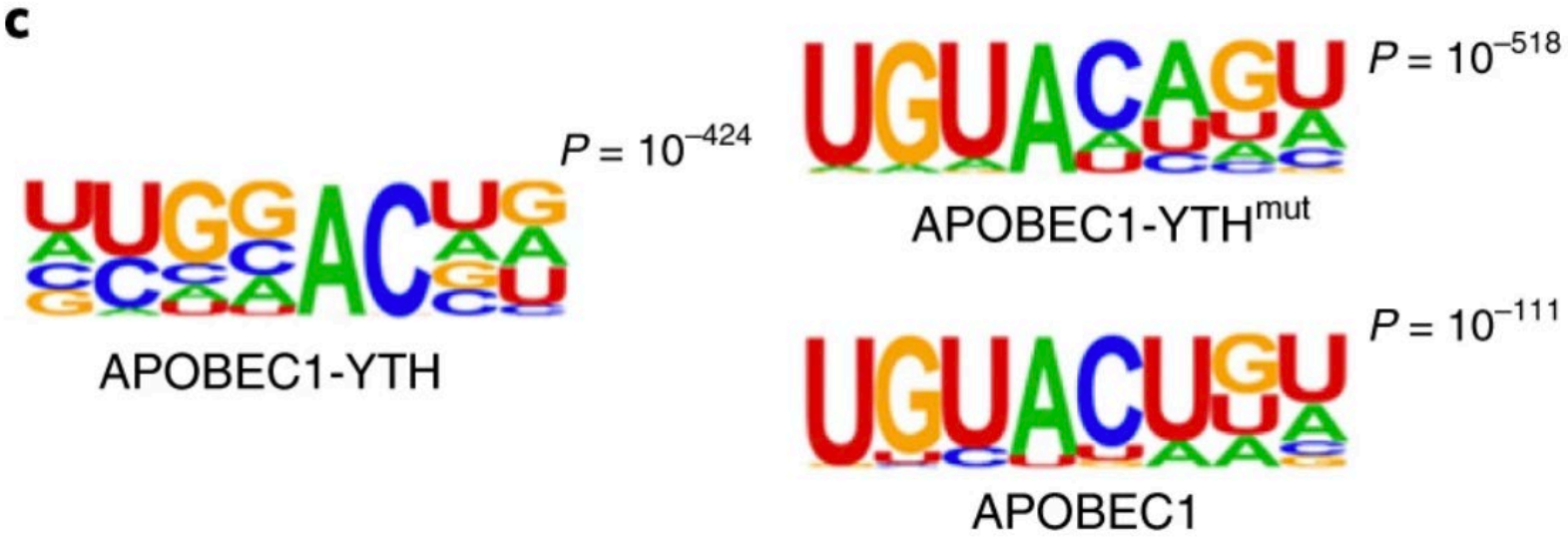
DART-seq identifies m^{6}A-containing RNAs transcriptome wide

- C>U editing sites



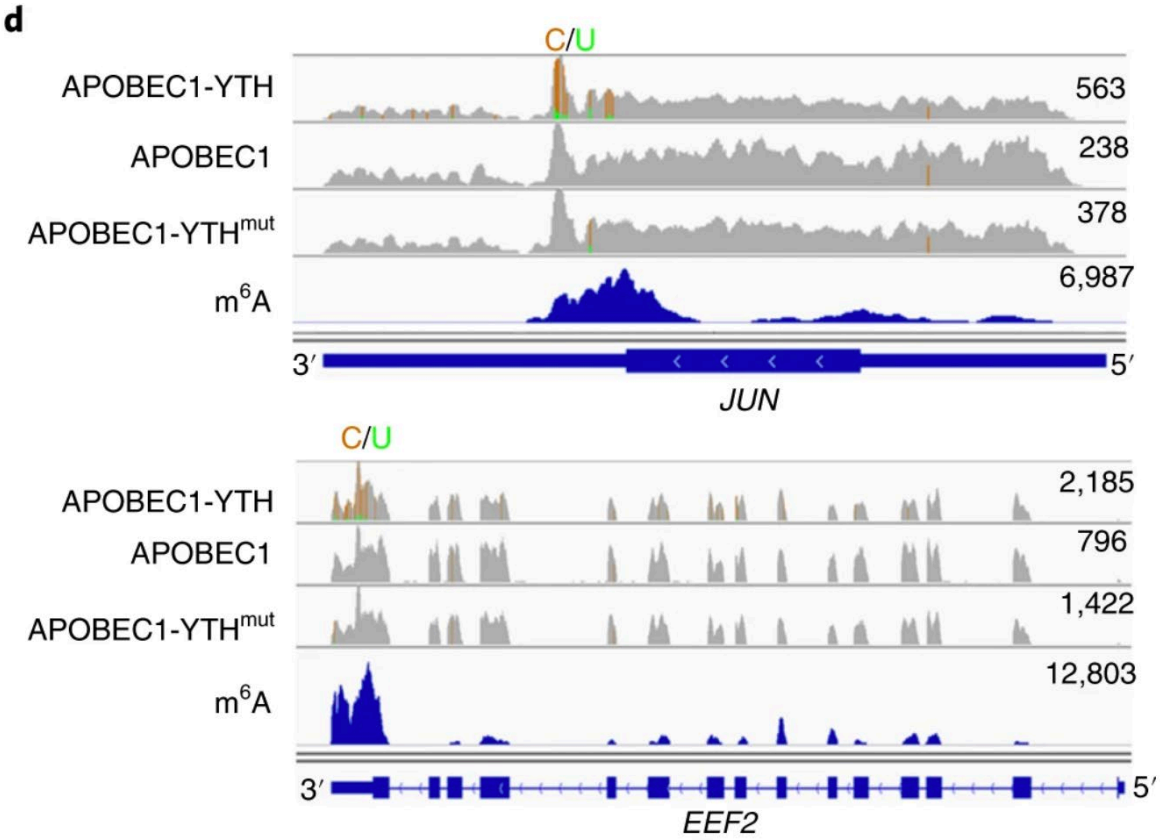
- Enrichment in 3'UTR, next to stop codons and long exons

GGACU is consensus motif of APOBEC1-YTH



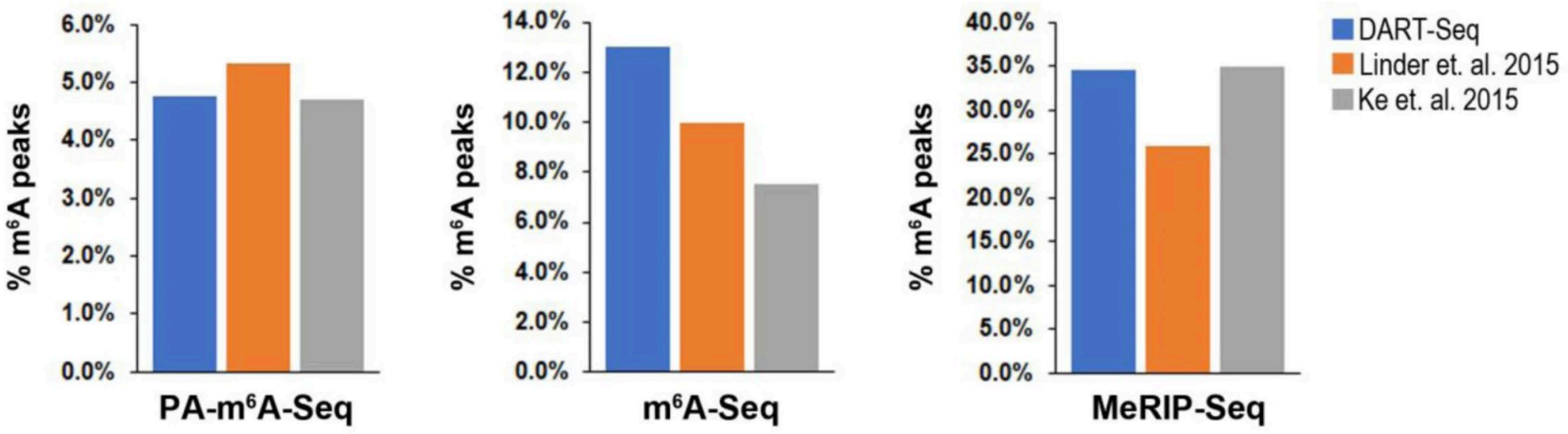
High overlap of DART-seq to MeRIP-seq

- 64% of m^{6}A-containing RNAs detected by DART-seq (3,679 of 5,768 RNAs)



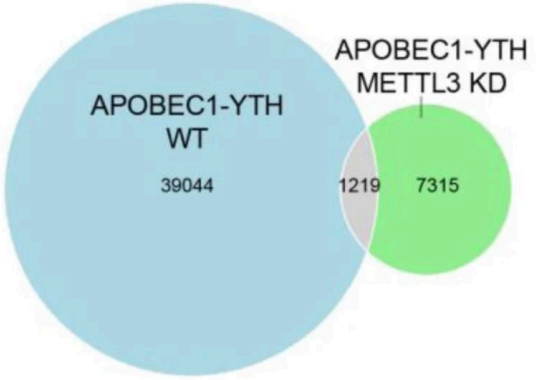
Comparision of single-nucleotide methods to immunoprecipitation-based methods shows high overlap

- PA-m^{6}A-seq, m^{6}A-Seq, MeRIP-Seq: IP
- Linder et al., Ke et al.: miCLIP



Validation of DART-seq

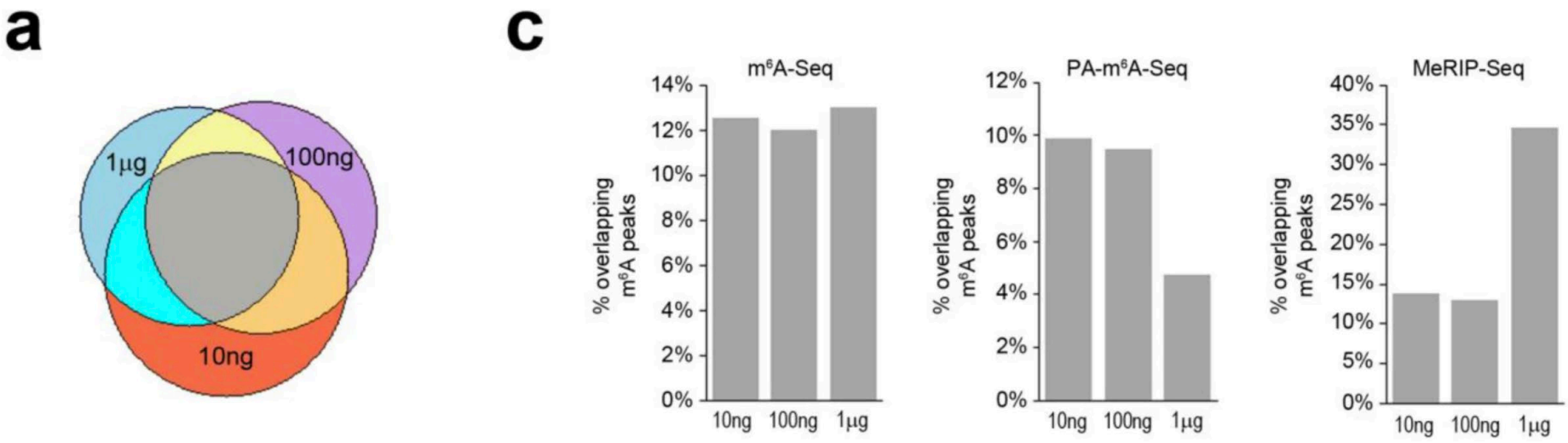
- Depletion of m^{6}A methyltransferase METTL3 should lead to fewer DART-seq events
- Loss of GGACU consensus site



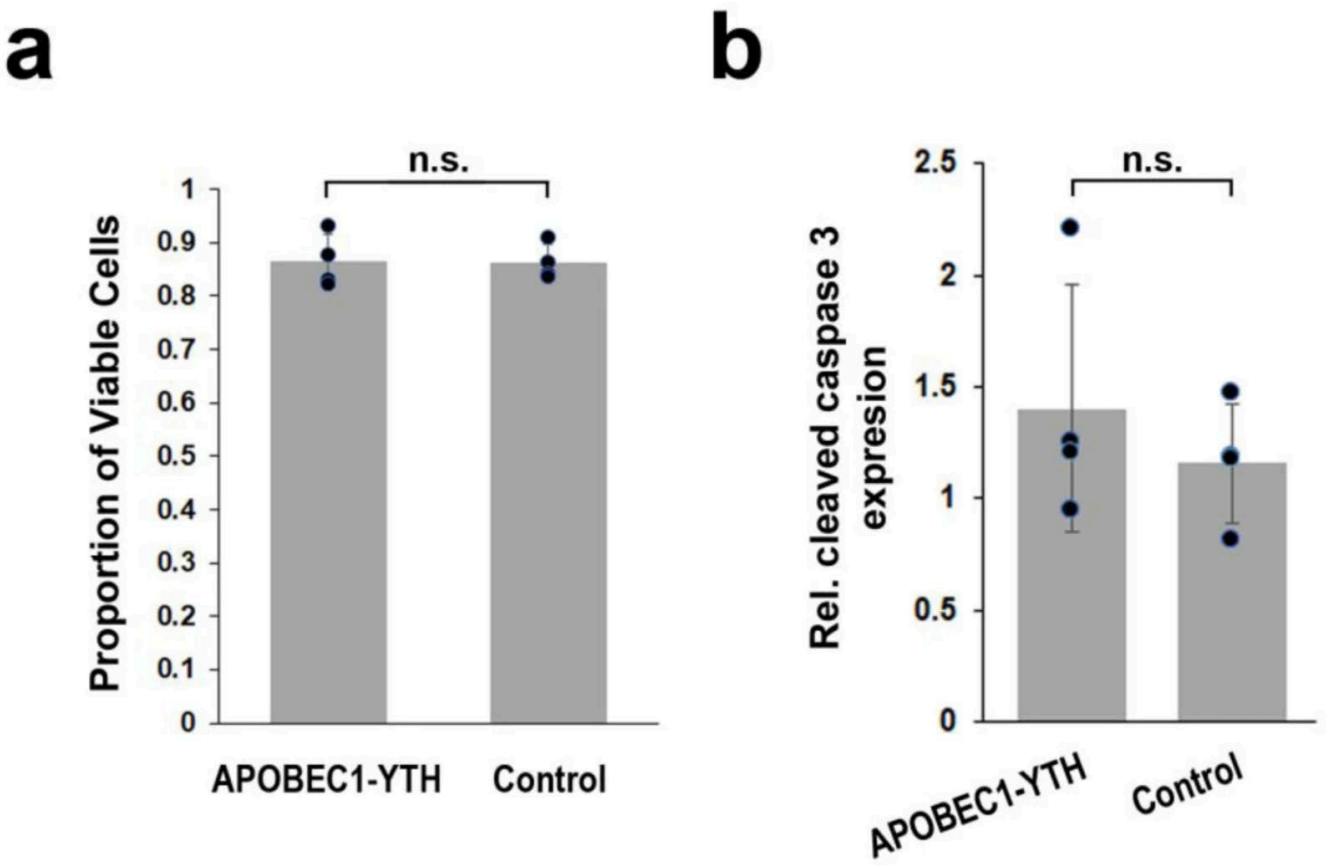
- 97% less m^{6}A sites detected

Low-input DART-seq

- Current limit for RNA input is around 150ng, 10 ng of RNA leads to detection of > 79% of high-input DART-seq sites

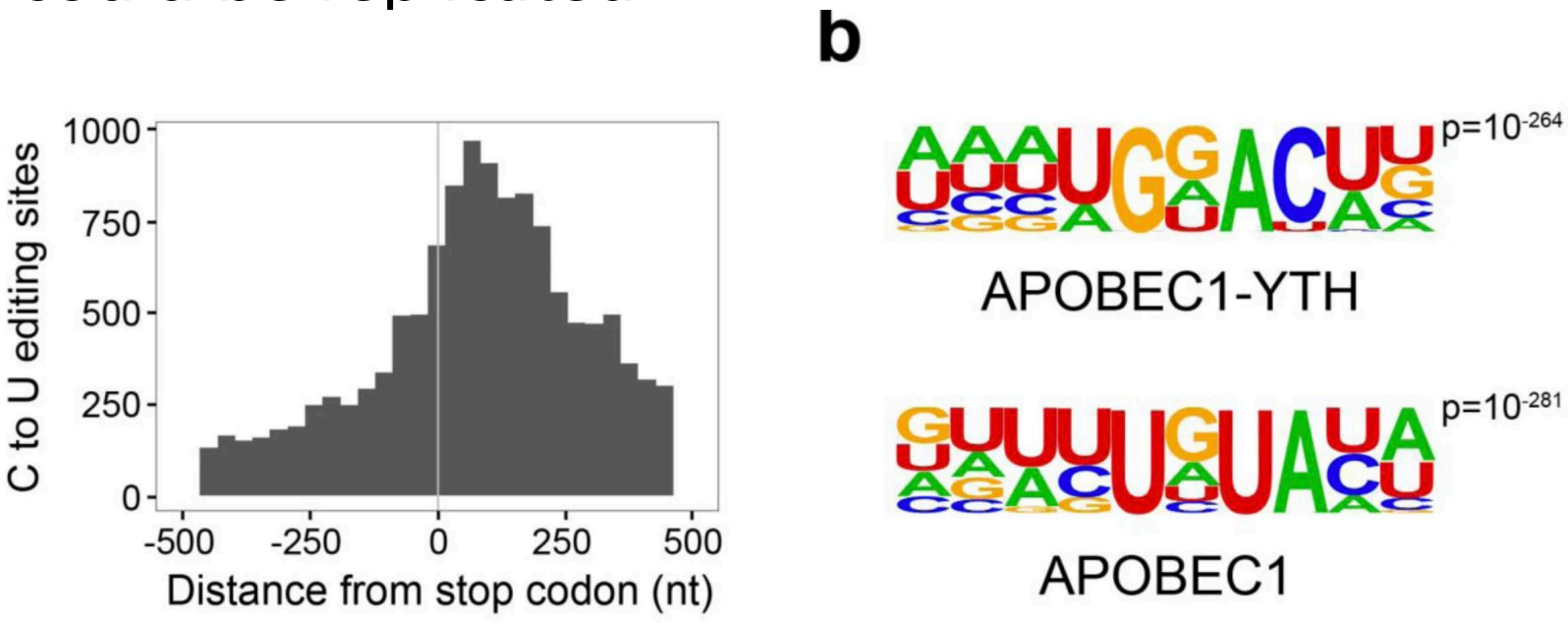


In vitro DART-seq



In vitro DART-seq

- In vitro transcribed/translated using the Promega TNT T7 Quick Coupled In Vitro Transcription/Translation kit.
- Majority, e.g. 91%, of cellular DART-seq sites could be replicated

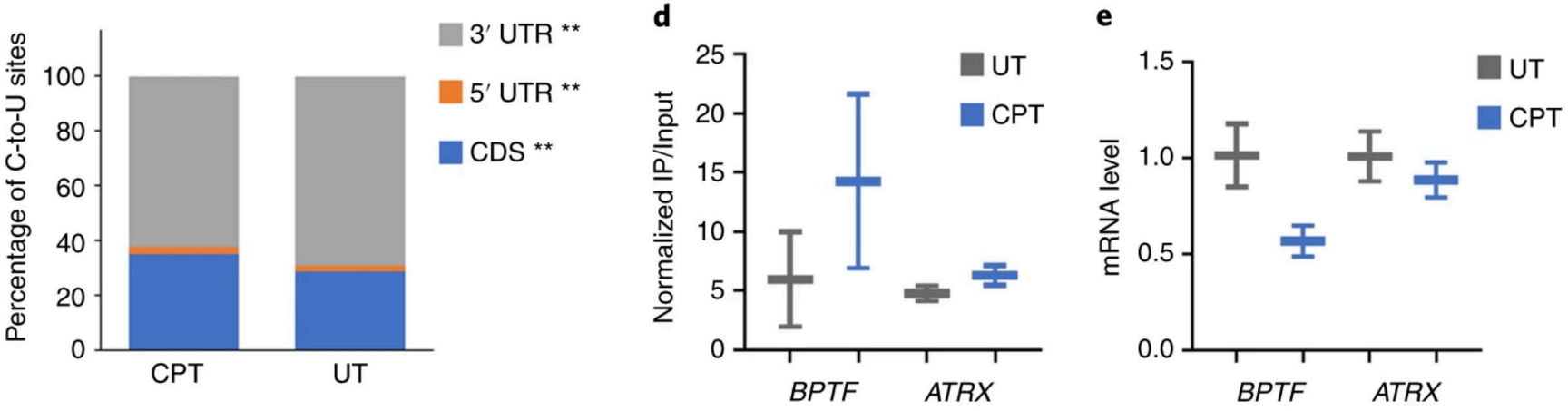


DART-seq distinguishes m^6A from m^6Am

- IP based approaches cannot distinguish m^6A from m^6Am and m^6Am is not invariably followed by a cytidine
- Only 1 RNA found after comparison of 3,431 sites with overlap between DART-seq sites and m^6Am sites
- However, this site was diminished in METTL3 KD cells, which makes it more likely an m^6A site

DART-seq identifies m^{6}A accumulation in cellular RNAs

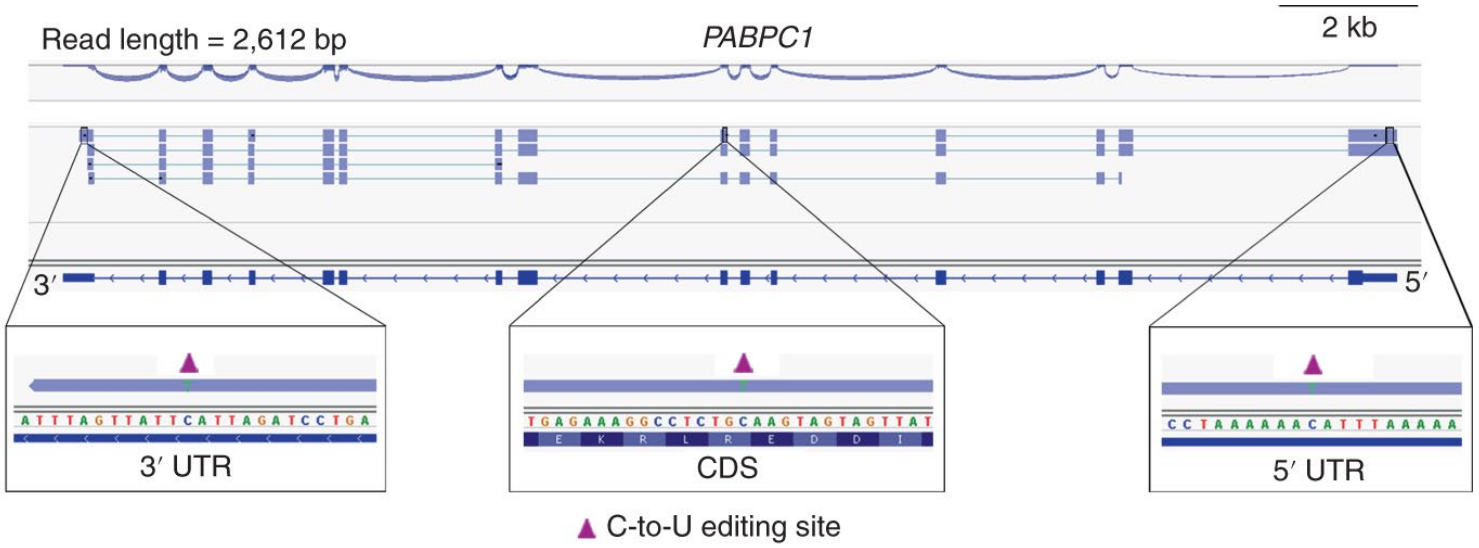
- Camptothecin (CPT) causes slowed transcription and an increase in m^{6}A abundance in the CDS
- Treatment of HEK293T cells with CPT for 5 h, followed by DART-seq
- 6,258 C-to-U sites that showed at least a

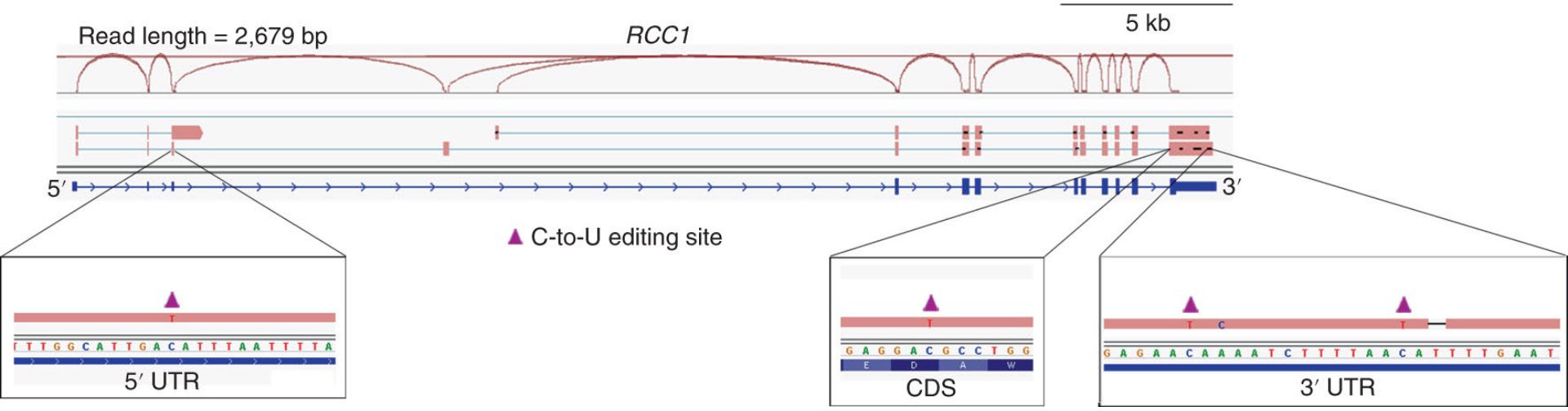


Long-read DART-seq

- Multiply methylated RNAs were described although it was unclear whether it was on the same or on different molecules
- Some transcripts exhibit isoform-specific regional editing, others contain DART-seq sites in the 5′ UTR, CDS and 3′ UTR
- In addition, 41% of reads spanning at least two editing sites contain two or more C-to-U editing events.
- Conclusion: most molecules harbor on m^6A site but many reads harbor multiple sites

Long-read DART-seq





Conclusion

- DART-seq works with low input amounts, down to 10 ng of RNA
- DART-seq detects more sites than previously used IP approaches
- DART-seq is dynamic, sets permanent mark on RNAs over multiple hours

- In vitro DART-seq lacks sensitivity to detect low abundance sites

- Outlook: compartmental, e.g. mitochondrial, nuclear, cytoplasmic DART-seq

Genome-wide quantification of ADAR adenosine-to-inosine RNA editing activity

- A-to-I editing is a common RNA modification, modifying false activation of immune response by endogenous dsRNA strands
- Methods to quantify global A-to-I editing index (AEI) are currently lacking and AEI is determined by differential editing of all transcripts between multiple treatment groups

Article | Published: 21 October 2019

Genome-wide quantification of ADAR adenosine-to-inosine RNA editing activity

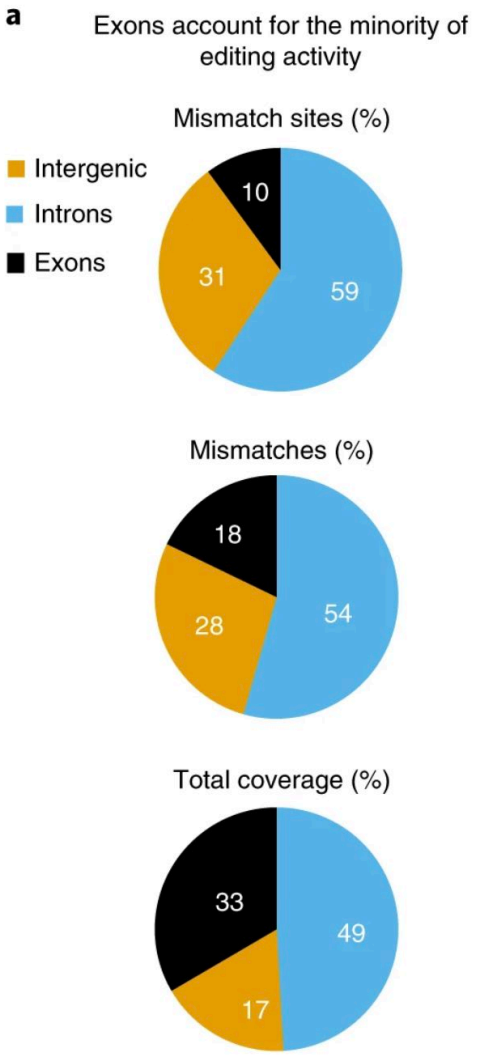
Shalom Hillel Roth, Erez Y. Levanon & Eli Eisenberg

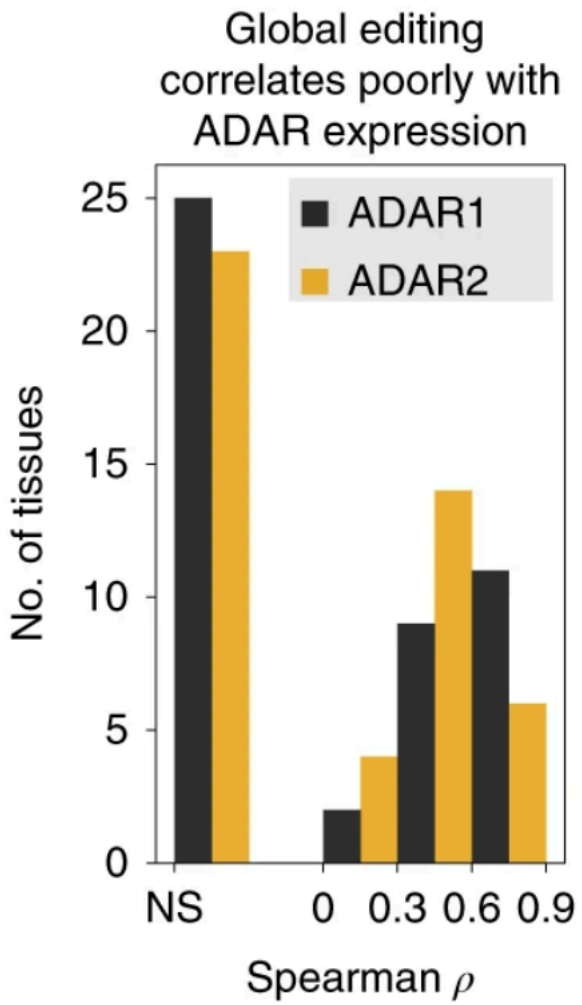
Nature Methods **16**, 1131–1138 (2019) | Cite this article

Caveats of current A-to-I editing measurements

- Most editing activities occur in Alu regions, which span ~ 11% of the human genome
- Thousands of Alu repeats may be edited heavily (> 10 %) but most of Alu editing occurs in millions of weakly edited sites
- Median editing in Alu sites is ~ 0.5%, e.g. only 1/200 reads, on average, exhibit editing at the typical site
- Reliable identification and quantification of such low levels of editing per-site requires coverage of hundreds of reads, which far extends the sequencing depth offered by standard RNA-seq datasets

Exons contribute poorly to total RNA editing



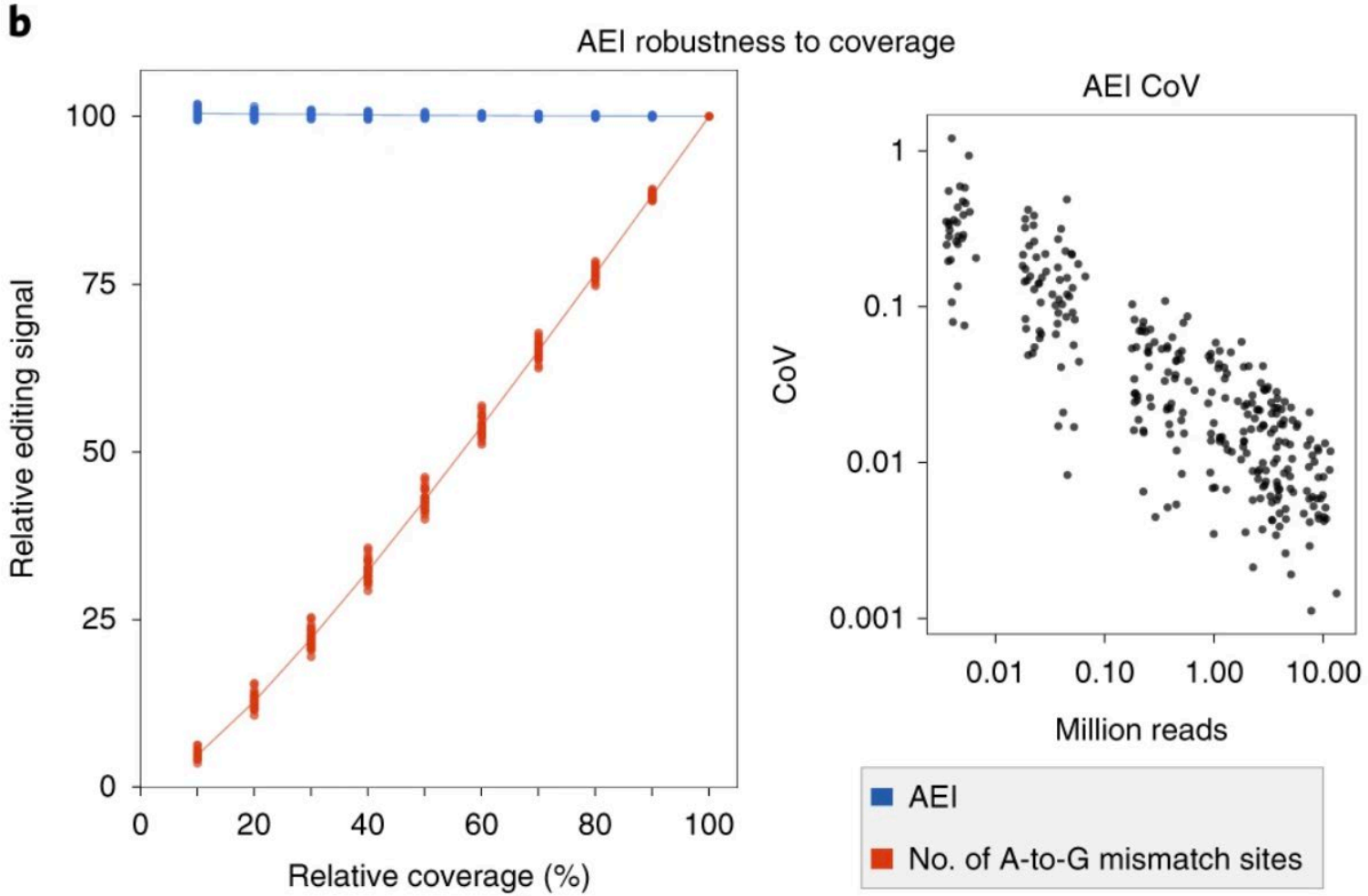


FDR>0.05 for all 47 samples

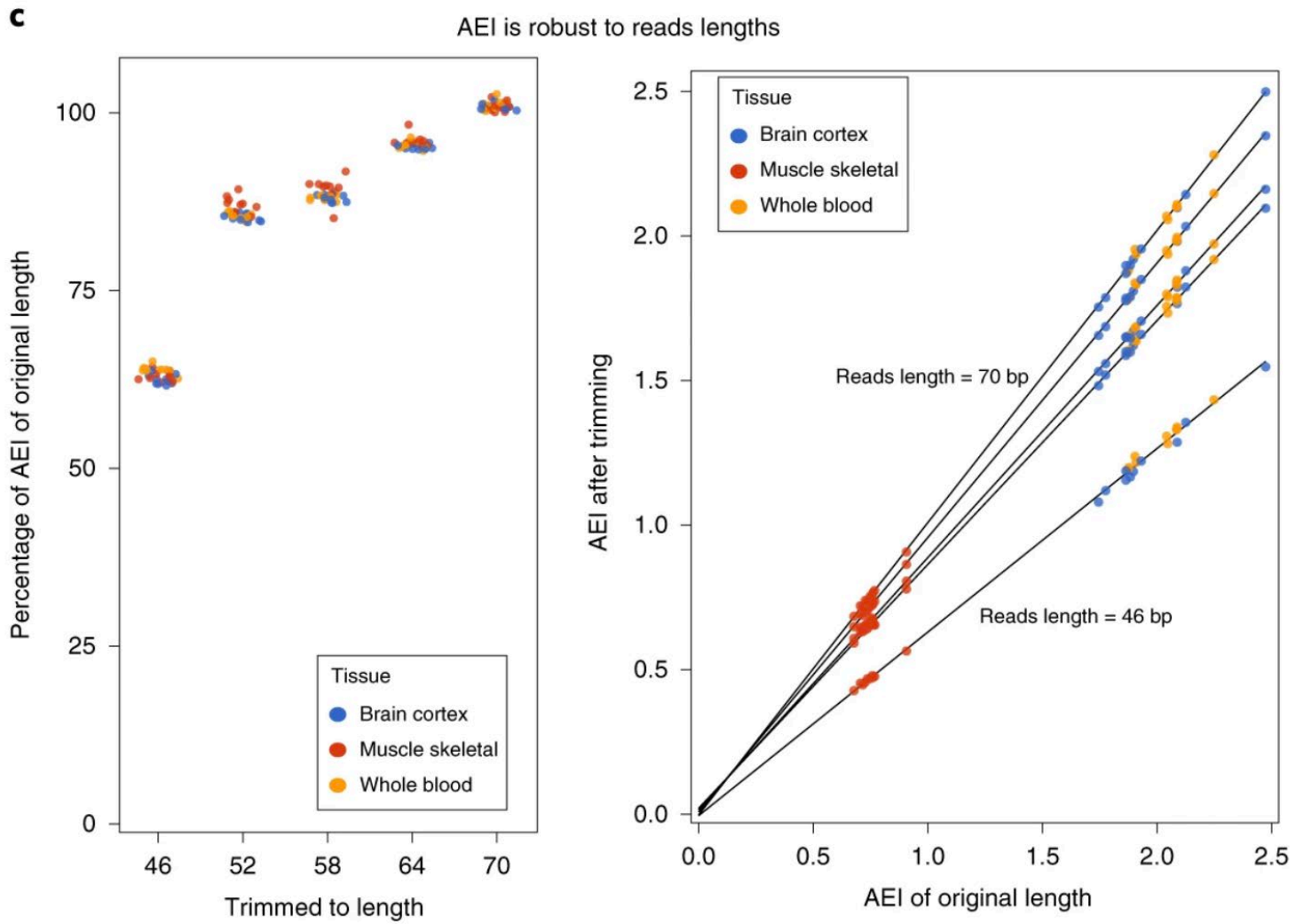
Editing across predefined regions poorly represents global editing levels

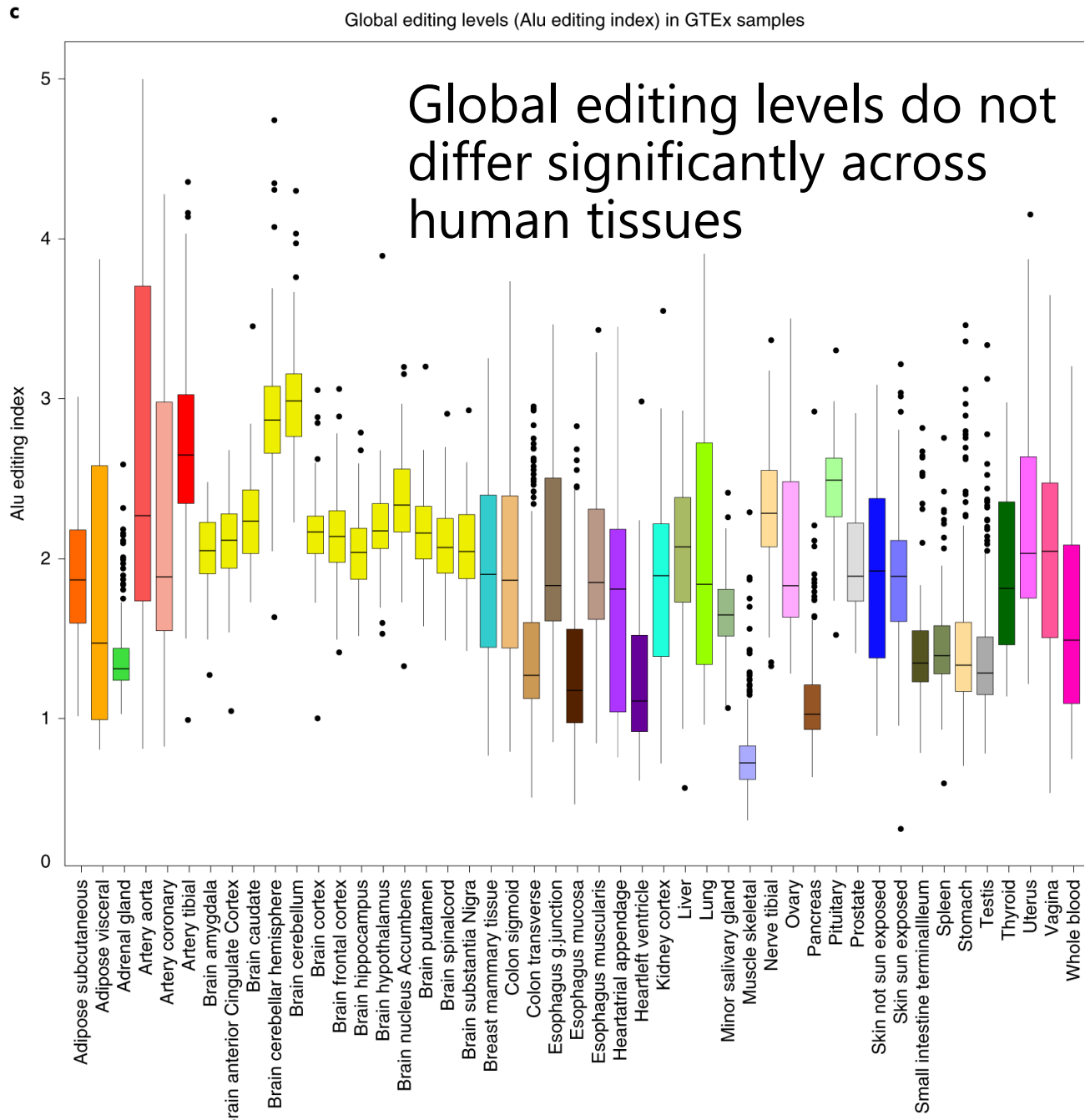
- Predefined sites are used for comparison across samples, however, datasets show either lots of false positives or are outdated
- Furthermore, these sites contribute less than 1% to global editing
- Authors propose an A-to-I editing index (AEI) focussing only on highly edited Alu repeats, which are comparable between samples

AEI is robust against variable coverage

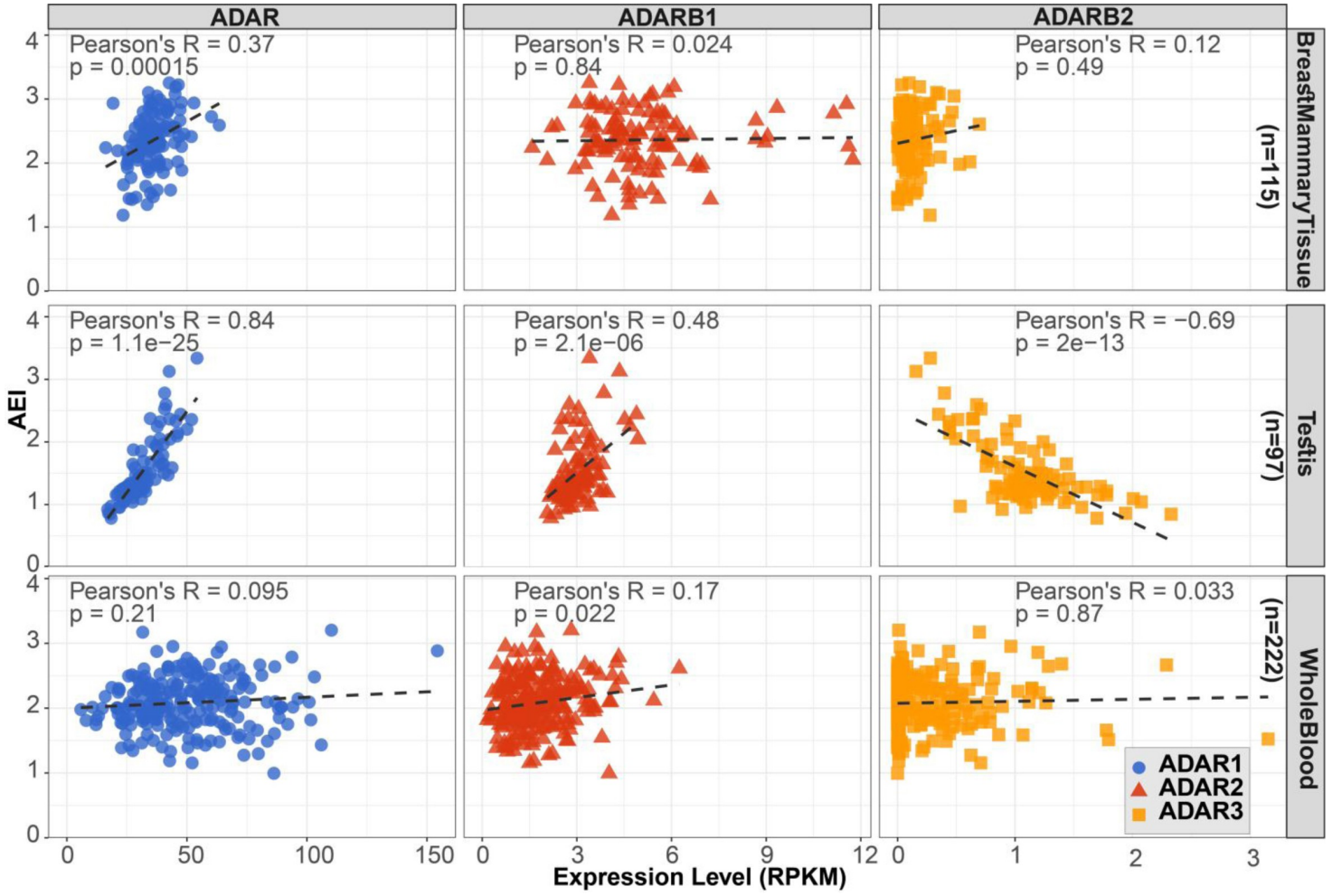


AEI is sensitive to reads' length (left), but the results for different lengths are highly correlated

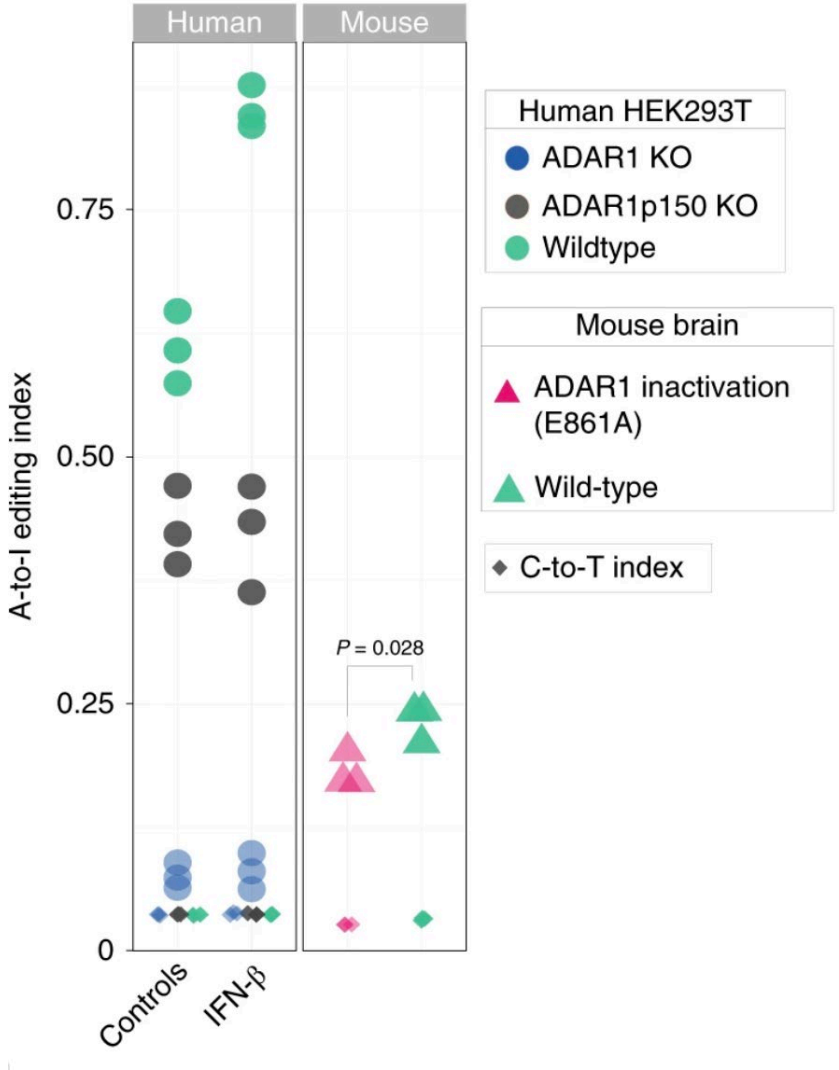



c
Global editing levels (Alu editing index) in GTEx samples
Global editing levels do not differ significantly across human tissues
Alu editing index
0
1
2
3
4
5
Adipose subcutaneous
Adipose visceral
Adrenal gland
Artery aorta
Artery coronary
Artery tibial
Brain amygdala
Brain anterior Cingulate Cortex
Brain caudate
Brain cerebellar hemisphere
Brain cerebellum
Brain cortex
Brain frontal cortex
Brain hippocampus
Brain hypothalamus
Brain nucleus Accumbens
Brain putamen
Brain spinalcord
Brain substantia Nigra
Breast mammary tissue
Colon sigmoid
Colon transverse
Esophagus g.junction
Esophagus mucosa
Esophagus muscularis
Heartatrial appendage
Heartleft ventricle
Kidney cortex
Liver
Lung
Minor salivary gland
Muscle skeletal
Nerve tibial
Ovary
Pancreas
Pituitary
Prostate
Skin not sun exposed
Skin sun exposed
Small intestine terminalIleum
Spleen
Stomach
Testis
Thyroid
Uterus
Vagina
Whole blood

ADAR enzyme activity, as measured by RNA levels, does not correlate with AEI

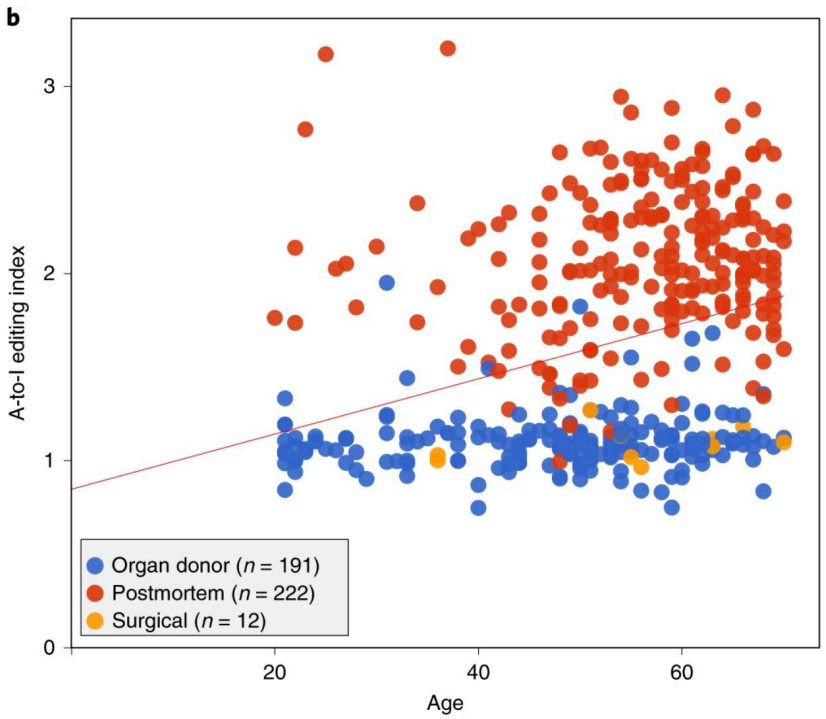


AEI recapituales ADAR1-induced RNA editing activity



Pitfalls comparing different samples

- Frozen > FFPE samples
- RiboZero > poly-A sequencing
- Origin of the sample, e.g. post-mortem, surgical, organ donor



Conclusions

- AEI is an easy to implement and robust marker of RNA editing
- Downside: authors claim false results from other methods, but may be more examples would have been nice