Signature of base editor RNA edits in the transcriptome and reduction of exogenous RNA editing

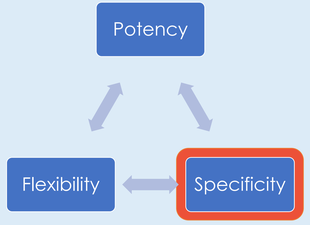
Isaiah Taylor, Lisle Mose, Joel Parker, Megan Williamson, Jake Deslauriers, Adrian Oliver, Aaron Crain, Mike Lewis, Colin Lim, and Allie Crawley

ElevateBio Life Edit, Durham, North Carolina, USA

life edit® an elevatebio company

Abstract

Genomic Medicines



The deaminases used for A-base editors are widely evolved from proteins that recognize and deaminate tRNAs. Activity at the DNA level is a feature that must be engineered into these proteins. Quantification of **RNA transcriptome deamination** by evolved deaminases remains difficult due to high rates of naturally occurring, endogenous RNA editing. Here we present a method that can **separate the signature** of RNA edits made by base editors from endogenous RNA editing. Using mammalian cell systems including primary cells, we demonstrate that RNA editing from base editors is **dose-dependent and transient**. This research can improve genomic medicines by standardizing a way to measure safety and specificity of base editors.

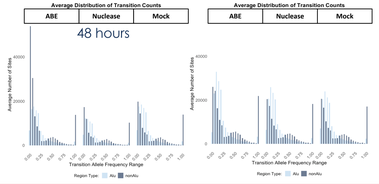
Separation of endogenous and base editor-derived RNA edits

Endogenous RNA editing

Endogenous RNA editing occurs at millions of known sites in the human transcriptome and is largely a result of ADAR family proteins1. ADAR proteins catalyze A-to-I edits in primarily dsRNA sequences2. Profiles of ADAR editing locations by tissue type has identified that the majority of ADAR edited sites are intronic within Alu retrotransposon mobile elements2,3,4,5. Alu elements are members of the SINE mobile element family and influence cellular processes including gene expression, gene splicing, polyadenylation, and ADAR editing6. Using the features of known ADAR editing, we can exclude RNA editing events likely to be caused by ADAR to enrich for sites that are edited due to exogenous deamination events.

A to G Transcriptomic substitutions allele frequency distribution in Alu vs non-Alu regions

- A to G substitutions are endogenous at Alu and non-Alu regions
- Base editors increase low AF A>G substitutions at 48 hrs
- Peak in A>G substitutions decreases at 96 hrs
- Signature is more pronounced in non-Alu regions



Metrics for quantifying RNA editing

Transcriptome Proportion

The "Transcriptome Proportion" metric was developed to describe the amount of base editor derived RNA editing present in cells. The transcriptome-wide RNA-editing is calculated by summing the read counts per position that support an A>G or T>C transition and dividing by total coverage across all positions. In effect, this outputs the proportion of the entire transcriptome exhibiting A>G or T>C transitions. This metric is simple to define but includes "all-cause" transitions: RNA-editing, alignment errors, sequencing errors, and genetic variation. To customize this metric to count for percent of transcriptome edits matching the signature for adenosine base editors, only A>G or T>C substitutions in non-Alu sites are used as the signal for the "Transcriptome Proportion" metric.

$$\text{Transcriptome Proportion} = \frac{\text{ABE-transition } (A \rightarrow G \text{ or } T \rightarrow C) \text{ base counts}}{\text{total base counts}}$$

In this example we assume a 12 base transcriptome sequenced at 8x coverage.

Reference: **AGCGTAGCTAGA**

Alignments:
AGCGTAGCTAGA
AGTGTAGCTAGA
AGCGTAGCTAGA
AGCGTAGCTAGA
AGCGTAGCTAGA
AGCGTAGCTAGA
AGCGTAGCTAGA

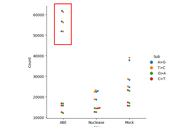
Transcriptome Proportion calculation:
We sum the two A>G transitions. We ignore the C>T transition. We then calculate total coverage (12 bases * 8 reads = 96).
transcriptome proportion = 2/96 = .02083

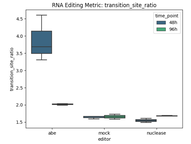
Transition Site Ratio

The "Site Ratio" metric counts the number of high-confidence sites with evidence of RNA editing (ABE A>G or T>C transitions). "High-confidence" sites are defined as follows: transitions observed at least twice, in sites covered at a depth of at least 10 reads, at a frequency of at least 0.01 (1%). First, the number of sites with non-expected ABE editing (G>A or C>T) in the allele frequency range 1-25% in non-Alu sites are counted as control counts to account for background substitution and error rates. Second, the number of sites with expected ABE editing (A>G or T>C) in the allele frequency range 1-25% in non-Alu sites are counted as the signal of interest. The "Site Ratio" for RNA editing signal divides the expected transition counts by the unexpected transition counts. This normalization is intended to mitigate the effect of sample-specific variable sequencing depth and variable background sequencing error rates.

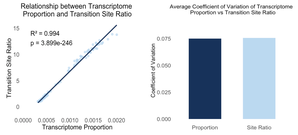
$$\text{Transition Site Ratio} = \frac{\#\text{ expected ABE-transition sites}}{\#\text{ non-expected ABE-transition sites}}$$

Count of transitions in non-Alu regions, in AF 1-25%





Comparison of Transcriptome Proportion and Transition Site Ratio Metrics

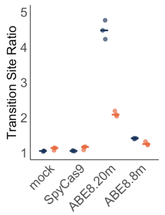


Key takeaway: The two metrics are highly correlated and possess similar levels of variability

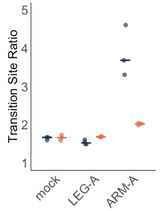
	"Transcriptome Proportion"	"Site Ratio"
Definition	Calculation of proportion of transcriptome edited by ABEs	Calculated by number of high-confidence sites in transcriptome edited by ABEs divided by background rate transitions
Advantage	• Direct biological interpretation of the metric • Less sensitive to low read counts	• Higher magnitude of signal allows more sensitive separation of measurements • Allele-frequency bound to enrich for ABE signal
Disadvantage	• Mainly driven by high expressed genes • Lower magnitude of signal range in metric	• Direct biological interpretation more difficult

Base editor RNA edits are transient

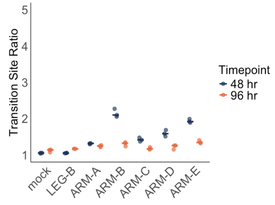
SpyCas9-based systems show transient RNA editing



LEG-A-based systems show transient RNA editing



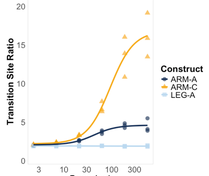
LEG-B-based systems show transient RNA editing



- Cells were harvested at 48 hrs and 96 hrs
- The highest RNA editing is present in ABE treated samples at 48 hrs
- RNA editing decreases in the ABE treated samples from 48 hrs to 96 hrs
- Detectable RNA editing still present with some deaminases at 96 hrs

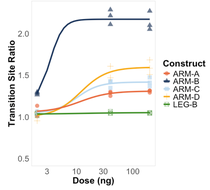
Base editor RNA edits are dose- and system- dependent

Dose-dependent RNA editing observed in a 6-point dose curve with LEG-A-based constructs



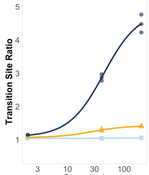
- RNA editing levels become saturated
- ARM-C has higher RNA editing than ARM-A
- At the low dose, RNA editing is not statistically higher than mock levels

Dose-dependent RNA editing observed in a 3-point dose curve with LEG-B-based constructs



- ARM-C shows higher RNA editing rates than ARM-A on the LEG-B construct
- The entire base editor, and not just the deaminase, contributes to RNA editing

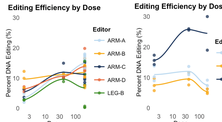
Dose-dependent RNA editing observed in a 3-point dose curve with SpyCas9-based constructs



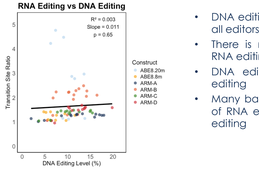
- ABE8.20 shows higher RNA editing than ABE8.8, which is consistent with literature

Magnitude of RNA editing is independent of DNA editing

DNA editing is dose-dependent



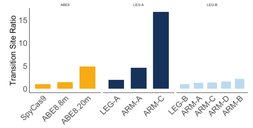
RNA editing is independent of DNA editing



- DNA editing rates increased with dose for all editors tested
- There is no global correlation between RNA editing and DNA editing rates
- DNA editing is not predictive of RNA editing
- Many base editors maintain similar levels of RNA editing despite increases in DNA editing

Deaminase variants have different RNA editing burdens

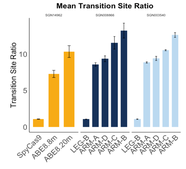
Saturation point can be used to rank RNA editing burden



- The same deaminase can have different RNA editing saturation points when fused to different nickases

Guide and deaminase affect RNA editing burden

- Plasmid delivery in HEK293T
- Two guides were used to identify guide-specific effects on RNA editing
- ARM-B showed higher RNA editing with SGN008866 compared to SGN003540 despite very similar transcript levels



Conclusion

- Life Edit has developed a metric for measuring base editor derived transcriptome editing using signatures that separate RNA base edits from endogenous ADAR editing
- RNA editing can be controlled with dose, system, guide and time
- RNA editing is transient and returns to mock levels as the editor levels decrease in cells

Learn more

acrawley@lifeeditinc.com

life edit® an elevatebio company

References

1. Roth, S.H., Levanon, E.Y. & Eisenberg, E. Genome-wide quantification of ADAR adenosine-to-inosine RNA editing activity. Nat Methods 16, 1131–1138 (2019). https://doi.org/10.1038/s41592-019-0610-9
2. Julien Häsler, Katharina Strub, Alu elements as regulators of gene expression, Nucleic Acids Research, Volume 34, Issue 19, 1 October 2006, Pages 5491–5497, https://doi.org/10.1093/nar/gkl706
3. Lo Giudice, C., Tangaro, M.A., Pesole, G. et al. Investigating RNA editing in deep transcriptome datasets with REDItools and REDIportal. Nat Protoc 15, 1098–1131 (2020). https://doi.org/10.1038/s41596-019-0279-7
4. Mansi L. Tangaro MA, Lo Giudice C, Flati T, Kopel E, Schaffer AA, Castrignanò T, Chillemi G, Pesole G, Picardi E. REDIportal: millions of novel A-to-I RNA editing events from thousands of RNAseq experiments. Nucleic Acids Res. 2021 Jan 8;49(D1):D1012-D1019. doi: 10.1093/nar/gkaa916. PMID: 33104797; PMCID: PMC7778987.
5. Ernesto Picardi, Anna Maria D'Erchia, Claudio Lo Giudice, Graziano Pesole, REDIportal: a comprehensive database of A-to-I RNA editing events in humans, Nucleic Acids Research, Volume 45, Issue D1, January 2017, Pages D750–D757, https://doi.org/10.1093/nar/gkw767
6. Deininger, P. *Alu* elements: know the SINEs. *Genome Biol* **12**, 236 (2011). https://doi.org/10.1186/gb-2011-12-12-236