CRISPR-Cas9 cytidine and adenosine base editing of splice-sites mediates highly-efficient disruption of proteins in primary and immortalized cells

Supplementary Figures Updated March 24th, 2021

AUTHORS

Mitchell G. Kluesner^{1,2,3,4*}, Walker S. Lahr^{1,2,3,4},* Cara-lin Lonetree^{1,2,3,4}, Branden A. Smeester^{1,2,3,4}, Xiaohong Qiu^{1,2,3,4}, Nicholas J. Slipek^{1,2,3,4}, Patricia N. Claudio Vázquez^{1,2,3,4,5}, Samuel P. Pitzen^{2,5}, Emily J. Pomeroy^{1,2,3,4}, Madison J. Vignes⁶, Samantha C. Lee^{5,6}, Samuel P. Bingea^{1,2,3,4}, Aneesha A. Andrew^{5,6}, Beau R. Webber^{1,2,3,4,†}, Branden S. Moriarity^{1,2,3,4,†}

*These authors contributed equally †These authors contributed equally

AFFILIATIONS

1Department of Pediatrics, University of Minnesota, Minneapolis, MN, USA 2Masonic Cancer Center, University of Minnesota, Minneapolis, MN, USA 3Center for Genome Engineering, University of Minnesota, Minneapolis, MN, USA 4Stem Cell Institute, University of Minnesota, Minneapolis, MN, USA 5Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, MN, USA 6College of Biological Sciences, University of Minnesota, Minneapolis, MN, USA

SpliceR communicates directly with Ensembl to pull genetic information and map guides. sgRNAs are then PAM, species being targeted, base editor class, desired splice-sites to target, and a base editing window. Supplementary Figure S1. Diagram of SpliceR v2.0.0. User provides an Ensembl transcript ID, a preferred scored by context motif PAM, species being targeted, base editor class, desired splice-sites to target, and a base editing window. **Supplementary Figure S1**communicates directly with Ensembl an
D . Diagram of SpliceR position of target base in the protopsacer. Users can download predicted sgRNAs.to pull genetic information and map guides. sgRNAs are then v2.0.0. User provides an Ensembl transcript ID, a preferred

Supplementary Figure S2. Distribution of BE-splice sgRNA density across all genes by BE-splice approach. (a) CBE splice donors, (b) ABE splice donors, (c) CBE splice acceptors, (d) ABE splice acceptors. Note that CBE and ABE splice donors utilize the same sgRNAs, hence the same guide density.

Supplementary Figure S3

Supplementary Figure S3. Distribution of the position of the first sgRNA across all genes by BE-splice approach. (a) CBE splice donors, (b) ABE splice donors, (c) CBE splice acceptors, (d) ABE splice acceptors. Note that CBE and ABE splice donors utilize the same sgRNAs, hence the same guide density.

Supplementary Figure S4. Representative gating strategies for flow cytometry. (a) Gating tree for B2M⁺ cells. (b) Gating tree for CD3⁺ cells. Gating strategy applied to data in Figure 2, 3, and 5.

a b

Supplementary Figure S5. Dinucleotide context dependencies of rAPOBEC1-BE4 and TadA^{WT}-TadA^{Evo}-ABE7.10. (a) Summary of linear model of rAPOBEC1-BE4 editing efficiency as a function of pre-dinucleotide context. (b) Summary of linear model of TadA^{WT}-TadA^{Evo}-ABE7.10 editing efficiency as a function of pre-dinucleotide context. (c) Distribution of APOBEC1-BE4 editing efficiency across the protospacer by postdinucleotide context. (d) Distribution of TadA^{WT}-TadA^{Evo}-ABE7.10 editing efficiency across the protospacer by post-dinucleotide context. Note that distributions are not as smooth as pre-dinucleotide context (Fig. 4). (e) Summary of linear model of rAPOBEC1-BE4 editing efficiency as a function of post-dinucleotide context. (f) Summary of linear model of TadA^{WT}-TadA^{Evo}-ABE7.10 editing efficiency as a function of post-dinucleotide context.

Supplementary Figure S6

Supplementary Figure S6. Comparison of context preferences from meta-analysis to BE-hive and the basis of the Honeycomb scoring algorithm. (a-b) BE-Hive predicted editing efficiency for base edits in meta-analysis grouped by preceding dinucleotide context. Predicted results observe identical trends to empirical results from meta-analysis (Figure 4) for context specificities of ABE7.10 (left) and BE4 (right). Boxplot center lines represent the median, box limits represent the upper and lower quartiles, and whiskers define the 1.5x interquartile range. N = 6 papers, 102 guides, 447 edits in total. (c-d) logistic regression context motif weights of ABE7.10 and BE4 first demonstrated in Arbab & Shen et al.40. The height of the logos signify the weight magnitude of having a base identity at a particular position relative to the target base. The direction of the logo signifies whether having that base at that position increases or decreases the odds of editing. Target base is centered in grey. Preceding base context preferences recapitulate results from Fig. 4 a-b. (e) Logistic regression weights for ABE7.10 and BE4 by position of target base in the protospacer. ABE7.10 exhibits a narrowed window relative to BE4, recapitulating results from Fig. 4 a-c. (f) Calculations of Honeycomb score. Logistic weights are summed and transformed into a probability to generate a score. A scaling constant C_{scaling} is added to equation (2) to allow for readily interpretable values. Addition of a scaling constant does not alter the rank of sgRNAs.

Supplementary Figure S7. Mapping of sgRNAs used in this work to the genomic loci of (a) B2M, (b) CD3D, (c) CD3E, (d) CD3G, (e) CD247, (f) TRAC. TRBC1 and TRBC2 were omitted from the BE-splice screen due to the inability to design single BE-splice sgRNAs to target both paralogs simultaneously.

B2M B2M Supplementary Figure S8

Supplementary Figure S8. Mapping of sgRNAs to B2M (β2M) protein structure (PDB 10GA).

Supplementary Figure S9. Mapping of sgRNAs to CD247 (CD3ζ) protein structure (PDB 6JXR).

Supplementary Figure S10. Mapping of sgRNAs to CD3D protein structure (PDB 6JXR).

Supplementary Figure S11. Mapping of sgRNAs to CD3E protein structure (PDB 6JXR).

TCRβ

Supplementary Figure S12

Supplementary Figure S12. Mapping of sgRNAs to TRBC (TCRβ) protein structure (PDB 6JXR).

Supplementary Figure S14. Mapping of sgRNAs to CD3G protein structure (PDB 6JXR).

Supplementary Figure S15

Supplementary Figure S15. Disrupting the immunoinhibitory protein CISH with BE-splice guides in K562 cell line. (a) Map of *CISH* locus with sgRNAs targeting known functional domains. (b) Comparison of editing efficiencies across ABE7.10, ABE8e, and BE4. N = 2 biological replicates. P-values represent results from one-way ANOVA followed by Tukey HSD test. (c) Relative splice site expression of all exon spanning Taqman assays across different treatments, N = 3 technical replicates per 2 biological replicates. (d) Expression of splice site junctions shown as the fold expression of β-Actin. Data demonstrates that the 1a-1b and 1b-2 junctions are nearly undetectable in K562, suggesting the presence of a single major CISH isoform. (e) Uncropped image of gel presented in figure 6e.

Supplementary Figure S17. Digital western blots for CISH following treatment with base editors. (a) Quantification of the relative expression of each band normalized to β-actin and the AAVS1 control. (b) Scatter plot of normalized protein expression vs. the remaining WT allele after treatment, showing a strong correlation between editing and protein knockout. Pearson's correlation coefficient (*r*) shown with two-tailed t-test of correlation coefficient. Error bands represent 95% CI of the mean. (c) Uncropped digital western blot for CISH. (d) Uncropped digital western blot for β-actin. (e) Pilot experiment in T cells before the development of ABE8e with CISH targeting guide. T cells show comparable CISH disruption to K562s. Representative gel of $N = 2$ biological replicates.

Supplementary Figure S18. (a) A standardized workflow for designing and testing base editor sgRNAs for gene disruption.

Supplementary Figure S19.

Table of sgRNA protospacer sequences and corresponding primers used for PCR and Sanger Sequencing.