Supplemental Information

Selective enrichment of A-to-I edited transcripts from cellular RNA using Endonuclease V

Authors: Steve D. Knutson¹, Robert A. Arthur², H. Richard Johnston³, and Jennifer M. Heemstra¹

To whom correspondence should be addressed: jen.heemstra@emory.edu

Affiliations: ¹Department of Chemistry, Emory University, Atlanta, GA 30322, USA; ²Emory Integrated Computational Core, Emory Integrated Core Facilities, Emory University, Atlanta, GA 30322, USA; ³Department of Human Genetics, Emory University, Atlanta, GA 30322, USA
Supplementary Figure 1. EndoV does not cleave RNA in the presence of Ca\textsuperscript{2+}. Cy5 labeled ssRNA I (sequence shown at top with cleavage site indicated by black arrow) was incubated with EndoV and either 10 mM Mg\textsubscript{2+}Cl or increasing amounts of Ca\textsubscript{2+}Cl as indicated. Cleavage is observed only in the presence of Mg\textsuperscript{2+}. Reactions were incubated at room temperature for 3 hours and resolved by 10 % denaturing PAGE.
Supplementary Figure 2. Effect of binding buffer ionic strength on EndoVIPER performance. a) Representative PAGE analysis of initial (I), flowthrough (FT) and eluate (E) EndoVIPER fractions when tested with varying concentrations of CaCl$_2$, NaCl, and KCl in binding buffer. b-c) Densitometric analysis of EndoVIPER efficiency and selectivity as a function of NaCl concentration. Values represent mean with standard deviation. (n = 2)
Supplementary Figure 3. Reduced ionic strength buffer (19 mM Tris, 100 mM NaCl, 1mM CaCl₂, pH 7.4 does not alleviate eEndoV structural binding preferences toward ssRNA. Using buffer formulations from Fig S1, we performed MST analysis of eEndoV binding affinity towards a) dsRNA A and b) dsRNA I targets. Values represent mean with standard deviation. (n = 3) c) Representative PAGE analysis of initial (I), flowthrough (FT) and eluate (E) EndoVIPER fractions when tested with various dsRNA targets. d) Densitometric analysis of EndoVIPER efficiency for dsRNA targets. Values represent mean with standard deviation (n = 2). Unpaired t-test was performed for all samples against ssRNA I pulldowns showing significantly different yields (** denotes p = 0.0045 for all comparisons).
Supplementary Figure 4. Glyoxal reversibly reacts with guanine residues on RNA and disrupts secondary structure. a) Oligoribonucleotide test sequences with glyoxalated guanosine residues (G*) and 20% PAGE analysis of reactions ssRNAs with glyoxal illustrating upward molecular weight gel shift. b) Kinetic analysis of glyoxal deprotection conditions. 10pmol of a glyoxalated ssRNA was incubated at 95 °C for the indicated time periods and analyzed with 20% PAGE.
Supplementary Figure 5. Glyoxal treatment is compatible with EndoVIPER. a) MST analysis of eEndoV binding affinity towards glyoxal treated and unmodified ssRNA A and ssRNA I. Values represent mean with standard deviation. $K_d$ denotes mean with 95% confidence interval (n = 3). b) Representative PAGE analysis of initial (I), flowthrough (FT) and eluate (E) EndoVIPER fractions when tested with increasing concentrations of eEndoV against glyoxal-treated ssRNA targets. c-d) Densitometric analysis of EndoVIPER efficiency and selectivity for glyoxal-treated ssRNA targets. Values represent mean with standard deviation. (n = 2)
Supplementary Figure 6. Glyoxal treatment with EndoVIPER enables robust binding and pulldown efficiency in RNAs with high degrees of secondary structure. a) Schematic of test hairpin RNA I substrate (hRNA I) labeled with Cy5, and b) NUPACK analysis illustrating calculated secondary structure and melting temperature. Inosine site is indicated by the red arrow. c) Glyoxal treatment disrupts secondary structure formation, unfolding the hairpin and providing access for inosine binding and recognition. d) Representative PAGE analysis of initial (I), flowthrough (FT) and eluate (E) EndoVIPER fractions when tested with glyoxal-treated or unmodified hRNA I targets. e) Densitometric analysis of EndoVIPER efficiency for glyoxal-treated or unmodified hRNA I targets. Values represent mean (n = 2) with standard deviation. Unpaired t-test was performed between untreated and glyoxal pulldowns (** denotes p = 0.004). f) MST analysis of eEndoV binding affinity towards glyoxal treated or unmodified hRNA I. Values represent mean with standard deviation. $K_d$ denotes mean with 95% confidence interval (n = 3).
Supplementary Figure 7. Minimal binding bias when using EndoVIPER on a G heavy RNA substrate. 

a) “G heavy” oligoribonucleotide test sequence (G ssRNA I) with glyoxalated G nucleotides (G*) highlighted surrounding an inosine site (red arrow). b) Representative PAGE analysis of initial (I), flowthrough (FT) and eluate (E) EndoVIPER fractions when tested with glyoxal-treated ssRNA I or G ssRNA I targets. c) Densitometric analysis of EndoVIPER efficiency for glyoxal-treated targets. Values represent mean (n = 2) with standard deviation. Unpaired t-test was performed between ssRNA I and G ssRNA I pulldowns (“ns” denotes no significant difference). d) MST analysis of eEndoV binding affinity towards glyoxal treated ssRNA I or G ssRNA I targets. Values represent mean with standard deviation. $K_d$ denotes mean with 95% confidence interval (n = 3).
Supplementary Figure 8. Characterizing RNA size distribution in both input and EndoVIPER pulldowns. Size distribution traces of human brain mRNA used in EndoVIPER and RNA-seq. Initial starting material (untreated) was assessed to confirm material was intact and not degraded. mRNA was then fragmented for ~1 minute at 94 °C with NEBNext® Magnesium RNA Fragmentation module (New England Biolabs) and assessed for size. Material collected as output from EndoVIPER pulldown was also assessed for size. Traces were obtained on an Agilent 2100 bioanalyzer using RNA 6000 Pico Kit (Agilent).
Supplementary Figure 9. Glyoxal treatment of mRNA is fully reversible and does not effect RT or PCR performance. a) PCR amplicons were generated to target known editing sites in both GRIA2 and KCNA1 mRNA transcripts. b) Untreated as well as glyoxal denatured and fully deprotected human brain mRNA was reverse transcribed and PCR amplified using gene specific primers for both GRIA2 and KCNA1. Amplicons were monitored using real-time PCR to determine cycle threshold (Ct) and overall performance. Traces were generated using a LightCycler® 96 instrument (Roche). c) Amplified DNA was purified and analyzed by 1% agarose gel to assess amplicon purity.
GRIA2 (continued)

untreated

glyoxalated, deprotected

untreated

glyoxalated, deprotected

untreated

glyoxalated, deprotected
Supplementary Figure 10. Glyoxal treatment of mRNA is reversible and does not affect sequencing performance. Purified RT-PCR amplicons for both GRIA2 and KCNA1 were subjected to Sanger sequencing. Representative traces were visualized and compared using SnapGene Viewer.

Supplementary Figure 11. Glyoxal treatment of mRNA is reversible and does not affect detection of A-to-I editing. Known protein recoding editing sites (yellow) in both GRIA2 and KCNA1 mRNA transcripts are visible in Sanger traces regardless of glyoxal treatment and removal.
Supplementary Figure 12. Semi-log scatter plot of transcript abundance vs fold enrichment. All enriched sites (supplementary table 8) found in RNA-seq replicates (n = 31,309 sites) were plotted against fold enrichment scores. Semi-log regression (black dashed line) was calculated using GraphPad Prism 8.
Supplementary Figure 13. EndoVIPER-seq enhances detection of edited transcript isoforms. Detection of "edited" reads in upregulated RNA editing sites of interest (supplementary tables 11-14), in a) brain development (462 sites), b) autism spectrum disorder (403 sites), c) schizophrenia (115 sites) and d) protein recoding events in glioblastoma (31 sites). Edited reads were calculated from total read coverage and calculated editing rate at each site. Heatmap columns display both replicate datasets for RNA-seq and EndoVIPER samples, and each row denotes an individual site and scaled to illustrate low (blue) and high (red) read coverage between groups.