Visualizing adenosine-to-inosine RNA editing in single mammalian cells

Ian A Mellis1,2, Rohit Gupte1, Arjun Raj1–3 & Sara H Rouhanifard1

Conversion of adenosine to inosine is a frequent type of RNA editing, but important details about the biology of this conversion remain unknown because of a lack of imaging tools. We developed inoFISH to directly visualize and quantify adenosine-to-inosine-edited transcripts in situ. We found that editing of the GRIA2, EIF2AK2, and NUP43 transcripts is uncorrelated with nuclear localization and paraspeckle association. Further, NUP43 exhibits constant editing levels between single cells, while GRIA2 editing levels vary.

Many RNA species are modified to contain noncanonical bases through a process known as RNA editing. The most prevalent type of modification is adenosine-to-inosine editing1, wherein adenosine deaminases (e.g., ADARs) enzymatically modify an adenosine base to an inosine base (Fig. 1a). The disruption of this process leads to defects in hematopoiesis2 and neurological function3. It has been speculated that adenosine-to-inosine RNA editing influences subcellular localization patterns like nuclear retention4–7, but the lack of tools for visualizing RNA editing has left this and other hypotheses untested. We thus developed inosineFISH (inoFISH), a method based on fluorescence hybridization (FISH) that directly images adenosine-to-inosine RNA editing events with single-molecule resolution.

Discriminating edited from unedited RNA via RNA FISH is difficult, because RNA FISH relies on the hybridization of oligonucleotide probes to visualize the target of interest8. Short oligonucleotides bind nonspecifically, while long oligonucleotides cannot discriminate single-base differences. We thus use a ‘toehold probe’ strategy9 to reduce the initial hybridization region in order to confer selectivity based on single-nucleotide differences (Fig. 1b and Supplementary Fig. 1). We use two detection probes that compete to target the unedited, adenosine-bearing sequence using a thymine and the edited, inosine-bearing sequence using a cytosine. Upon specific binding, the ‘mask’ sequence is released by strand displacement to stabilize hybridization. However, single oligonucleotides are still prone to nonspecific binding, so we simultaneously used smFISH (the ‘mRNA guide’ probe coupled to a unique fluorophore) to target a constant region of mRNA in the vicinity of the edited base (Fig. 1b).

To test inoFISH we chose the canonical, well-studied example of the glutamate receptor 2 transcript (GRIA2) (GRIA2 editing is critical for maintaining neuronal function10, and defects in GRIA2 editing have been associated with amyotrophic lateral sclerosis (ALS)11). We confirmed that GRIA2 was edited by comparing genomic DNA and cDNA sequences in SH-SY5Y cells (Supplementary Fig. 2), and we verified that GRIA2 is a viable target for smFISH (Supplementary Fig. 3). Combining four biological replicates, 10.53% of mRNA guides uniquely colocalized with adenosine- or inosine-detection probes, with 5.25% and 5.28% of GRIA2 guides colocalizing with the adenosine- and inosine-detection probes, respectively (Fig. 1c,d). The estimated mean editing level for GRIA2 was 57.3% (95% confidence interval, 45.1%, 69.5%; full statistical model in Supplementary Note).

To confirm that detection probes did not colocalize with guide probes by random chance, we measured the rate of random colocalization by computationally shifting guide spots by 5 pixels in both the X and Y direction (‘pixel-shift’ analysis), thereby moving them outside the range of any true colocalization events (see Online Methods; Fig. 1d and Supplementary Fig. 4). Pixel-shift analysis reduced colocalization to 1.83% and 1.16% for adenosine and inosine, respectively, and showed that most colocalization events were specific. (Substituting an unrelated guide probe yielded similar results; Supplementary Fig. 5.) To check for dye-specific effects, we swapped fluorophores on the detection probes (Fig. 1d and Supplementary Fig. 4), and this revealed variation in the estimated mean editing level of 22% (Fig. 1c,d; Supplementary Figs. 6 and 7; and Supplementary Note). Together, these findings show that inoFISH can measure editing levels, provided that researchers check for dye-specific biases in detection probe sets.

We validated inoFISH estimates of editing levels by comparing them with three established population-based methods (Supplementary Fig. 8). We generated GRIA2 cDNA and estimated editing ratio either through Sanger sequencing or digestion with a restriction enzyme specific to cDNA from the edited transcript12. We also cloned and sequenced individual GRIA2 cDNA molecules. We found editing levels of 59%, 54.9%, and 50%, respectively, consistent with the 57.3% mean estimated editing level (95% confidence interval, 45.1%, 69.5%) measured by inoFISH. Publicly available RNA-sequencing data from untreated SH-SY5Y cells also revealed GRIA2 editing (see Online Methods and Supplementary Fig. 8).

1Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania, USA. 2Genomics and Computational Biology Group, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. 3Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. Correspondence should be addressed to S.H.R. (sara.rouhanifard@gmail.com).

RECEIVED 22 MARCH; ACCEPTED 5 MAY; PUBLISHED ONLINE 12 JUNE 2017; CORRECTED ONLINE 20 JULY 2017; DOI:10.1038/NMETH.4332
To verify that inoFISH signals were specific to inosine bases and not adenosine or guanosine bases, we altered the frequency of inosines in two different ways. GRIA2 mRNA is primarily edited by the enzyme ADAR2, so we used siRNA to knock down ADAR2 mRNA levels by 60% in SH-SY5Y cells. We observed a concomitant reduction in mean estimated GRIA2 editing level from 65% to 14% (parametric bootstrap \( P = 0.0004 \); Supplementary Note and Supplementary Fig. 6). We also chemically modified inosine bases with acrylonitrile on the \( N^1 \) position to prevent base pairing to cytosine \( ^{15,16} \); this reduced observed editing level from 52.1% to 13.5% (parametric bootstrap \( P = 0.0006 \); Supplementary Figs. 6 and 7).

Additionally, we designed a guanosine-carrying ‘false detection’ probe that should not bind to either the edited or unedited transcript; and it did not bind more than expected by chance (Supplementary Fig. 9). These results show that inoFISH specifically discriminates between adenosine and inosine bases.

We next measured the subcellular localization of edited and unedited transcripts. Previous studies used cell fractionation to show that unmodified RNAs exist in both the nucleus and the cytoplasm, whereas hyperedited RNAs— but not selectively edited RNAs— were predominantly nuclear\(^4,6\). Other studies have shown that mRNAs containing Alu repeats, which are prone to adenosine-to-inosine editing, are inefficiently exported to the cytoplasm\(^5\).

We therefore looked for associations between editing status and subcellular localization of GRIA2 transcripts. We classified GRIA2 transcripts as nuclear if they overlapped with the nuclear stain DAPI. Estimated GRIA2 editing levels were roughly equal in the nucleus and cytoplasm (\( P = 0.38 \); Fig. 2a and Supplementary Fig. 10). (Uncharacteristically of most mRNAs, 93.4% of GRIA2 transcripts localized to the nucleus, though they were still translated; see Supplementary Fig. 3.)

We also used inoFISH to visualize adenosine-to-inosine editing in two additional targets, the hyperedited transcript EIF2AK2 (ref. 17) (Fig. 2b) and the Alu-bearing NUP43 (ref. 5) (Fig. 2c) in U-87 MG cells (Supplementary Fig. 11), which we validated as editing targets by Sanger sequencing of genomic DNA and cDNA (Supplementary Fig. 2). We found that 6.91% and 5.57% of EIF2AK2 guide spots colocalized with adenosine- and inosine-specific
detection spots, respectively (Fig. 2b and Supplementary Fig. 4), giving a population-wide mean editing level estimate of 36.4% (95% confidence interval, 20.4%, 53.1%). For NUP43, 11.3% and 12.4% of guide spots colocalized with the adenosine- and inosine-specific detection spots (Fig. 2c and Supplementary Fig. 4), respectively, giving a population-wide editing level estimate of 53.2% (95% confidence interval, 45.1%, 61.2%). In both cases, the editing level did not vary between nucleus and cytoplasm (P = 0.18, 0.81 for EIF2AK2, NUP43, respectively; Fig. 2b,c and Supplementary Fig. 10). Note that the three inoFISH targets studied had detection efficiencies of 10%, 12%, and 24%; the reasons for this variability is unknown, but efficiencies are within previous bounds. As before, cyanoethylation reduced the percentage of colocalization of inosine-detection probe with guide probe for both EIF2AK2 and NUP43, showing specificity (Supplementary Fig. 7).

inoFISH also allowed us to test whether edited transcripts are trafficked to nuclear paraspeckles. We performed inoFISH together with single-molecule RNA FISH of NEAT1 RNA, a marker of nuclear paraspeckles, in SH-SY5Y cells (Fig. 2d); this revealed that 8.57% of all GRIA2 transcripts colocalized with paraspeckles (Fig. 2d). Simulations (see Online Methods) showed that the observed rate of GRIA2–paraspeckle association was 1.7-fold greater than expected by random chance (simulation of GRIA2–paraspeckle association rate null distribution for one representative replicate P < 0.001; see Online Methods).

We then used inoFISH to determine whether edited or unedited GRIA2 transcripts were preferentially associated with paraspeckles. We found no significant differences in the editing status in paraspeckles for GRIA2 in SH-SY5Y cells (P = 0.44, determined via simulation, for one representative replicate; Fig. 2d), thus demonstrating that edited GRIA2 transcripts in SH-SY5Y cells do not necessarily preferentially associate with paraspeckles.

We next used inoFISH to determine whether adenosine-to-inosine editing is cotranscriptional or post-transcriptional. Introns mark transcription sites, and colocalization of edited transcripts with intron signal would suggest that editing can occur cotranscriptionally. We concurrently performed NUP43 inoFISH with single-molecule FISH targeting NUP43 introns in 212 U-87 MG cells (Fig. 2e); and we observed 17 total transcription sites—of

Figure 2 | Analysis of subcellular localization using inoFISH. Nuclear localization analysis reveals no significant differences (parametric bootstrapping) in editing levels for each target between nucleus and cytoplasm. (a) GRIA2 (n = 4 biological replicates; P = 0.38), (b) EIF2AK2 (n = 3; P = 0.18), and (c) NUP43 (n = 2; P = 0.81) transcripts; representative overlays and fractions of labeled transcripts found to be unedited or edited (inlay). (d) GRIA2 inoFISH results with NEAT1 colocalization (coloc.) pooled over n = 4 biological replicates; representative overlays and counts of inoFISH-labeled, paraspeckle-associated transcripts. (e) NUP43 inoFISH results with transcription site localization analysis (n = 2); representative images and counts of inoFISH-labeled, transcription-site-associated NUP43 transcripts. Scale bars, 5 μm. Blue, adenosine; orange, inosine.

Figure 3 | Single-cell analysis of inoFISH. Simulated inoFISH results assuming (a) binomially distributed per-cell counts of edited and unedited transcripts or (b) two populations of cells, one population with 95% editing and the other with 5% editing, mixed in proportion according to the population-wide editing level. (c) Single-cell analysis of GRIA2 inoFISH results pooled over all four replicates and simulation of the exact conditional null distribution of −log(likelihood) of the data under the binomial model specified in a (inset). (d) Single-cell analysis of NUP43 inoFISH results pooled over all four replicates (left) and simulation of the exact conditional null distribution of −log(likelihood) of the data under the binomial model specified in a (inset; see Online Methods). Obs., −log(likelihood) of observed results under binomial model in a.
these, 5 transcription sites contained unedited NUP43, and none containing edited NUP43 (Fig. 2e). This result does not rule out cotranscriptional editing of NUP43 altogether, but it does suggest that some NUP43 editing may be post-transcriptional.

We also looked for evidence of fluctuations in editing level from cell to cell. We simulated inoFISH results in the cases of uniform (Fig. 3a) or variable editing levels in single cells (Fig. 3b). We found that GRIA2 editing in single cells did not fit the constant editing level model, suggesting per-cell heterogeneity in GRIA2 editing levels (Fig. 3c). NUP43 editing in U-87 MG cells, however, was consistent with the constant-editing level model (Fig. 3d).

Thus, single-cell fluctuations in the level of editing may occur in a target-specific manner.

inoFISH provides a direct method for visualizing adenosine-to-inosine RNA editing in single cells with single-nucleotide resolution. Cell-population-wide studies lack the resolution to provide information such as subcellular localization and cell-to-cell variability of RNA editing. This new tool will enable researchers to answer basic questions about edited RNA species and will enable a deeper understanding of the biology of adenosine-to-inosine editing.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank K. Nishikura and M. Sakurai for helpful discussions, and we thank C. Bartman and B. Emert for comments and suggestions. I.A.M. acknowledges support from NIH/NIGMS T32GM007170 (University of Pennsylvania MSTP), NIH/NHGRI T32HG000046 and NIH/NINDS F30NS100595; S.H.R. from NIH 1F32GM120929-01A1; and A.R. from NSF CAREER Award 1350601, NIH New Innovator 1DP2OD008514, NIH/NIBIB R33 EB019767, NIH 4DN U01 HL129998, and NIH Center for Photogenomics RM1 HG007743.

AUTHOR CONTRIBUTIONS

S.H.R., I.A.M., and A.R. conceived of the paper; S.H.R. and I.A.M. performed experiments; I.A.M., R.G., and A.R. wrote custom software; I.A.M., S.H.R., and A.R. analyzed the data; I.A.M., S.H.R., and A.R. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Bass, B.L. & Weintraub, H. Cell 48, 607–613 (1987).
2. Liddicoat, B.J. et al. Science 349, 1115–1120 (2015).
3. Higuchi, M. et al. Nature 406, 78–81 (2000).
4. Kumar, M. & Carmichael, G.G. Proc. Natl. Acad. Sci. USA 94, 3542–3547 (1997).
5. Chen, L.-L. & Carmichael, G.G. Mol. Cell 35, 467–478 (2009).
6. Zhang, Z. & Carmichael, G.G. Cell 106, 465–475 (2001).
7. Prasanth, K.V. et al. Cell 123, 249–263 (2005).
8. Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A. & Tyagi, S. Nat. Methods 5, 877–879 (2008).
9. Levesque, M.J., Ginart, P., Wei, Y. & Raj, A. Nat. Methods 10, 865–867 (2013).
10. Seeburg, P.H., Single, F., Kuner, T., Higuchi, M. & Sprengel, R. Brain Res. 907, 233–243 (2001).
11. Yamashita, T. et al. Neurosci. Res. 73, 42–48 (2012).
12. Paschen, W., Hedreen, J.C. & Ross, C.A. J. Neurochem. 63, 1596–1602 (1994).
13. Melcher, T. et al. Nature 379, 460–464 (1996).
14. O’Connell, M.A., Gerber, A. & Keller, W. J. Biol. Chem. 272, 473–478 (1997).
15. Yoshida, M. & Ukita, T. Biochim. Biophys. Acta 157, 455–465 (1968).
16. Sakurai, M., Yano, T., Kawabata, H., Ueda, H. & Suzuki, T. Nat. Chem. Biol. 6, 733–740 (2010).
17. Wang, I.X. et al. Cell Rep. 5, 849–860 (2013).
18. Shaffer, S.M. et al. Lab Chip 15, 3170–3182 (2015).
19. Sunwoo, H. et al. Genome Res. 19, 347–359 (2009).
Online methods

inoFISH protocol. The step-by-step inoFISH protocol (description below) can be found in the Supplementary Protocol and the open resource Protocol Exchange20.

Cell culture. We grew human neuroblastoma cells (SH-SY5Y, ATCC CRL-2266) in a 1:1 mixture of Eagle’s Minimum Essential Medium and F12 Medium supplemented with 10% fetal bovine serum (FBS) and 50 𝜇/mL penicillin–streptomycin. We grew human glioblastoma cells (U-87 MG, ATCC HTB-14) in Eagle’s Minimum Essential Medium supplemented with 10% FBS and 50 𝜇/mL penicillin–streptomycin. Note that some SH-SY5Y and U87 cells can be autofluorescent when grown on glass slides; this is improved when the cells are ~70% confluent. Cells were initially checked for mycoplasma contamination and stored as frozen stocks.

Selection of targets for inoFISH. Besides the well-studied editing target GRIA2, we wanted to test inoFISH on editing targets that are less commonly studied in the literature. For these we referred to the literature, to the RADAR database of RNA editing, and to publicly available RNA-seq data for screening. We identified NUP43 and EIF2AK2 as targets that are studied in a relatively small number of research groups, are candidate editing targets in published transcriptome-wide adenosine-to-inosine editing screens, and have editing sites amenable to inoFISH probe designs. In order to systematically identify inoFISH targets, we found it useful to first filter targets by those with enough nonrepetitive sequence to design a full guide probe (>24 20-mer oligonucleotides). Then, we performed transcriptome-wide screens aimed at studying editing targets in the RADAR database that are conserved across humans, nonhuman primates, and mice (http://rnaedit.com/; accessed 28 October 2016). We expect that conserved targets are more likely to be both biologically interesting and observable in our cell lines than targets that only appear in screening experiments in humans.

RNA-sequencing-based screen for candidate editing targets. We downloaded publicly available RNA-seq screening data from EBI ArrayExpress for total RNA from two biological replicates of SH-SY5Y cells21 and three replicates of U-87 MG cells22. We also downloaded RNA editing candidates’ positions (in human genome build hg19) from the RADAR database, with a focus on conserved positions and those derived from neural lineage samples. We aligned raw reads to hg19 with STAR v2.3.0e with default parameters, except runThreadN was set to 4. Next, we used picard tools v1.96 (http://broadinstitute.github.io/picard/). Accessed: 28th October 2016) to interconvert between aligned SAM and BAM file formats (SamFormatConverter), to sort BAMs by coordinate (SortSam), to index sorted BAMs (BuildBamIndex), and to remove PCR duplicates (MarkDuplicates). We then trimmed 5’ ends of all aligned and PCR-deduplicated reads with samUtil v1.0.13 (http://genome.sph.umich.edu/wiki/SamUtil). Accessed: 28th October 2016). Finally, using bam-readcount v0.7.4 we filtered aligned reads by those overlapping candidate editing sites of interest, by their overall mapping (phred) score, and by individual base-call qualities using a custom bed file and parameters -q 25 and -b 20. We quantified RNA-seq-based editing level estimates in R by calculating the fraction of filtered reads with ‘G’ base calls at each editing site. We considered for downstream RNA editing verification and inoFISH probe design any sites in these screens with at least seven overlapping filtered reads, of which at least one was called ‘G’ at the editing site.

Verification of RNA editing of candidate targets. As described below, we used RT-PCR of total RNA and PCR of genomic DNA in cell lines of interest to further check that candidate editing sites were in fact RNA edited in our cell lines.

Optimization of inoFISH targets. We then chose targets for inoFISH by checking RT-PCR and genomic DNA PCR Sanger sequencing results for each candidate site to ensure that there would be no additional polymorphisms in transcripts, resulting either from RNA editing or single-nucleotide polymorphisms, in the regions flanking the editing site up to 30 bp upstream or downstream (Supplementary Fig. 4). We ultimately designed inoFISH probe sets against one editing site in NUP43 and two editing sites in EIF2AK2 (both using the same guide probe set) that appeared to be amenable to inoFISH guide and detection probe set design. We were able to verify inoFISH probe binding with detection efficiencies greater than those expected by random colocalization for the one NUP43 candidate editing site and one of the two EIF2AK2 editing sites. (See below for experimental methods.)

Genotyping of edited regions. We extracted genomic DNA from SH-SY5Y cells and U-87 MG cells using the Qiagen DNeasy Blood & Tissue kit. We used Platinum Taq (Invitrogen cat. no. 10966-018) for PCR amplification of the genomic regions of interest for each target, following manufacturer’s recommendations for reaction component concentrations. We PCR amplified two biological replicates, each with two technical PCR replicates. For GRIA2, we used primers GRIA2-F1 and GRIA2-R2 (Supplementary Table 1). For EIF2AK2, we used EIF2AK2_20-F1 and EIF2AK2_20-R1; and for NUP43 we used NUP43-F1 and NUP43-R1 (Supplementary Table 1). We confirmed PCR product sizes by gel electrophoresis, using a 1.5% agarose gel in TAE. Then, we treated these PCR products with ExoSAP-IT (Affymetrix 78200) according to manufacturer’s instructions, and we submitted them for Sanger sequencing at the University of Pennsylvania DNA sequencing facility.

Estimation of editing efficiency by RT-PCR and Sanger sequencing. We extracted total RNA from SH-SY5Y and U-87 MG cells using miRNeasy kits (Qiagen cat. no. 217004) according to manufacturer’s instructions. Then, we reverse transcribed target transcripts around editing sites of interest using Superscript III first strand RT kit (ThermoFisher cat. no. 18080044) according to manufacturer’s instructions. In separate reactions for RNA from each cell type, we used both oligo-dT and transcript-specific primers for reverse transcription. Briefly, we used 50 ng of RNA per reaction for reverse transcription with either oligo-dT or transcript-specific primers (Supplementary Table 1). Then, we performed PCR with transcript-specific primers (Supplementary Table 1) using Platinum Taq (Invitrogen cat. no. 10966-018) according to manufacturer’s instructions. We completed biological replicates, each with technical PCR replicates for these reactions. We confirmed product sizes by gel electrophoresis on 1.5% agarose gels in TAE. Then, we treated these products with ExoSAP-IT (Affymetrix cat. no. 78200) according to manufacturer’s instructions and submitted for sequencing at the University of Pennsylvania DNA sequencing facility.

DOI:10.1038/nmeth.4332
Estimation of editing efficiency by clonal analysis of GRIA2 RT-PCR product. The amplified GRIA2 cDNA was cloned into a vector using the TOPO TA cloning kit (ThermoFisher), transformed into chemically competent Escherichia coli cells, and plated on LB plates with 0.1 mg/mL ampicillin. We isolated DNA from >20 individual colonies and submitted this DNA for sequencing at the University of Pennsylvania DNA sequencing facility. We performed sequence alignment at the editing site using MAFFT in Benchling to determine the ratio of edited and unedited transcripts.

Estimation of editing efficiency by restriction digest and bioanalyzer analysis. Edited and unedited GRIA2 cDNAs yield distinct restriction fragment patterns upon digestion with BbvI3. Edited GRIA2 cDNA yields two DNA fragments upon digestion (225 bp and 46 bp); and unedited GRIA2 cDNA yields three DNA fragments (145 bp, 80 bp, and 46 bp). Following BbvI digestion (NEB cat. no. R0173S) of GRIA2 cDNA, according to manufacturer’s instructions, we submitted digestion products for fragment sizing analysis on an Agilent 2100 Bioanalyzer at the University of Pennsylvania DNA sequencing facility.

RNA probe design and synthesis. For each of the validated editing sites, we designed probes by matching free energies of hybridization as specified in Levesque et al.9 (2013b). We optimized mask oligonucleotides to leave 8-bp overhangs for each of the detection probes and pooled all five together to act as the complete allele-specific probe. We provide all oligonucleotide sequences in Supplementary Table 1. We coupled 3 amine-labeled adenosine- and inosine-detection probes to NHS-Cy3 or NHS-Cy5 fluorophores (GE Healthcare) and purchased guide probes labeled with Cal Fluor 610 (Biosearch Technologies). We also acquired images in the Atto 488 channel with a 1,000 ms exposure as a marker of autofluorescence.

inoFISH procedure. We grew cells on glass coverslides until ~70% confluent. We washed the cells twice with 1× PBS, then fixed for 10 min with 4% formaldehyde–1× PBS at room temperature. We aspirated off the formaldehyde and rinsed twice with 1× PBS prior to adding 70% ethanol for storage at 4 °C or inoFISH after a 1 hour permeabilization in 70% ethanol. We incubated our cells overnight at 37 °C in hybridization buffer (10% dextran sulfate, 2× SSC, 10% formamide) with 100 nM concentration of guide probe, 24 nM concentration of the adenosine- and inosine-detection probes, and 72 nM concentration of the mask probe—ensuring excess mask for complete hybridization to the detection probes. The following morning, we performed two washes in wash buffer (2× SSC, 10% formamide), each consisting of a 30-min incubation at 37 °C. After the second wash, we rinsed once with 2× SCC/DAPI and once with antifade buffer (10 mM Tris (pH 8.0), 2× SSC, 1% w/v glucose). Finally, we mounted the sample for imaging in an antifade buffer with catalase and glucose oxidase (Raj et al.9) to prevent photobleaching. We performed RNA FISH on cell culture samples grown on a Lab-Tek chambered coverglass using 50 µL of hybridization solution spread into a thin layer with a coverslip and placed in a parafilm-covered culture dish with a moistened Kimwipe to prevent excessive evaporation.

Imaging. We imaged each sample on a Nikon Ti-E inverted fluorescence microscope using a 100× Plan-Apo objective (numerical aperture of 1.40) and a cooled CCD camera (Andor iKon 934). For 100× imaging, we acquired z-stacks (0.3 µm spacing between stacks) of stained cells in five different fluorescence channels using filter sets for DAPI, Cy3, Calfluor 610, Cy5, and Atto 700. The filter sets we used were 31000v2 (Chroma), 41028 (Chroma), SP102v1 (Chroma), 17 SP104v2 (Chroma), and SP105 (Chroma) for DAPI, Atto 488, Cy3, Atto 647N/Cy5, and Atto 700, respectively. A custom filter set was used for Alexa 594/CalFluor610 (Omega). We tuned the exposure times depending on the dyes used: 4 s for each guide probe; 4,000 ms for each of the detection probes; 5,000 ms for the NEAT1 probe; and 7,000 ms for ADAR1 and ADAR2 probes. We also acquired images in the Atto 488 channel with a 1,000 ms exposure as a marker of autofluorescence.

Image analysis. We first segmented and thresholded images using a custom Matlab software suite (downloadable at https://bitbucket.org/arjunrajlabatory/rajlabimagetools/wiki/Home). Segmentation of cells was done manually by drawing a boundary around nonoverlapping cells. The software then fits each spot to a two-dimensional Gaussian profile specifically on the z plane on which it occurs in order to ascertain subpixel-resolution spot locations. Colocalization took place in two stages. In the first stage, guide spots searched for the nearest-neighbor detection probes within a 2.5-pixel (360-nm) window. We ascended the median displacement vector field for each match and subsequently used it to correct for chromatic aberrations. After this correction, we used a more stringent 1.5-pixel (195-nm) radius to make the final determination of colocalization. In order to test random colocalization caused by spots occurring randomly by chance, we took our images and shifted the guide channel by adding 5 pixels (1.3 µm) to the X and Y coordinates and then performing colocalization analysis.

Autofluorescence subtraction. For U-87 MG cells, we controlled for punctate autofluorescence by imaging with the 41028 (Chroma) filter set, the ‘gfp channel’, which we have previously found to be sensitive for autofluorescence in this cell line (data not shown). We performed colocalization as previously described between guide spots and any spot-like autofluorescence called in the gfp channel. In R, we excluded spots colocalizing with this autofluorescence from all inoFISH analyses.

Subcellular localization. Nuclear localization. We extracted a DAPI nuclear mask as previously described8. We call a spot as localized to the nucleus if the guide spot X and Y coordinates overlap with the 2D nuclear mask.

Localization to transcription sites. We visualized NUP43 introns by probing with intron-specific probes coupled to Atto 700 and imaging with SP105 filter set. We used the txnSiteGUI2 interface within rajlabimagetools to manually curate calls of exon-intron spot colocalization.

Localization to paraspeckles. We visualized paraspeckles by probing with NEAT1-specific probes coupled to Atto 700 and imaging with SP105 filter set. We used the txnSiteGUI2 interface within rajlabimagetools to manually curate calls of transcript–paraspeckle association.
**In situ cyanoethylation.** Cyanoethylation was performed in a manner similar to that of previous descriptions. We aspirated the 70% ethanol off of the fixed cells and added cyanoethylation solution (1.1 M triethylammoniumacetate (pH 8.6) resuspended in 100% ethanol) with or without 1.6 M acrylonitrile at 70 °C for 15 min. We used a large volume to prevent drying from evaporation. We removed the treated sample from heat after 15 min (30 min incubation abolishes guide probe signal) and washed twice with wash buffer (2× SSC, 10% formamide) before beginning inoFISH procedure.

**Statistical analysis. Detection efficiency.** For each label (edited or unedited) in each experiment we calculated the mean fraction of transcripts colocalized with a spot that label over all replicates (excluding three-color spots). For complete details of this analysis, please see the Supplementary Note.

**Population-wide editing level estimation by inoFISH.** We define the population-wide editing level estimate as the average over all replicates of the inferred fraction of uniquely labeled guide spots labeled as edited. For a complete description of our estimation of population-wide editing level, please see the Supplementary Note.

**Paraspeckle–transcript association rates.** In MATLAB, we simulated the exact conditional null distribution of paraspeckle–transcript association rates for each experiment under the null hypothesis that a paraspeckle and a nuclear-localized transcript will only colocalize by chance. For each cell in the experiment, we conditioned on (i) the shape and size of that cell’s nucleus, (ii) the locations of all paraspeckles in that nucleus, and (iii) the number of transcripts of interest (GRIA2 or EIF2AK2) retained in the nucleus. In order to efficiently simulate these distributions, rather than using txnSiteGUI as above, we generated 2D masks for paraspeckle locations and called paraspeckle–transcript association when a randomly placed transcript spot overlapped with this mask. We selected the mask size as 25 pixels per paraspeckle spot called—roughly the mean paraspeckle size—based on our inspection of paraspeckles while calling spots (as in image analysis). For each experiment, we simulated draws from the exact conditional null distribution 1,000 times. A raw $P$ value for paraspeckle–transcript association rate is equal to the fraction of simulations with a higher paraspeckle–transcript association rate. We similarly simulated exact conditional null distributions for paraspeckle-edited-transcript and paraspeckle-unedited-transcript association rates.

**Single-cell editing level distributions.** In R, we assessed single-cell spot counts after inoFISH colocalization as reported by rajlabimagetools (https://bitbucket.org/arjuna/rajlabimagetools/wiki/Home; in Image analysis), as well as after autofluorescence subtraction (for U-87 MG data). We simulated the null distribution of data likelihood under a null model wherein all cells share the same effective editing level; for an experiment with overall estimated editing level equal to $p_e$ (above), let $n_{ij}$ be the number of edited transcripts detected in cell $j$ and $n_{ij}^u$ be the number the number of unedited transcripts detected in cell $j$. Under the null model, $n_{ij}$ is drawn from a binomial with $(n_{ij} + n_{ij}^u)$ draws and probability $p_e$. We simulated single-cell label counts for cells by drawing from these conditional null distributions for each cell 100,000 times. We then compared the negative log-likelihood of the observed data, combined over all replicates, with the distribution of negative log-likelihoods of each simulation iteration. A $P$ value of 0.12 indicates that 12% of the simulated iterations had a negative log-likelihood that was greater than that of the observed data. Note that the negative log-likelihood density plots in Figure 3 are subsampled to 3,000 of the aforementioned 100,000 such iterations per plot in order to facilitate figure generation. For complete details of this analysis, please see the Supplementary Note.

**siRNA knockdowns of ADAR2.** Briefly, we used lipofectamine RNAlmex to transfect SH-SY5Y cells with Silencer Select siRNAs targeting ADARB1 (ADAR; ID:s1011, Ambion) and a negative control siRNA (#1, Ambion) for 72 h, verifying knockdown via RNA FISH of ADAR2.

**SFPQ-guided GRIA2 inoFISH.** We performed GRIA2 inoFISH, as described above, but substituted an smFISH probe set for SFPQ for the GRIA2 guide. Like GRIA2 transcripts, SFPQ transcripts are localized to the nucleus in SH-SY5Y cells. In parallel, we performed regular GRIA2 inoFISH on a sample of cells from the same passage. We counted the number of GRIA2 mRNA per cell in the regular sample, and we subsampled the SFPQ ‘guide’ spots from that distribution of mRNA counts per cell. In this way, we could more directly compare colocalization rates with GRIA2 detection probes between SFPQ-guided and GRIA2-guided experiments. We then performed colocalization as described above.

**Cell-cycle inhibitor.** We measured nuclear retention of GRIA2 mRNA by inhibiting transcription for 24 h by applying aphidicolin at 1 μg/ml.

**Code availability.** Scripts for all analyses presented in this paper—including all data extraction, processing, and graphing steps—are freely accessible at the following url: https://www.dropbox.com/sh/j5umeneit1ck9/AAA4W4I648glUUhePjIxyaRaa?dl=0.

**Data availability statement.** All imaging and other non-RNA-sequence data associated with this paper are also freely available at https://www.dropbox.com/sh/vwwf9mgg72o75c/AACsFk6VbjHY2S5MK8jLR2jNa?dl=0. For publicly available RNA sequencing data discussed in Supplementary Figure 1, please see EBI ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) accession numbers E-MTAB-2690 (SH-SY5Y samples “SY5Y_A” and “SY5Y_B”) and E-MTAB-1875 (U-87 MG samples “s_2_78”, “s_2_82”, and “s_2_88”).

20. Mellis, I.A. et al. Protocol Exchange http://dx.doi.org/10.1038/proteRx.2017.056 (2017).
21. Duffy, D.J. et al. Oncotarget 6, 43182–43201 (2015).
22. Kuilwilm, M., Davierwala, A. & Pääbo, S. PloS One 8, e83218 (2013).
23. Whitney, N.P. et al. FASEB J. 22, 2888–2900 (2008).
24. Gabill, M.N. et al. Genome Biol. 16, 20 (2015).

doi:10.1038/nmeth.4332
Corrigendum: Visualizing adenosine-to-inosine RNA editing in single mammalian cells

Ian A Mellis, Rohit Gupte, Arjun Raj & Sara H Rouhanifard

Nat. Methods doi:10.1038/nmeth.4332; corrected online 20 July 2017

In the version of this article initially published online, a statement in the introductory paragraph incorrectly implied that deamination leads to defects in hematopoiesis and neurological functions. It is the lack of deamination that causes these defects. The error has been corrected in the print, PDF and HTML versions of this article.