

Supplementary Information I

SABER enables highly multiplexed and amplified detection of DNA and RNA in cells and tissues

Jocelyn Y. Kishi^{1,2,†} Brian J. Beliveau^{1,2,†} Sylvain W. Lapan^{3,†} Emma R. West³ Allen Zhu^{1,2} Hiroshi M. Sasaki^{1,2}
Sinem K. Saka^{1,2} Yu Wang^{1,2} Constance L. Cepko^{3,*} Peng Yin^{1,2,*}

Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115¹

Department of Systems Biology, Harvard Medical School, Boston, MA 02115²,

Department of Genetics Harvard Medical School, Boston, MA 02115³,

These authors contributed equally[†],

Correspondence*: py@hms.harvard.edu (P.Y.), cepko@genetics.med.harvard.edu (C.L.C.)

Contents

1	Supplemental information for Fig. 1	1
2	Supplemental information for Fig. 2	2
3	Supplemental information for Fig. 3	3
4	Supplemental information for Fig. 4	4
5	Supplemental information for Fig. 5	5
6	Supplemental information for Fig. 6	6
7	Supplemental information for Fig. 7	7
8	Supplemental Experimental Procedures	8
9	Supplemental References	9

List of Figures

S1	Related to Fig. 1.	1
S2	Related to Fig. 2.	2
S3	Related to Fig. 3.	3
S4	Related to Fig. 4.	4
S5	Related to Fig. 5.	5
S6	Related to Fig. 6.	6
S7	Related to Fig. 7.	7

1 Supplemental information for Fig. 1

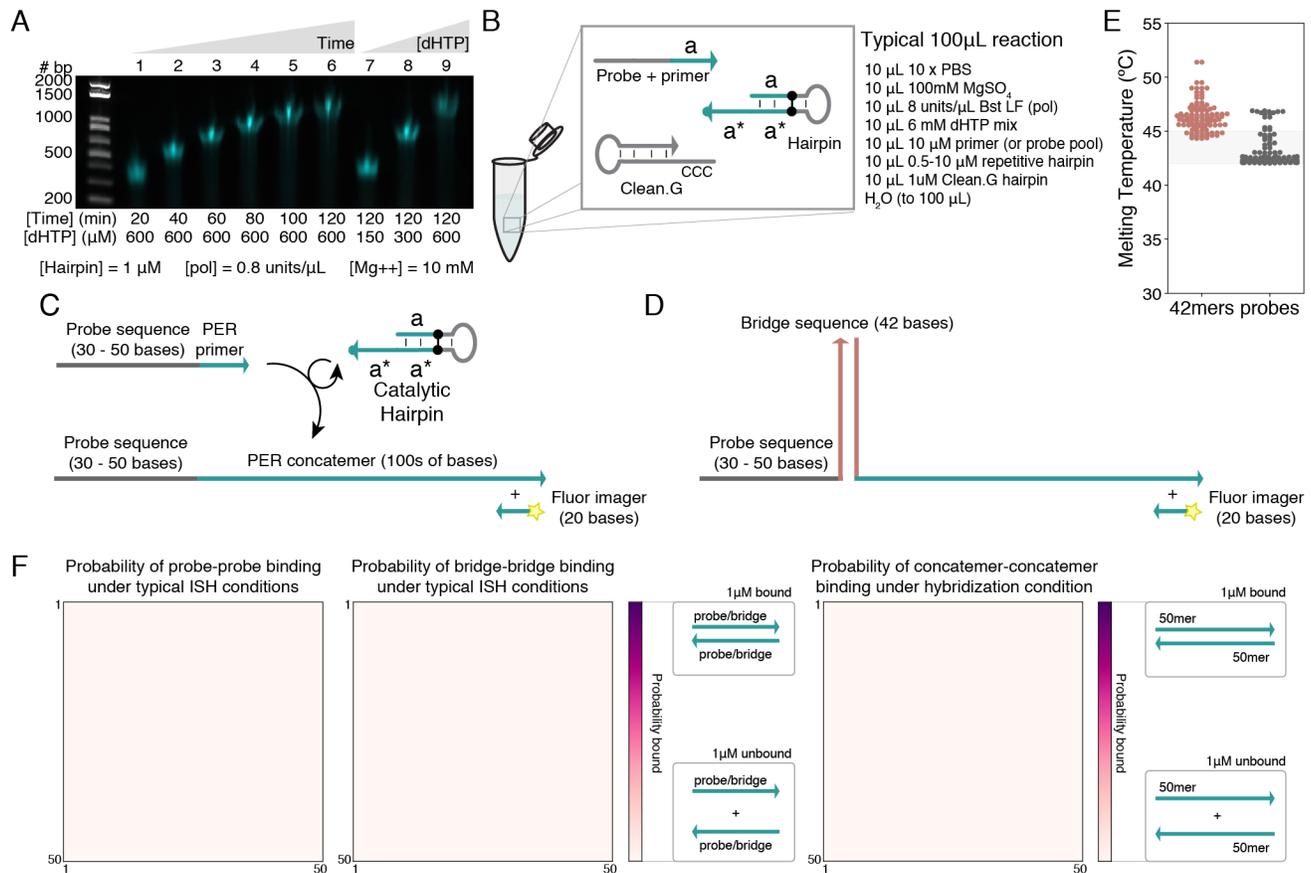


Figure S1: Related to Fig. 1. (A) PER concatemer length control by varying incubation time and dHTP (dATP, dTTP, and dCTP) concentration. (B) Schematic representation of typical 100 μL PER concatemerization reaction. (C) Probe sequences are typically computationally designed using the OligMiner pipeline¹ to have 30-50 bases of homology to targets of interest, and they are chemically synthesized with additional short PER primer sequences on their 3' ends. A catalytic hairpin is then used to extend the concatemer sequence *in vitro*. Complementary fluorescent 'fluor' imagers that have 20 bases of homology to the concatemer are typically used for imaging. (D) Probes may instead be appended with one of 84 available 42mer bridge sequences. The complement of the 42mer bridge sequence can then be appended with a PER primer and concatemerized *in vitro*, followed by co-hybridization with the probe-bridge strands *in situ*. (E) Melting temperatures of the fifty 42mer bridge and probe sequences from Fig. 1F as calculated with Biopython.² The typical FISH temperature range (42°C-45°C) is highlighted. (F) Orthogonality plots calculated with NUPACK³⁻⁵ of 50 probe, 42mer, and 50mer PER concatemer sequences show that homodimerization and heterodimerization are negligible under typical hybridization conditions.

2 Supplemental information for Fig. 2

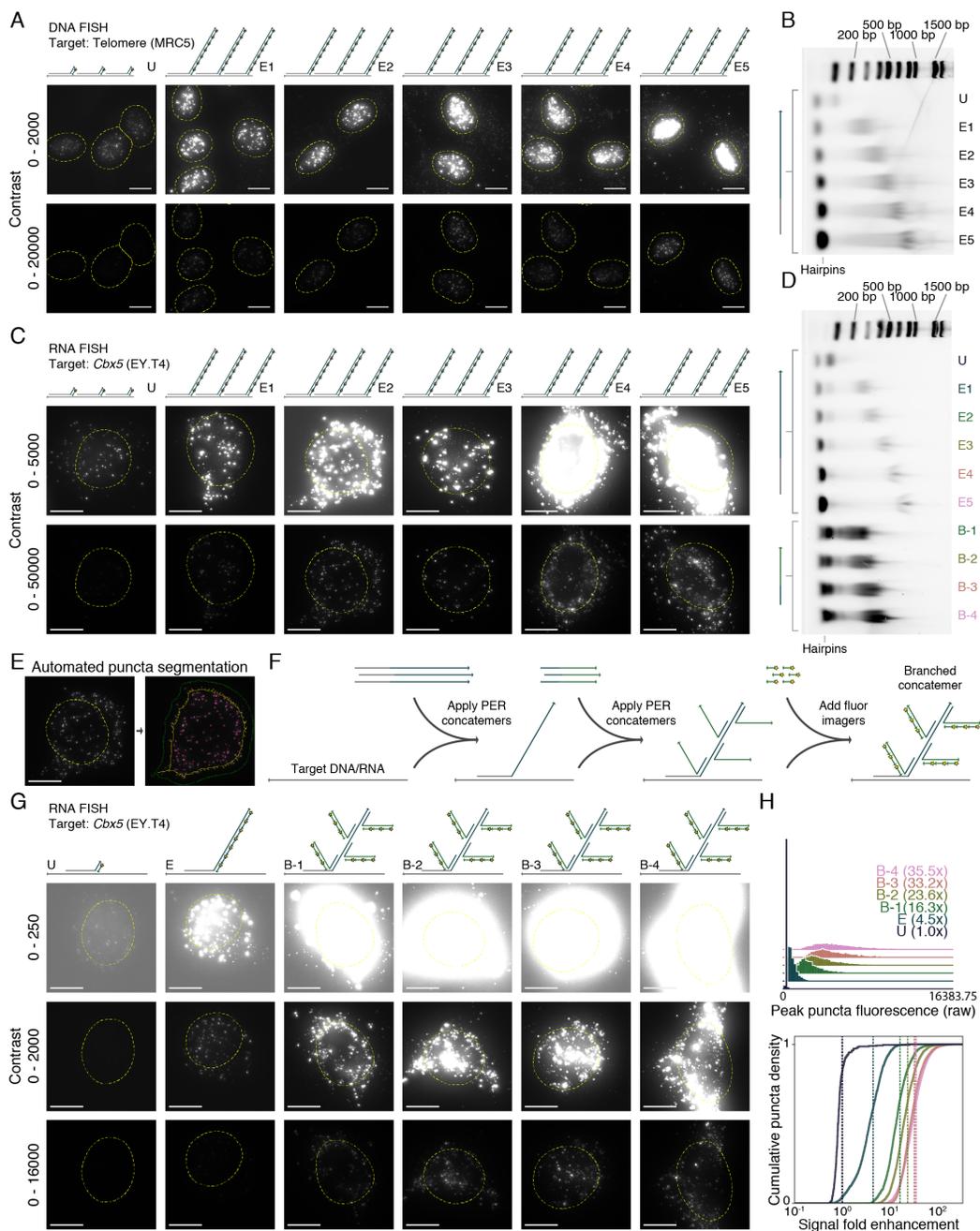


Figure S2: Related to Fig. 2. (A) Representative images of all 5 extension length conditions (E1-E5) compared to unextended (E) corresponding to Fig. 2A-B are shown under two contrast settings. (B) Gel electrophoresis shows the lengths of the extensions shown in (A) range from ~250 to 900 nt. (C) Representative images of all 5 extension length conditions (E1-E5) compared to unextended (E) corresponding to Fig. 2C-D are shown under two contrast settings. (D) Gel electrophoresis shows the lengths of the extensions shown in (B) range from ~250 to 900 nt. Shorter lengths (~250 to 500 nt) were used to characterize one round of branching (see parts (F)-(H)). (E) An example microscopy image (left) shown beside an overlay output (right) by the CellProfiler⁶ pipeline, which automatically detects cell bodies (green), shrunken cell bodies (yellow), and puncta within these shrunken cell bodies (pink). (F) Schematic depiction of one round of branching, where a second set of PER concatemers is hybridized to a first set of hybridized concatemr extensions. (G) Representative images depicting a single round of branching using four different branch sequence lengths (B1-B4, see part (D)) compared to single extension (E) and unextended (U) conditions. (H) Quantification as done in Fig. 1D was used to quantify one round of branching. In total, the longest branch condition (B4) showed >30 fold amplification over the unextended condition. Scale bars: 10 μ m.

3 Supplemental information for Fig. 3

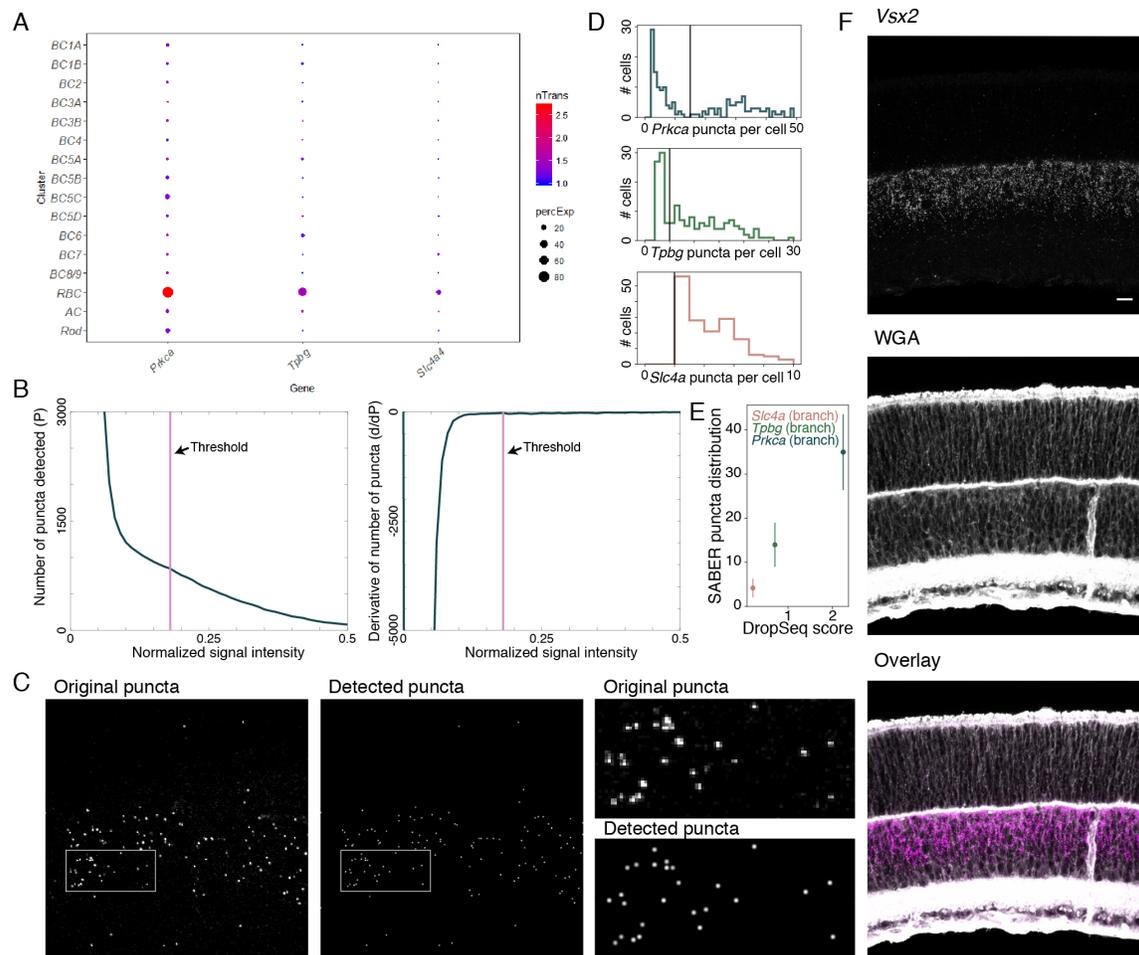


Figure S3: Related to Fig. 3. (A) Dot plot displaying average transcripts per cell (nTrans) in expressing cells and percent of cells in each cluster with detection of the indicated transcript (percExp) (data from ref⁷). Plot displays all individual bipolar cell types and two additional populations: amacrine cells (AC), a highly heterogeneous population located basal to the bipolar cells, and rods. Note that indicated transcripts are highly enriched in RBCs relative to other bipolar cell types, but may be abundantly expressed in subpopulations of amacrine cells and other spatially separated retinal populations not deeply sampled in the dataset. (B-C) Puncta were detected in 3-D retinal tissue using a custom MATLAB Laplacian of Gaussian (LoG) method, requiring a threshold on the LoG-filtered image. To determine the appropriate threshold for each condition, we counted the number of puncta detected across the range of possible thresholds and plot the change in number of puncta detected as a function of threshold (B, left). Generally, we observed a plateau or elbow in the graph where the number of puncta is relatively constant. For a more obvious threshold choice, we plotted the derivative graph, which rapidly approaches zero as the number of transcripts becomes constant (B, right). Choosing a threshold just after the graph plummets to zero generated reliable segmentation results that were consistent with visual verification. The chosen threshold is depicted by the vertical line in each graph and the resulting puncta detection can be compared to the original SABER image in a 2-D projection (C). (D) Full distributions of the number of puncta per cell for each RBC probe after cell and puncta segmentation. Thresholds for calling positive cells are shown with a vertical line and were chosen to eliminate cells with background levels of transcript. For transcripts of high abundance like *Prkca* and *Tpbp*, puncta from positive cells located near cell-cell boundaries were occasionally miss-assigned to adjacent, negative cells. Thresholding in this way allows us to exclude these cases from the analysis, yielding a more accurate average value. (E) Quantification of RBC marker transcripts with branching recapitulates relative abundance of transcripts (average number of transcripts per cell) observed from >10,000 single RBCs profiled in the Drop-seq dataset.⁷ n=65-78 cells. (F) Detection of *Vsx2* transcript using a probe pool of 12 extended 47 nt oligos, with a single round of branch amplification. Signal can be compared to *Vsx2* detection in Figure 5 performed using a set of 40 probes without branch amplification. Retina is P9. Scale bar: 10 μ m.

4 Supplemental information for Fig. 4

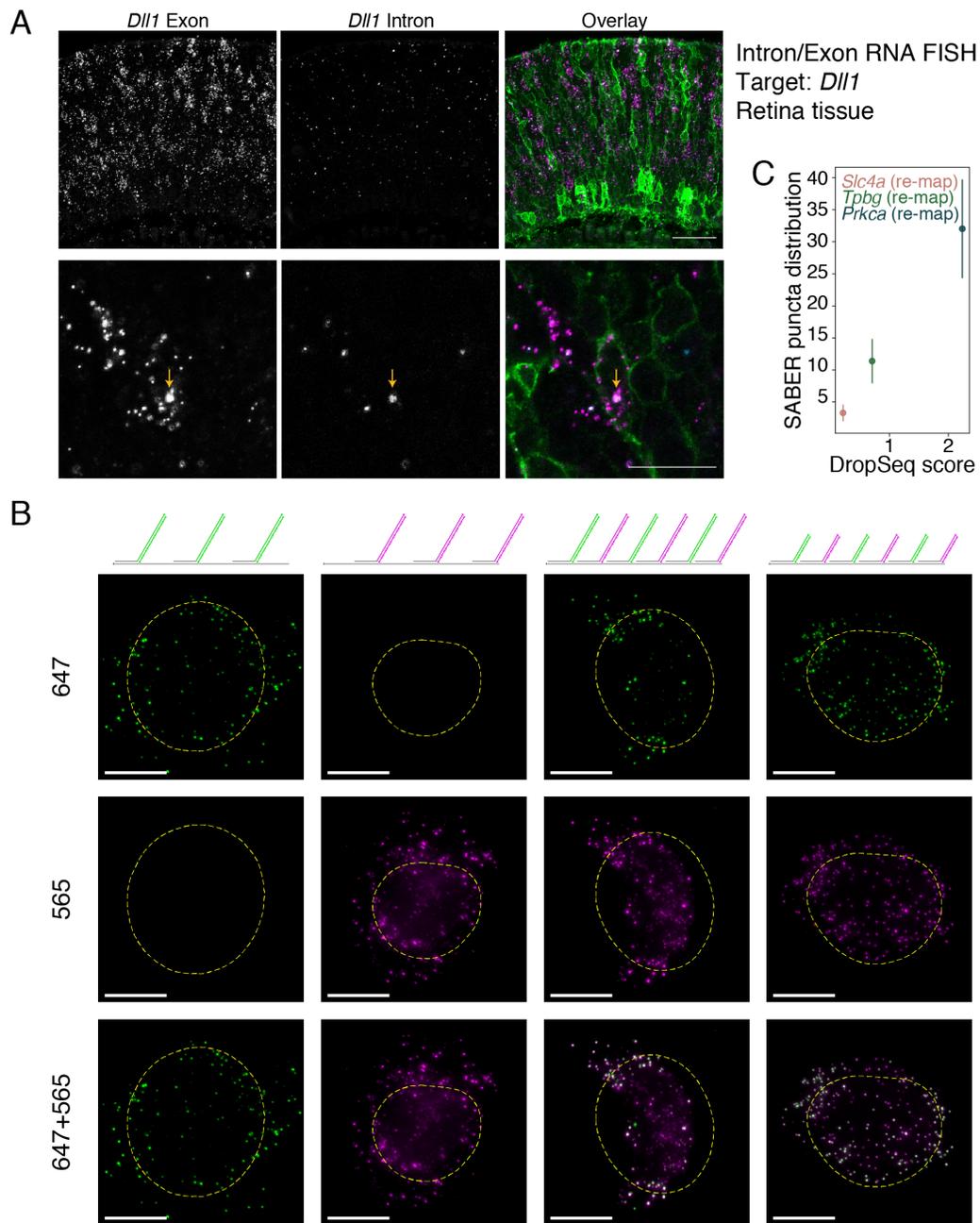


Figure S4: Related to Fig. 4. (A) Multiplexed SABER enables co-detection of intronic and exonic sequences. Bottom panel shows magnified view. Retinas were electroporated with CAG-mGFP, detected after in situ by IHC for GFP, to visualize cell boundaries. Arrow indicates presumed site of transcription based on overlap of intronic and exonic sequence detection. (B) Control conditions for Fig. 4D show no bleedthrough between the 565/647 detection channels. The longer concatemer condition showed reduced co-localization of channel puncta (67-82%) than the shorter extensions (92-95%) as determined by our automated segmentation pipeline, once again demonstrating the importance of length programmability of concatemers. See Supplemental Experimental Procedures for cell and puncta counts. (C) Detection of transcripts using remapped probes for *Prkca* and *Tpbp*, along with *Slc4a* (Fig. 4E-F), compared against average number of transcripts per cell measured by Drop-Seq.⁷ n=38-52 cells. Scale bars: 10 μ m.

5 Supplemental information for Fig. 5

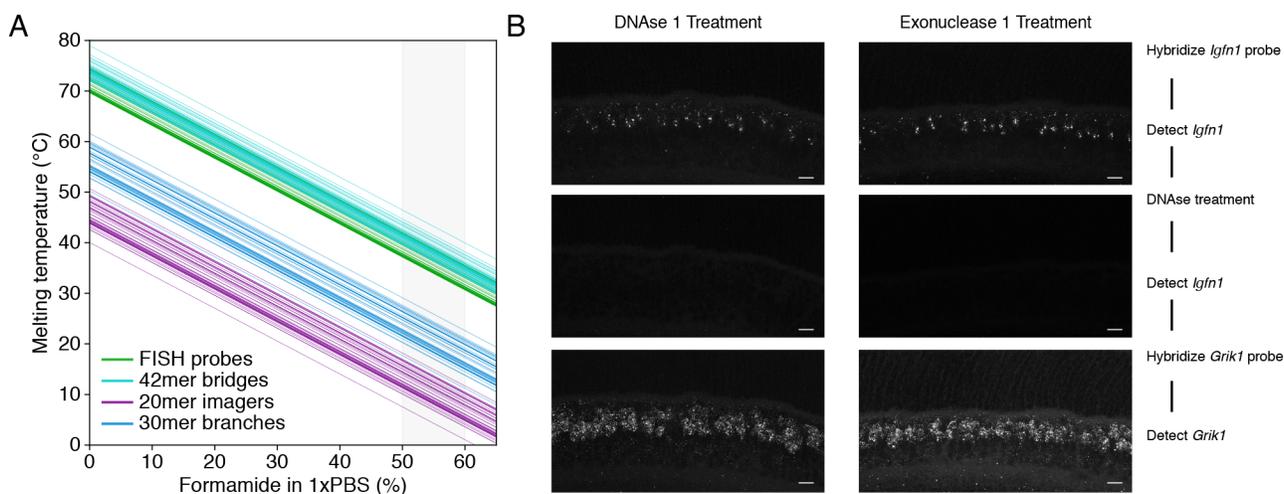


Figure S5: Related to Fig. 5. (A) Melting temperatures calculated using Biopython² for 20mer imager sequences, 30mer branch sequences, 42mer bridge sequences, and probes (corresponding to the sequences analyzed in Fig. 1F) were modeled under different formamide concentrations in 1xPBS. Typically, 50%-60% formamide in 1xPBS is effective in rapidly destabilizing imager strands while leaving underlying bridge and probe sequences hybridized. (B) Sequential hybridization demonstrating recycling of concatemer sequences by digestion with DNase I or Exonuclease I. First, type 7 ON bipolar cells are detected by probing for the marker gene *Igfn1*.⁷ After detection, DNases are applied and successful probe degradation is confirmed by reapplication of fluorescent imagers. Finally, a second round of probe hybridization is performed using a probe set bearing the same concatemer to detect a different bipolar cell population (OFF bipolar cells, *Grik1* transcript), to confirm stability of RNA for FISH detection after DNase treatment. Scale bars: 10 μm .

6 Supplemental information for Fig. 6

A

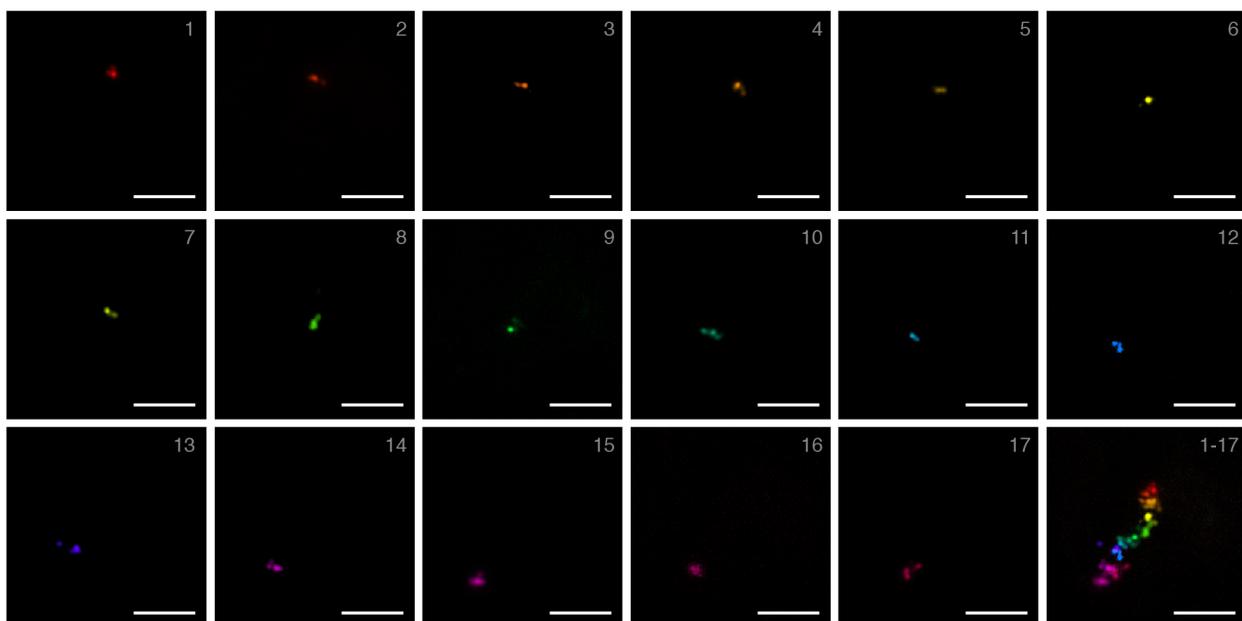


Figure S6: Related to Fig. 6. (A) Individual channel images corresponding to the experiment Fig. 6A-B, along with an overlay of all 17 positions, are shown. Scale bars: 5 μm .

7 Supplemental information for Fig. 7

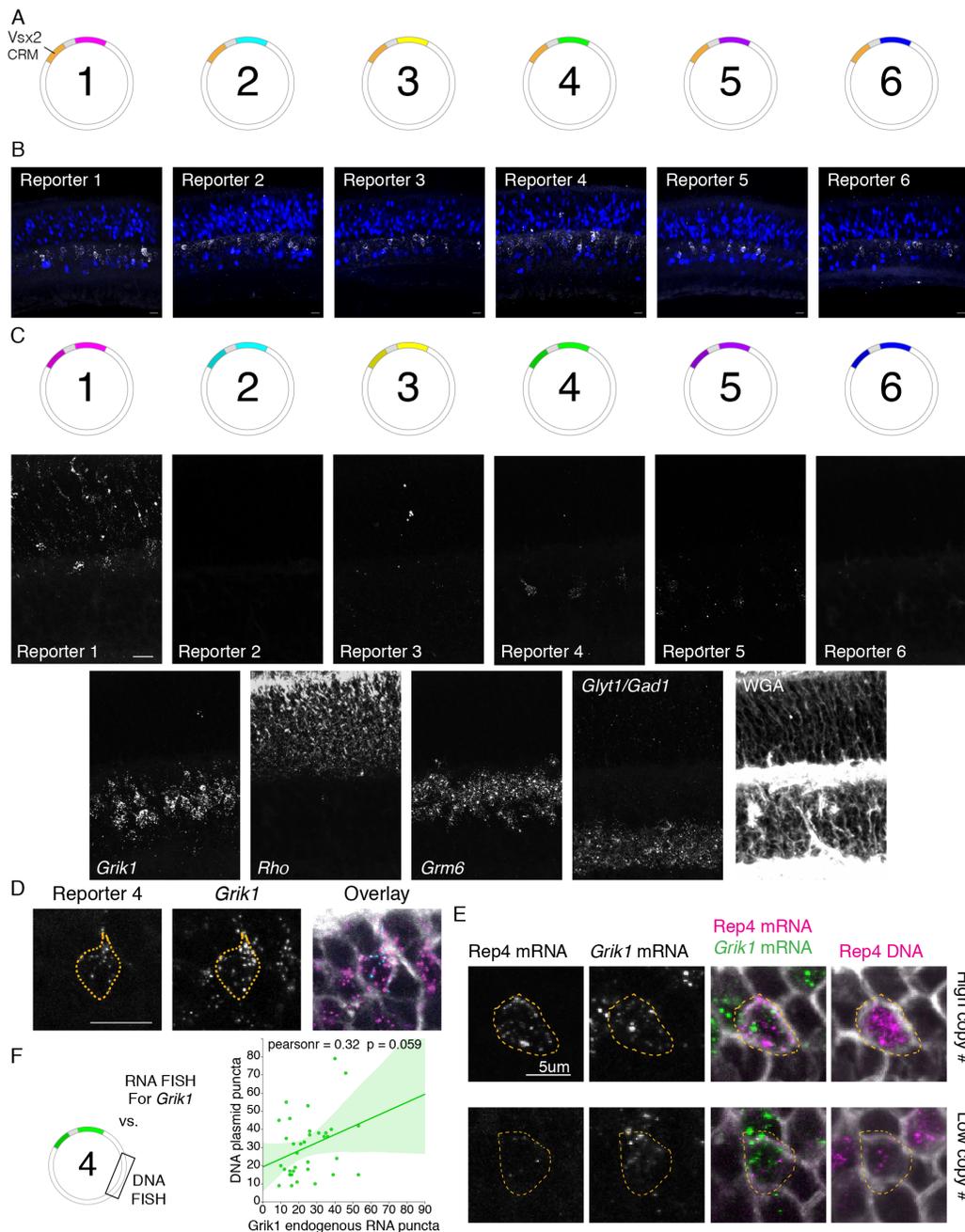


Figure S7: Related to Fig. 7. (A) Plasmid design for reporter testing, wherein each reporter is inserted downstream of a common enhancer that drives expression in bipolar cells. (B) Detection of each reporter RNA shows expression confined to the bipolar cell layer, consistent with the described activity of the Vsx2 enhancer used. Shown in blue is a co-electroporation marker (CAG-*nl*sTagBFP) for patch identification. (C) All channels detected by spectral and serial multiplexing for the experiment described in Fig. 7. *Glyt1* and *Gad1* were detected with the same concatemer to label the amacrine population, composed of both glycinergic and gabaergic cells. (D) Magnified image of a CRM-4/reporter4-expressing cell showing co-expression with *Grik1*. (E) Magnified images for sequential RNA/DNA FISH prior to puncta detection, corresponding to Fig. 7J. Dashed outlines are manually drawn to highlight representative electroporated cells. (F) Plot of plasmid copy number against *Grik1* expression shows no correlation, in contrast to the correlation observed for plasmid copy number and reporter expression (Fig. 7K). n=35 cells.

8 Supplemental Experimental Procedures

PER reaction conditions

Panel-by-panel PER extension conditions can be found in the ‘PER experiments’ tab of the SABER experiments Excel file.

FISH incubation conditions

In situ hybridization (ISH), branch hybridization, and fluorescent hybridization condition details can be found in the ‘FISH experiments’ tab of the provided SABER experiments Excel file. Z depths of the maximum intensity projections displayed for tissue images are also listed here.

Cell and puncta counts

Numbers of cells/tissues analyzed, puncta counts, and mean and standard deviations of puncta values for amplification fold enhancement experiments are reported in the provided SABER puncta counts Excel file.

Sequences

Sequences for all experiments, along with those used to model orthogonality in Fig. 1 and Fig. S1 can be found in the provided SABER sequences Excel file.

9 Supplemental References

1. B. J. Beliveau, J. Y. Kishi, G. Nir, H. M. Sasaki, S. K. Saka, S. C. Nguyen, C. Wu, and P. Yin. Oligominer provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes. *Proceedings of the National Academy of Sciences*, 2018.
2. P. J. A. Cock, T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski, and M. J. L. de Hoon. Biopython: freely available python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25(11):1422–1423, 2009.
3. R. M. Dirks and N. A. Pierce. A partition function algorithm for nucleic acid secondary structure including pseudoknots. *Journal of Computational Chemistry*, 24(13):1664–1677, 2003.
4. R. M. Dirks and N. A. Pierce. An algorithm for computing nucleic acid base-pairing probabilities including pseudoknots. *Journal of Computational Chemistry*, 25(10):1295–1304, 2004.
5. R. M. Dirks, J. S. Bois, J. M. Schaeffer, E. Winfree, and N. A. Pierce. Thermodynamic Analysis of Interacting Nucleic Acid Strands. *SIAM Review*, 49(1):65–88, 2007.
6. A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. a Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome biology*, 7(10):R100, 2006.
7. K. Shekhar, S. W. Lapan, I. E. Whitney, N. M. Tran, E. Z. Macosko, M. Kowalczyk, X. Adiconis, J. Z. Levin, J. Nemes, M. Goldman, S. A. McCarroll, C. L. Cepko, A. Regev, and J. R. Sanes. Comprehensive classification of retinal bipolar neurons by single-cell transcriptomics. *Cell*, 166(5):1308–1323.e30, 2018/07/20 2016.