Supplementary Information for:

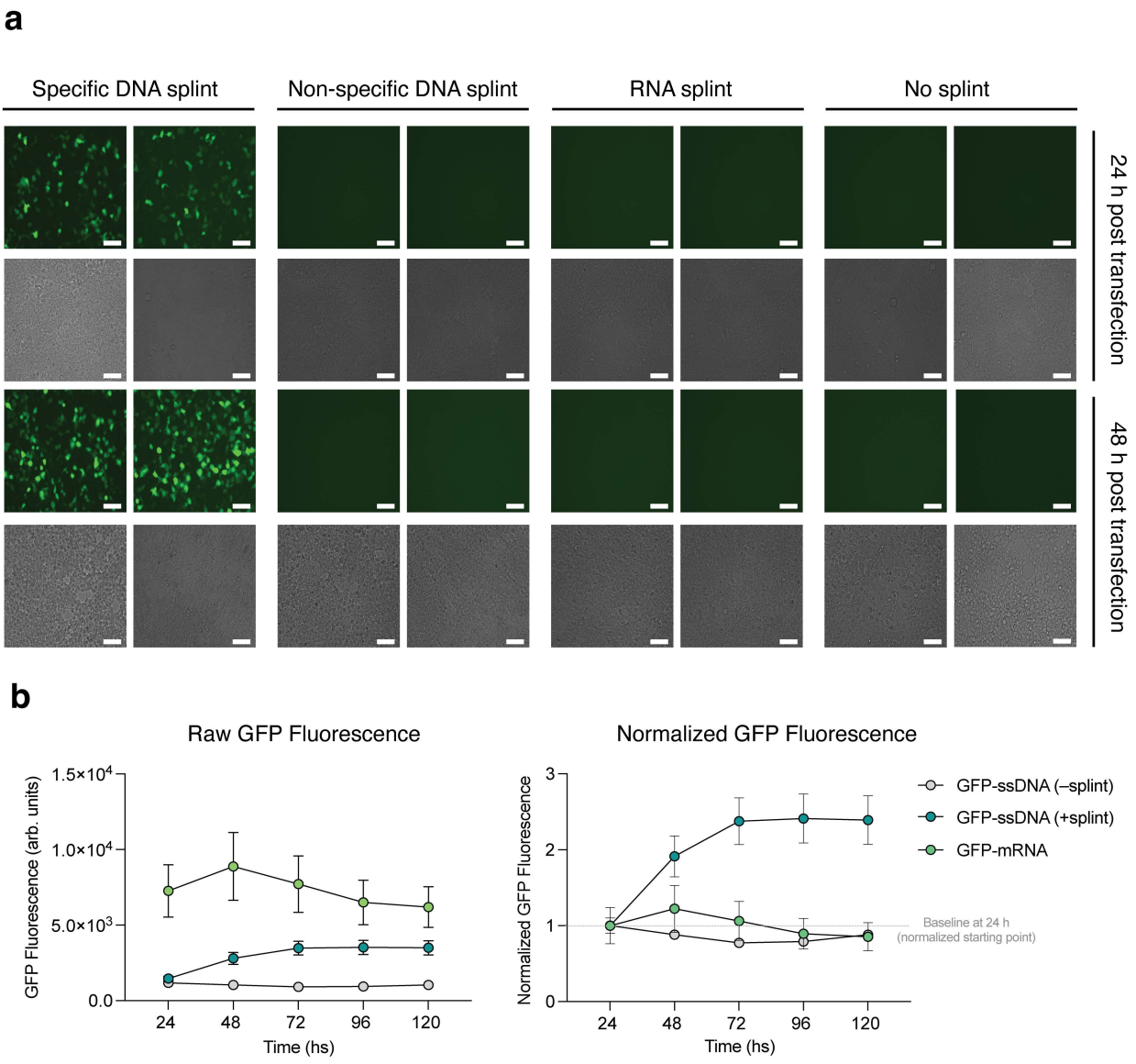
Programmable Nucleic Acid Sensing in Human Cells Using Circularizable ssDNA

Ahmed Mahas[1,2*†], Raphael Ferreira[1,2,3*], Lisa M. Riedmayr[1,2,4], George M. Church[1,2†]

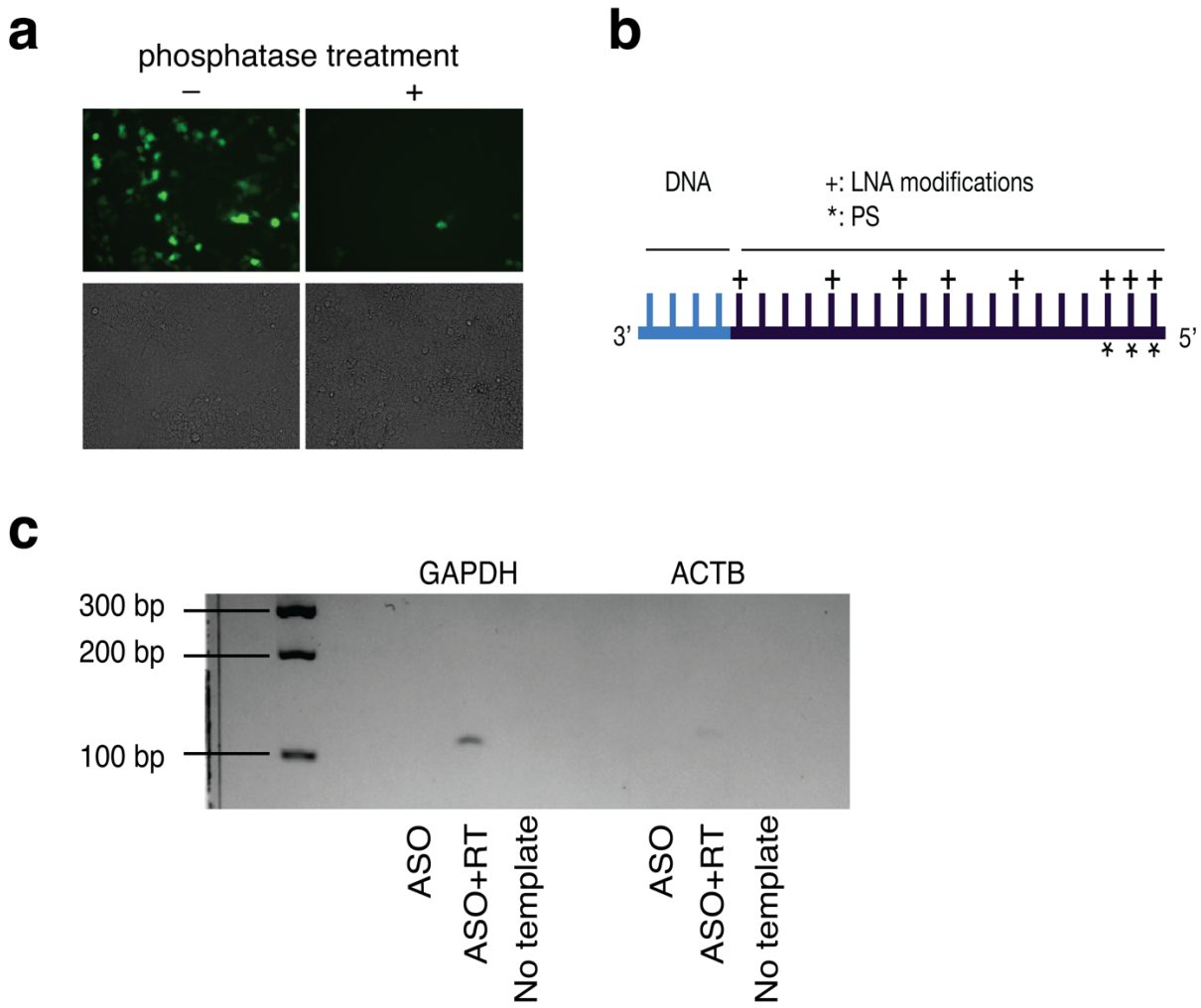
[1] Harvard Medical School, Department of Genetics, Boston, MA, USA
[2] Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA, USA
[3] Present address: Department of Health Technology, Technical University of Denmark, Kongens Lyngby, Denmark
[4] Present address: Department of Biotechnology and Biomedicine, Technical University of Denmark (DTU), Kongens Lyngby, Denmark

* These authors contributed equally to this work
† Corresponding: ahmed_mahas@hms.harvard.edu, gchurch@genetics.med.harvard.edu

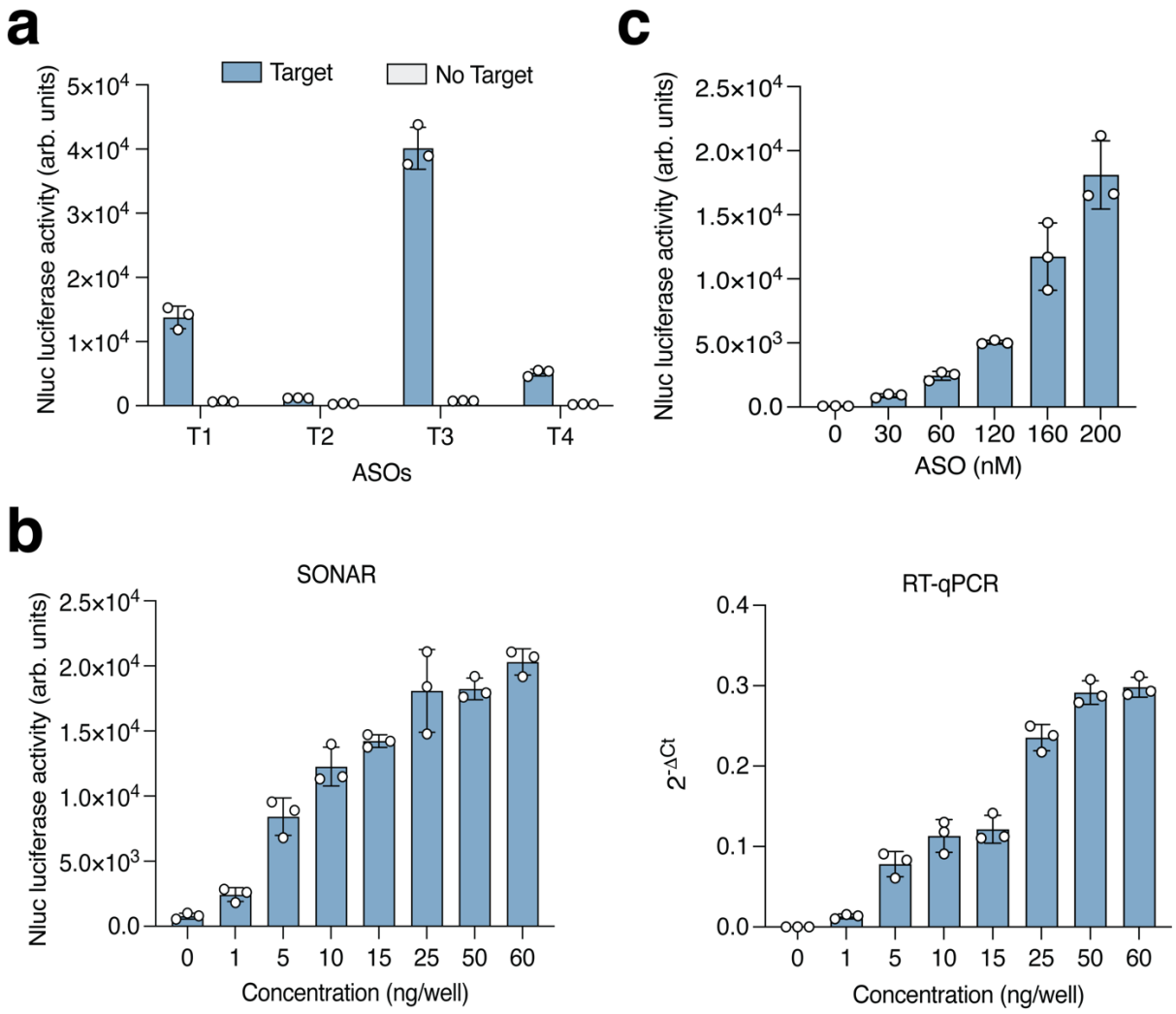
Supplementary Figures:



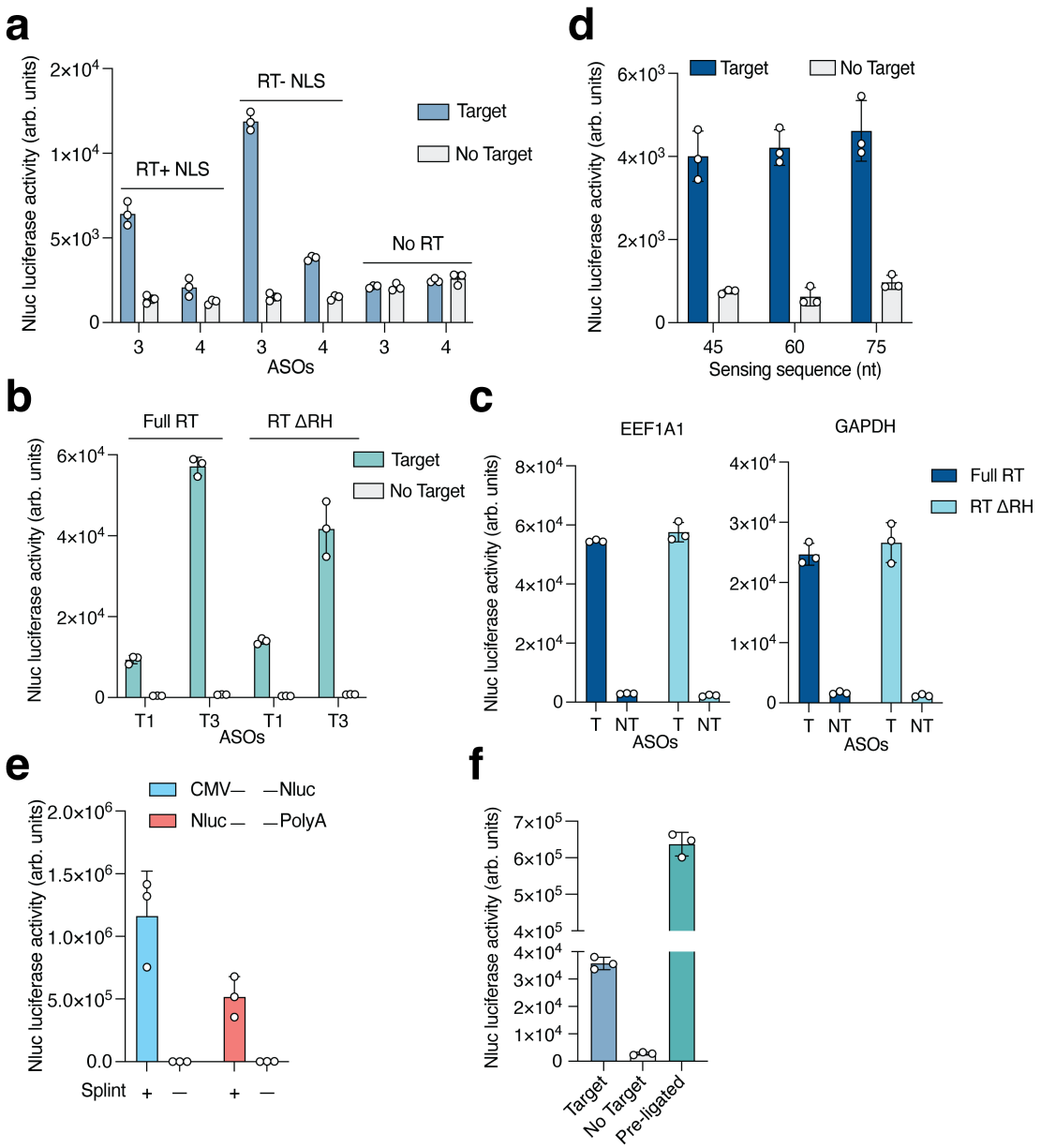
Supplementary Fig. 1. Evaluation of ssDNA sensor activity and specificity. a. Fluorescence images of HEK293T cells transfected with the ssDNA sensor and either a complementary DNA splint or control conditions after 24 h (top) and 48 h (bottom). Scale bar: 125 µm. Similar results were observed in two independent experiments. **b.** Time course of ssDNA sensor activation following transfection of GFP-ssDNA (−splint) and GFP-ssDNA (+splint). GFP mRNA was included for comparison of signal kinetics. Left: Raw GFP fluorescence over time. Right: Normalized GFP fluorescence relative to 24 h (mean ± s.d., *n* = 5 independent biological replicates). arb. units: arbitrary units. (b) Source data are provided as a Source Data file.



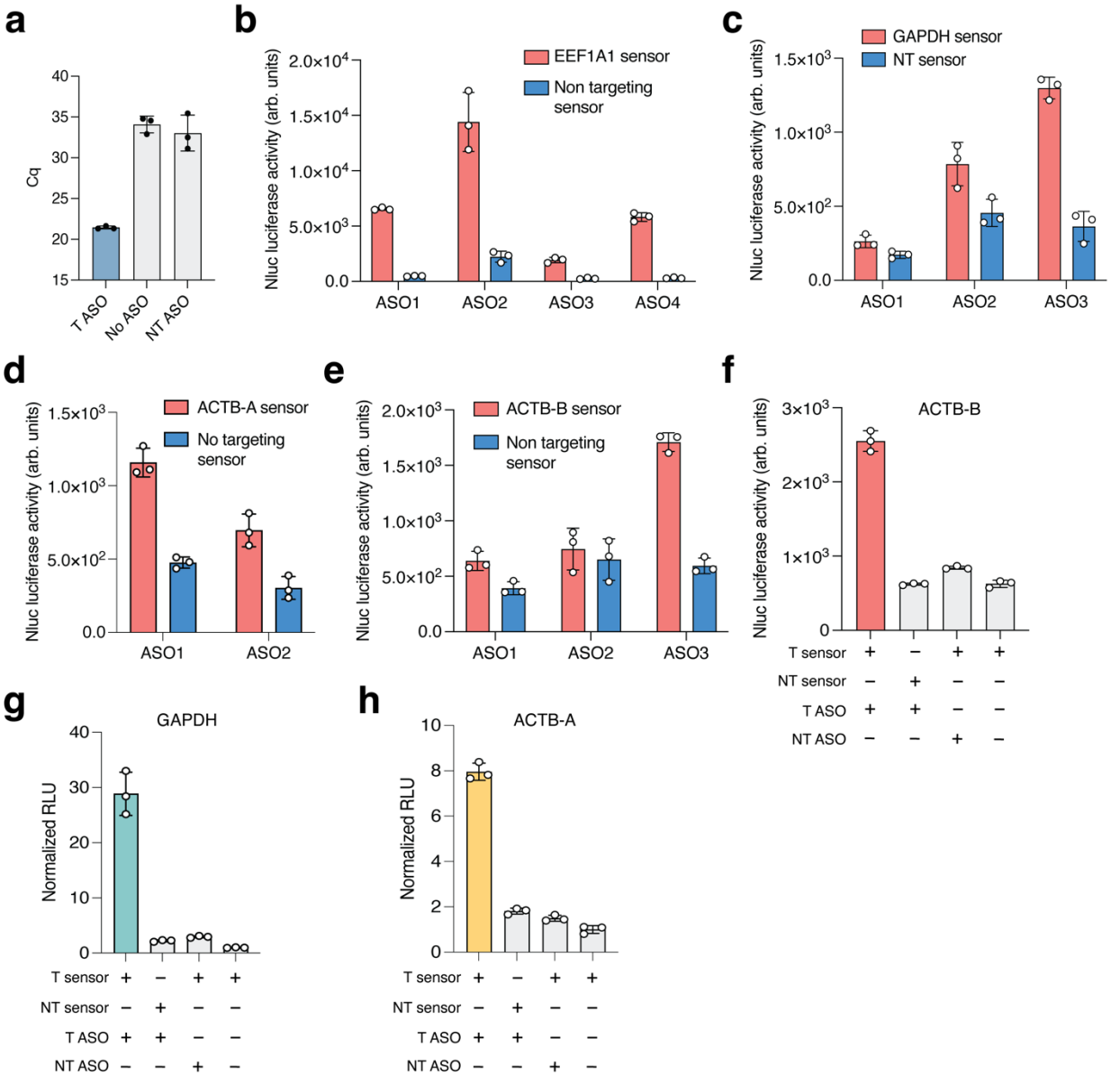
Supplementary Fig. 2. Characterization of ssDNA sensor. a. Impact of phosphatase treatment on the sensor's activity. Similar results were observed in two independent experiments. **b.** Schematic of the ASOs used in the study. **c.** Representative gel electrophoresis of PCR product of GAPDH and ACTB experiments described in Fig. 1d. Similar results were observed in three independent experiments. The uncropped gel is shown in Supplementary Fig. 14b. Panel **b** was created in BioRender. Mahas, A. (2026) https://BioRender.com/7rx0nbb.



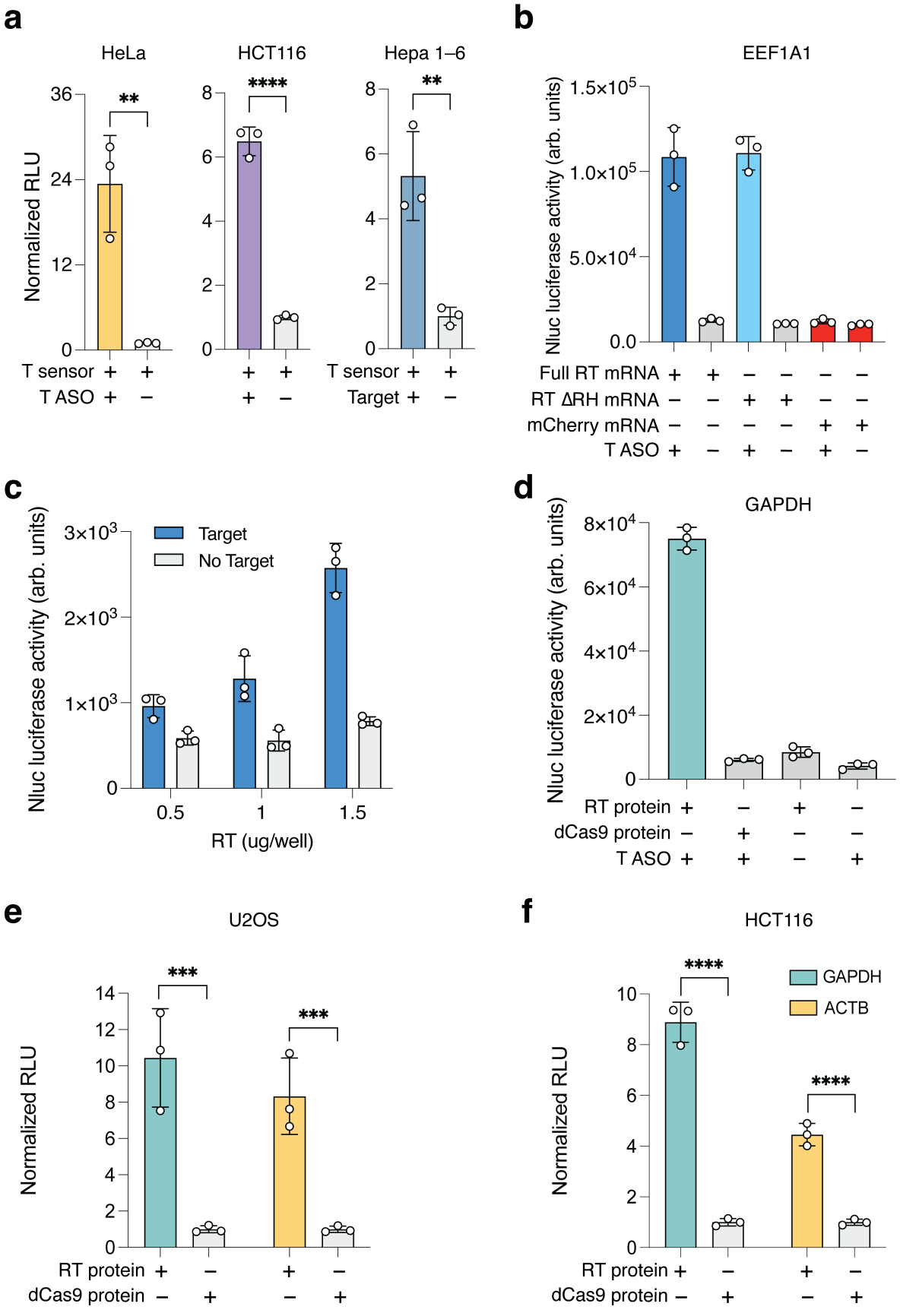
Supplementary Fig. 3. Evaluation and benchmarking of SONAR for RNA detection in human cells. Raw data of sensor efficiency in detecting target RNA in human cells (from Fig. 1f, g, and h). In a, b and c, data plotted as mean ± s.d. of *n* = 3 independent biological replicates. arb. units: arbitrary units. (a-c) Source data are provided as a Source Data file.



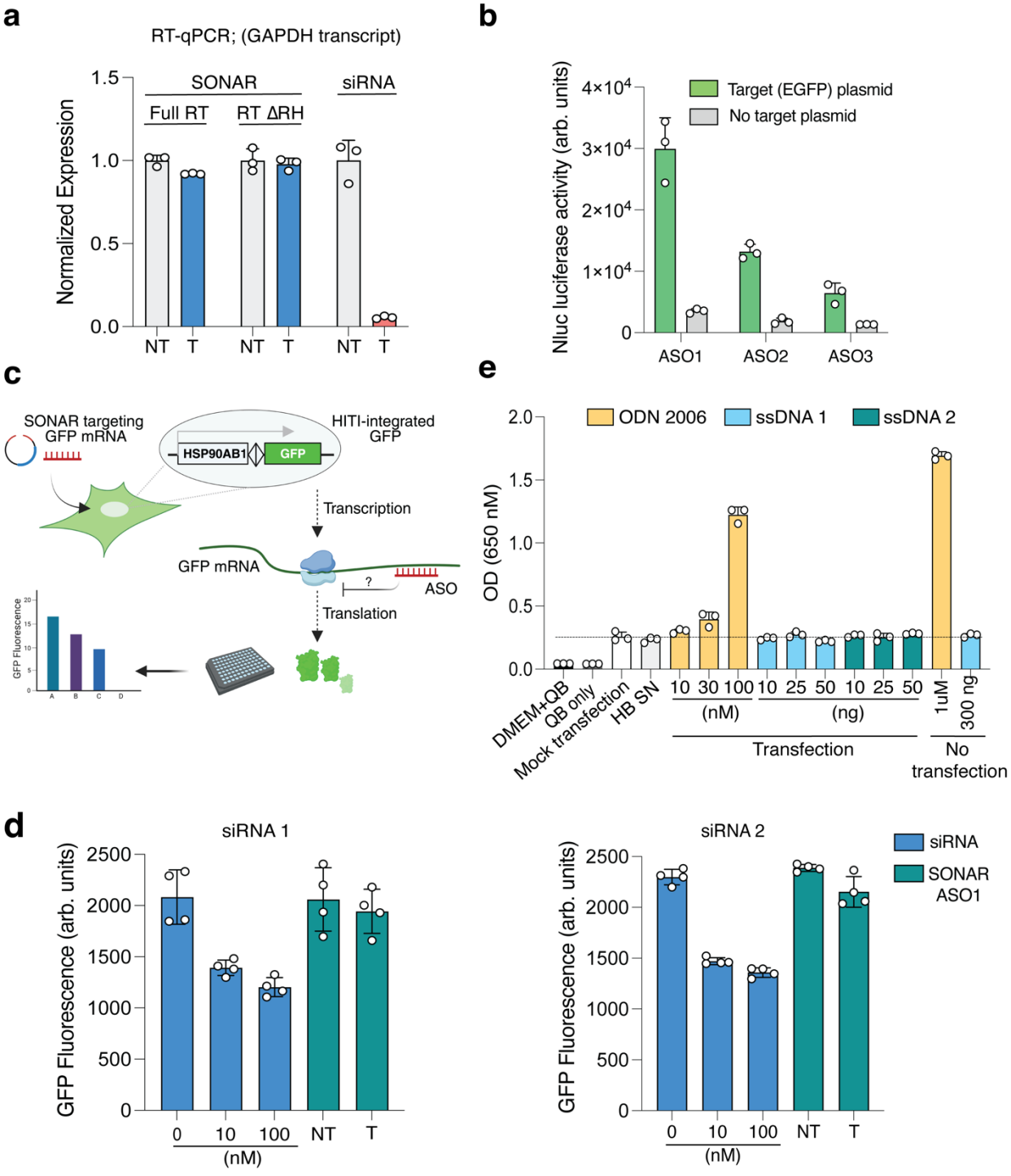
Supplementary Fig. 4. Activity of different RTs, length and positioning of the sensing sequence. a. Evaluation of the SONAR efficiency using RT with nuclear localization signal (RT+NLS), without NLS (RT-NLS), or with No RT. **b.** Evaluation of the SONAR efficiency using full RT and RT without its RNase H domain (RT ΔRH) with plasmid target, or endogenous targets (**c**). **d.** Impact of nucleotide length of the sensing sequence on SONAR activation. **e.** Quantification of SONAR activity with two different ssDNA sensor designs. CMV- -Nluc: ssDNA sensor with the sensing region placed between the promoter and the Nluc payload sequence. Nluc- -PolyA: ssDNA sensor with the sensing region placed between the Nluc payload sequence and the polyA tail sequence. **f**. Comparison of SONAR activity with a pre-ligated (constitutively active) ssDNA sensor. In **a-f**, data plotted as mean ± s.d. of *n* = 3 independent biological replicates. arb. units: arbitrary units. (a-f) Source data are provided as a Source Data file.



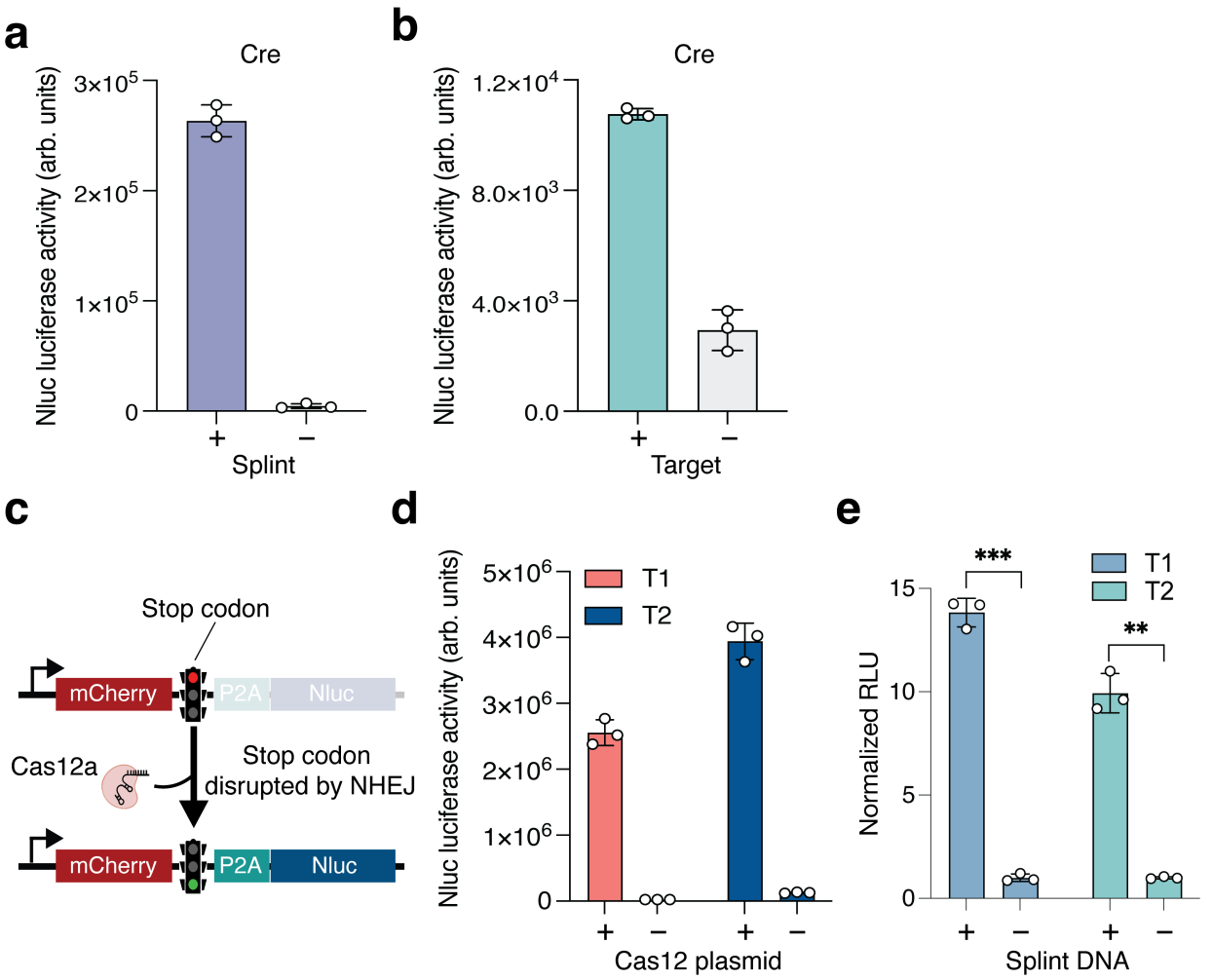
Supplementary Fig. 5. SONAR activity for detection of endogenous transcripts. a. Quantitative PCR (qPCR) analysis of in-cell-synthesized RT-DNA from a genomically integrated target. T ASO: targeting ASO, NT ASO: non targeting ASO. **b.** Raw luciferase values for screening ASO efficiency for *EEF1A1*, *GAPDH* **(c)**, *ACTB* sensor A **(d)**, and *ACTB* sensor B **(e)**. **f.** Extended control data for *ACTB*-B using ASO3, with raw luciferase measurements under various conditions. **g–h.** SONAR activity for detection of endogenous *GAPDH* with ASO3 (**g**) and *ACTB*-A with ASO1 (**h**) transcripts measured at 72 h post-transfection. For **a**-**h**, data plotted as mean ± s.d. of $n = 3$ independent biological replicates. Data in **b**, **c**, **d** and **e** were collected 24 h after SONAR transfection. arb. units: arbitrary units. RLU: Relative Luminescence Units. (a-h) Source data are provided as a Source Data file.


a
HeLa
HCT116
Hepa 1–6
Normalized RLU
36
24
12
0
8
6
4
2
0
**
****
**
T sensor + +
T ASO + –
+ +
+ –
T sensor + +
Target + –
b
EEF1A1
Nluc luciferase activity (arb. units)
1.5×10^5
1.0×10^5
5.0×10^4
0.0
Full RT mRNA + + – – – –
RT ΔRH mRNA – – + + – –
mCherry mRNA – – – – + +
T ASO + – + – + –
c
Nluc luciferase activity (arb. units)
3×10^3
2×10^3
1×10^3
0
Target
No Target
0.5
1
1.5
RT (ug/well)
d
GAPDH
Nluc luciferase activity (arb. units)
8×10^4
6×10^4
4×10^4
2×10^4
0
RT protein + – + –
dCas9 protein – + – –
T ASO + + – +
e
U2OS
Normalized RLU
14
12
10
8
6
4
2
0
***
***
RT protein + – + –
dCas9 protein – + – +
f
HCT116
Normalized RLU
10
8
6
4
2
0
****
****
GAPDH
ACTB
RT protein + – + –
dCas9 protein – + – +

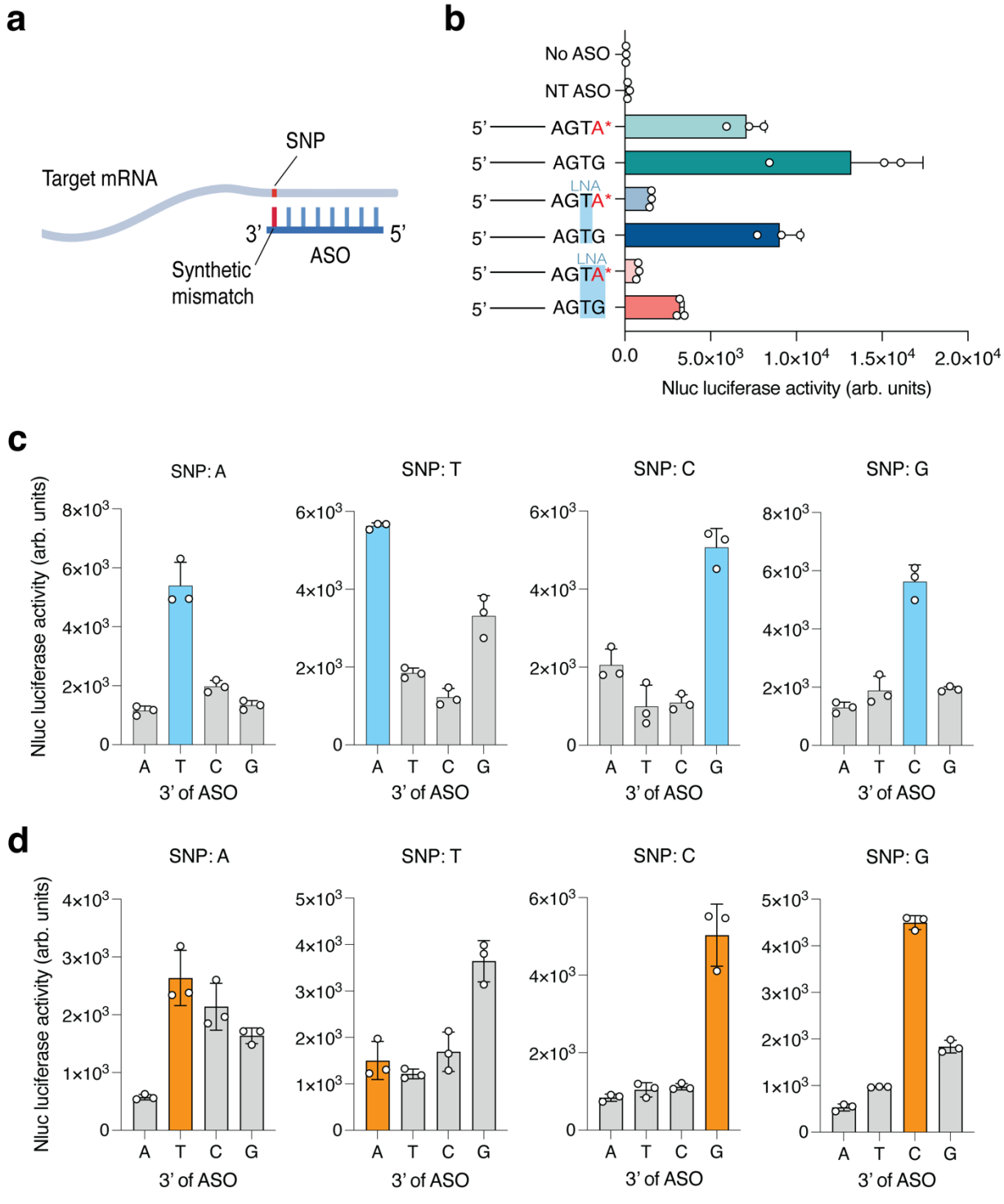
Supplementary Fig. 6. Validation of SONAR activity across multiple cell types and RT delivery formats. a. SONAR activity in HeLa (targeting *ACTB-A*), HCT116 (targeting *EEF1A1*), and Hepa 1–6 cells (plasmid target). Statistical significance determined by two-tailed lognormal Welch's *t*-test; HeLa: ***P* = 0.0021, HCT116: *****P* < 0.0001, Hepa 1–6: ***P* = 0.0019. **b.** SONAR activation with RT supplied as mRNA; *mCherry* mRNA was used as a negative control. The data in Fig. 2f were generated from this experiment. **c.** Dose-dependent activation of SONAR in response to increasing amounts of RT protein using a target expressed from a plasmid. **d.** SONAR activation with RT supplied as purified protein when targeting the endogenous *GAPDH* transcript; dCas9 protein was used as a negative control. The data in Fig. 2g were generated from this experiment. **e, f.** SONAR activity driven by RT-protein delivery in U2OS (**e**) and HCT116 (**f**) cells, targeting *GAPDH* and *ACTB* transcripts. Statistical significance determined by two-tailed lognormal Welch's *t*-test; U2OS: ****P* = 0.0005; HCT116: *****P* < 0.0001. In **a-f**, data are shown as mean ± s.d. (*n* = 3 biological replicates). T sensor, targeting sensor; T ASO, targeting antisense oligonucleotide. arb. units: arbitrary units. RLU: Relative Luminescence Units. (a-f) Source data are provided as a Source Data file.


a
RT-qPCR; (GAPDH transcript)
Normalized Expression
SONAR
Full RT
RT ΔRH
siRNA
NT
T
b
Nluc luciferase activity (arb. units)
Target (EGFP) plasmid
No target plasmid
ASO1
ASO2
ASO3
c
SONAR targeting GFP mRNA
HITI-integrated GFP
HSP90AB1
GFP
Transcription
GFP mRNA
ASO
Translation
GFP Fluorescence
e
ODN 2006
ssDNA 1
ssDNA 2
OD (650 nM)
DMEM+QB
QB only
Mock transfection
HB SN
(nM)
(ng)
Transfection
No transfection
1uM
300 ng
d
siRNA 1
siRNA 2
GFP Fluorescence (arb. units)
siRNA
SONAR ASO1

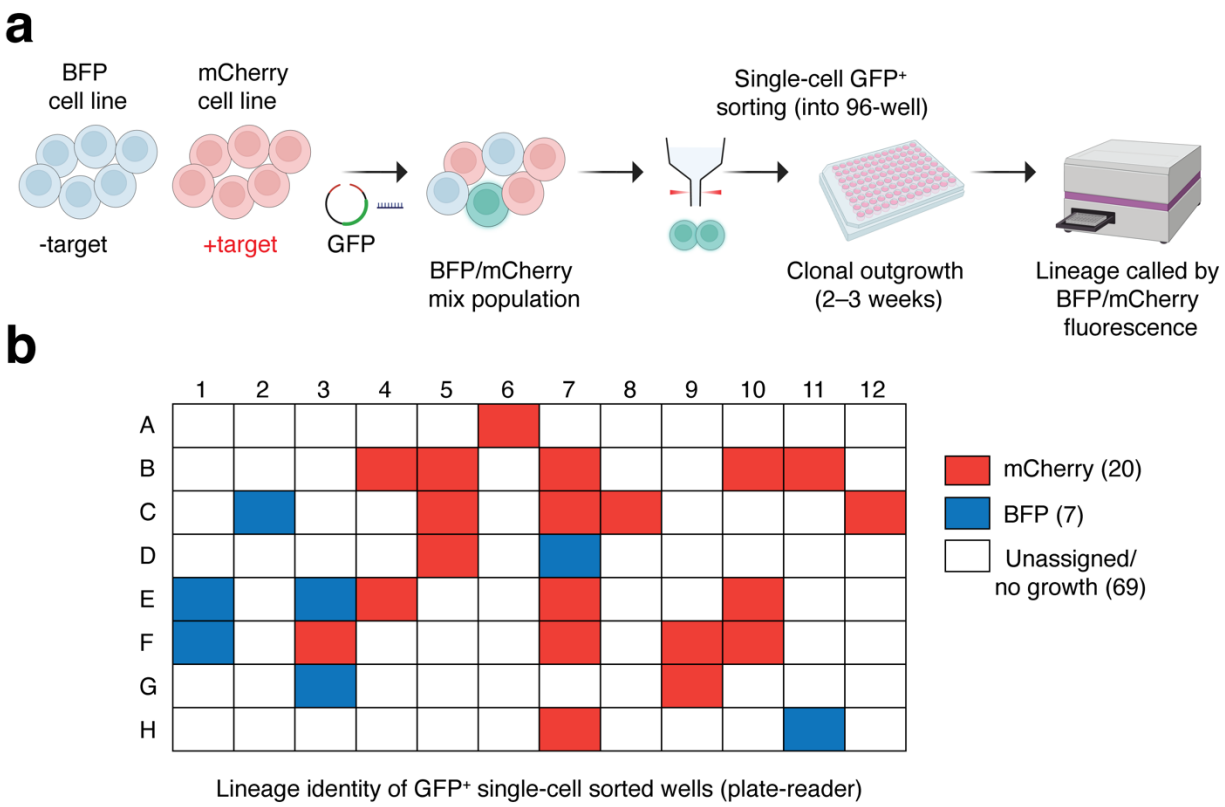
Supplementary Fig. 7. Evaluation of SONAR effects on target RNA expression, translation, and cellular response. a. RT-qPCR analysis of *GAPDH* mRNA levels in cells transfected with SONAR or siRNA constructs targeting *GAPDH*. Full-length RT or truncated RT lacking the RNase H domain (ΔRH) variants were used as indicated. Expression was normalized to non-targeting (NT) controls (NT ASO and ssDNA for SONAR; NT siRNA for siRNA experiments). **b.** Screening of ASO efficiency for SONAR targeting a plasmid-encoded GFP reporter. **c.** Schematic of design and experimental workflow for assessing SONAR's effect on mRNA target translation. A GFP reporter was genomically integrated via HITI in-frame with the *HSP90AB1* gene, and its mRNA transcripts served as targets for SONAR or siRNA. **d.** Results from the experimental design illustrated in (**c**), comparison of siRNA and SONAR (using ASO1 from **b**) effects on GFP fluorescence across a range of siRNA concentrations, showing that SONAR detection does not reduce target expression. Data are shown as mean ± s.d. (n = 4 biological replicates). **e.** Assessment of innate immune activation in HEK-Blue (hTLR9) cells transfected with SONAR ssDNA components compared to the positive control ODN2006 (TLR9 agonist). QB, QUANTI-Blue™; HB SN, HEK-Blue supernatant. In **a, b,** and **e,** data are shown as mean ± s.d. (n = 3 biological replicates). arb. units: arbitrary units. (a-b, d-e) Source data are provided as a Source Data file. Panel **c** was created in BioRender. Mahas, A. (2026) https://BioRender.com/7rx0nbb.



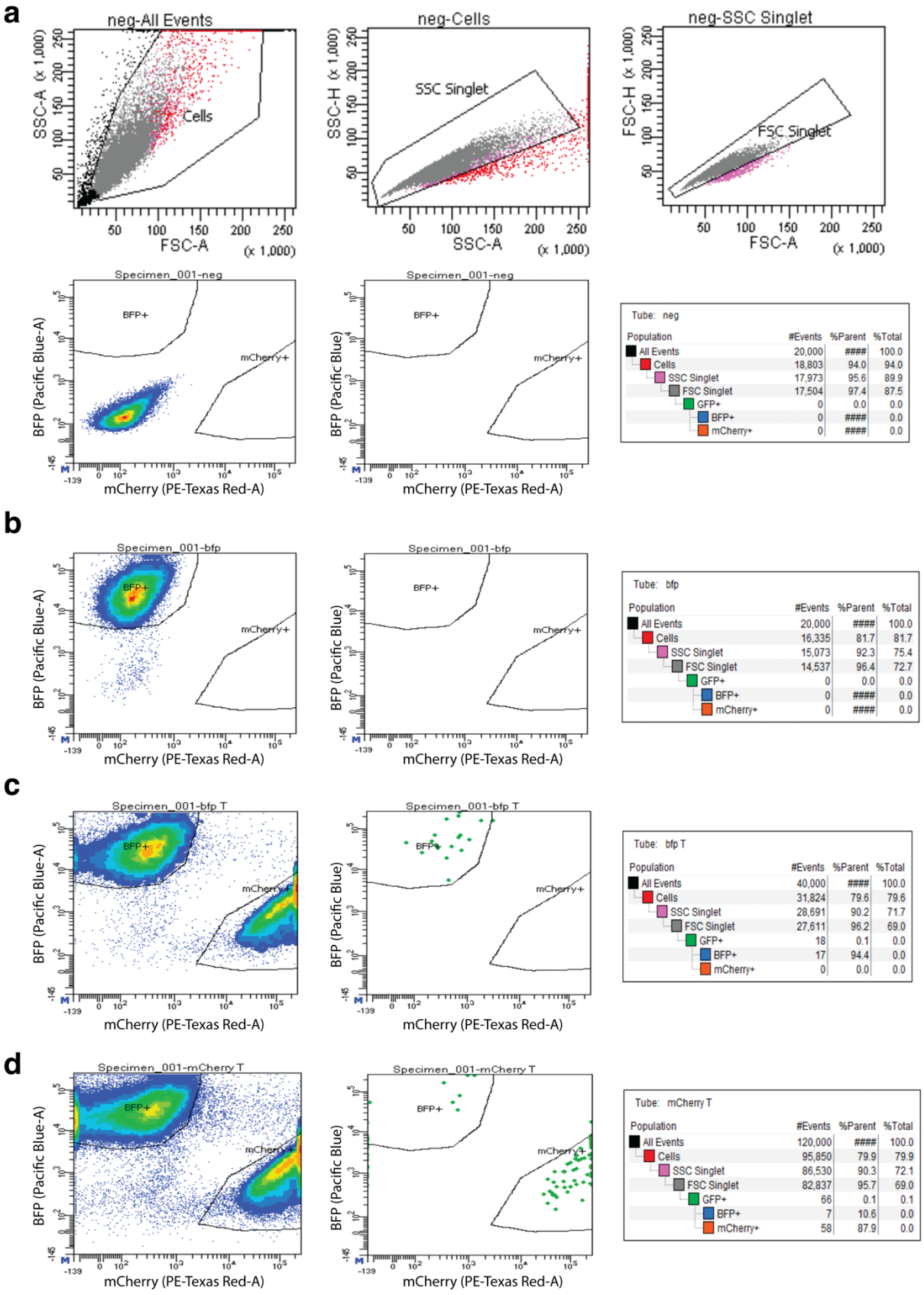
Supplementary Fig. 8. SONAR Applications. a. Assessment of the activity and target responsiveness of Cre-containing ssDNA sensor, evaluated through sensor-driven Cre expression in response to a specific splint DNA sequence in human cells. Activation was measured via Nluc reporter luciferase activity, with luminescence readings collected 24 h post-transfection. **b.** Quantification of sensor-driven Cre expression in detecting plasmid-expressed target RNA in human cells, measured as luciferase activity. The data were collected 24 h after SONAR transfection. **c.** Schematic of the Cas12a reporter plasmid, where a stop codon prevents downstream Nluc expression. Upon Cas12a and CRISPR RNA (crRNA) expression, targeted cutting near the stop codon followed by non-homologous end joining (NHEJ) repair can restore the Nluc reading frame. **d.** Assessment of the Cas12 Nluc reporter functionality using Cas12-expressing plasmid and two different reporter plasmids containing different Cas12 target sites (T1 and T2) and a Nluc sequence. Luciferase activity was measured to evaluate Cas12-induced reporter activation with luminescence readings collected 48 h post-transfection. **e.** Quantification of sensor efficiency in cells with or without the specific splint DNA. Significance determined by unpaired two-tailed Welch's *t*-test. $^{**}P = 0.0037$, $^{***}P = 0.0005$. In **a-e**, data plotted as mean ± s.d. of $n = 3$ independent biological replicates. arb. units: arbitrary units. RLU: Relative Luminescence Units. (a-b, d-e) Source data are provided as a Source Data file.


a
Target mRNA
SNP
3'
5'
ASO
Synthetic
mismatch
b
No ASO
NT ASO
5' AGTA*
5' AGTG
LNA
5' AGTA*
5' AGTG
LNA
5' AGTA*
5' AGTG
0.0
5.0×10^3
1.0×10^4
1.5×10^4
2.0×10^4
Nluc luciferase activity (arb. units)
c
SNP: A
SNP: T
SNP: C
SNP: G
Nluc luciferase activity (arb. units)
A T C G
3' of ASO
d
SNP: A
SNP: T
SNP: C
SNP: G
Nluc luciferase activity (arb. units)
A T C G
3' of ASO

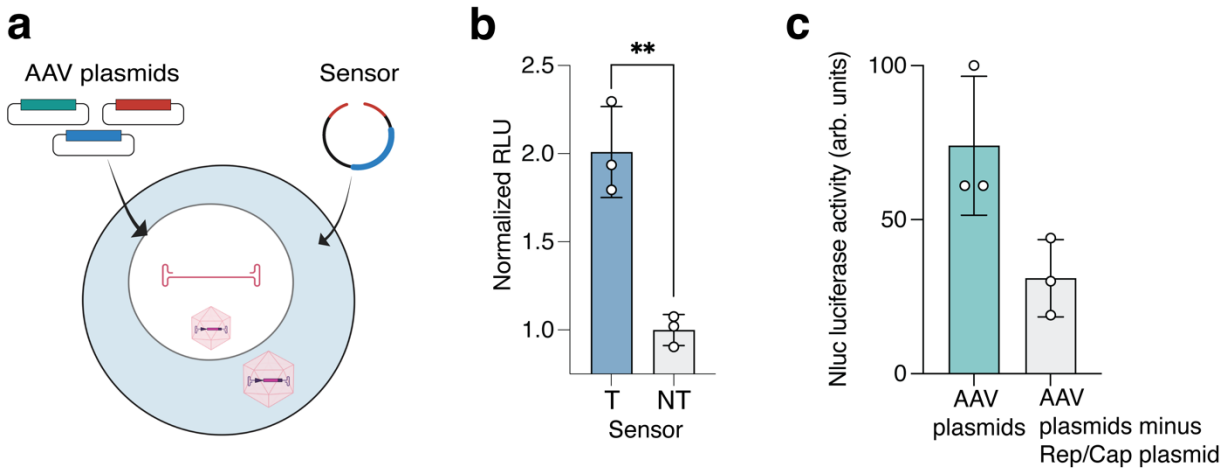
Supplementary Fig. 9. Evaluation of SONAR SNP discrimination activity. a. Schematic illustrating SONAR's SNP detection strategy, in which a single synthetic mismatch is introduced in the ASO near the SNP site to enhance discrimination between variant and wild-type sequences**. b.** Impact of mismatches and LNA modifications on sensing specificity. Mismatches highlighted in red with an asterisk. LNA modifications are highlighted in light blue. Data represent sensor efficiency with ASOs containing mismatches relative to the corresponding perfectly matched ASO. The data were collected 24 h after SONAR transfection. **c-d.** Evaluation of SONAR SNP discrimination in detecting SNP variants (A, T, C, or G) for target 1 (**c**) and target 2 (**d**) using ASOs with matched or synthetic mismatched 3′ bases. Data in (**c**) represent the raw data underlying Fig. 2i. In **b-d**, data are shown as mean ± s.d. (*n* = 3 biological replicates). arb. units: arbitrary units. (b-d) Source data are provided as a Source Data file. Panel **a** was created in BioRender. Mahas, A. (2026) https://BioRender.com/7rx0nbb.



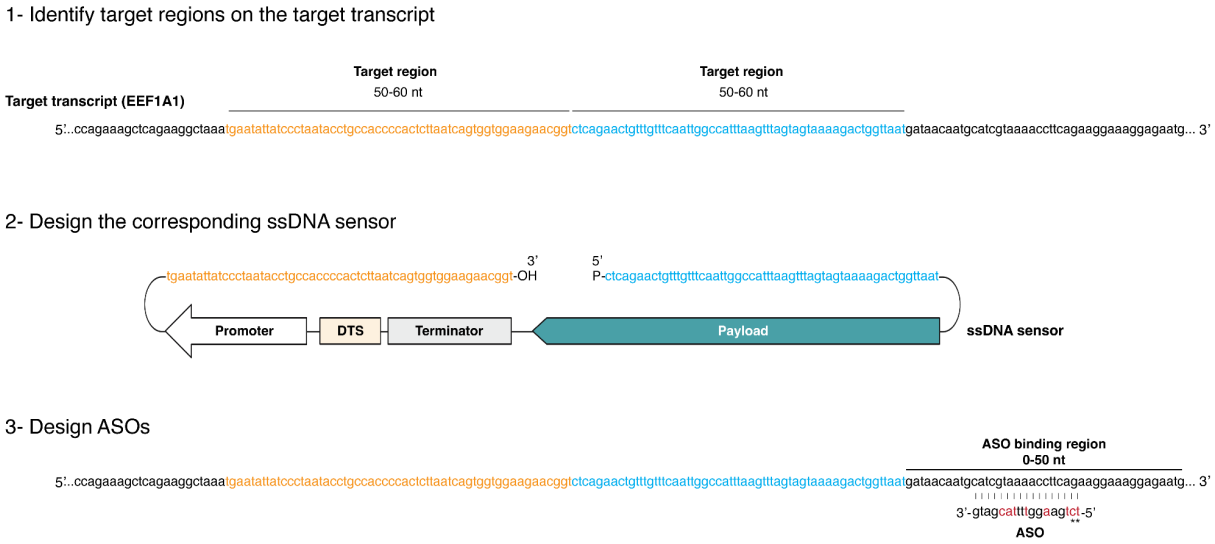
Supplementary Fig. 10. Sequence-guided single-cell isolation following SONAR activation. **a.** Workflow schematic. The RNA target was introduced into the mCherry cell line, and mCherry (+target) cells were mixed with BFP (–target) cells. After SONAR delivery, target-dependent SONAR activation drove GFP reporter expression, and GFP$^+$ cells were single-cell sorted into 96-well plates. After 2–3 weeks of incubation, wells were read by plate-reader fluorescence in the BFP and mCherry channels to assign cell line identity. **b.** Plate map of cell line calls for wells seeded with single GFP$^+$ sorted cells. Wells classified as mCherry (red, n = 20) or BFP (blue, n = 7) met the cell line-calling criteria based on plate-reader fluorescence. Wells labeled unassigned/no growth (white, n = 69) had BFP and mCherry signals below the calling thresholds, consistent with no detectable outgrowth, and were not assigned to either cell line. (b) Source data are provided as a Source Data file. Panel **a** was created in BioRender. Mahas, A. (2026) https://BioRender.com/7rx0nbb.


a
neg-All Events
SSC-A (x 1,000)
FSC-A (x 1,000)
Cells
neg-Cells
SSC-H (x 1,000)
SSC-A (x 1,000)
SSC Singlet
neg-SSC Singlet
FSC-H (x 1,000)
FSC-A (x 1,000)
FSC Singlet
Specimen_001-neg
BFP+
mCherry+
BFP (Pacific Blue-A)
mCherry (PE-Texas Red-A)
Tube: neg
Population #Events %Parent %Total
All Events 20,000 #### 100.0
Cells 18,803 94.0 94.0
SSC Singlet 17,973 95.6 89.9
FSC Singlet 17,504 97.4 87.5
GFP+ 0 0.0 0.0
BFP+ 0 #### 0.0
mCherry+ 0 #### 0.0
b
Specimen_001-bfp
Tube: bfp
Population #Events %Parent %Total
All Events 20,000 #### 100.0
Cells 16,335 81.7 81.7
SSC Singlet 15,073 92.3 75.4
FSC Singlet 14,537 96.4 72.7
GFP+ 0 0.0 0.0
BFP+ 0 #### 0.0
mCherry+ 0 #### 0.0
c
Specimen_001-bfp T
Tube: bfp T
Population #Events %Parent %Total
All Events 40,000 #### 100.0
Cells 31,824 79.6 79.6
SSC Singlet 28,691 90.2 71.7
FSC Singlet 27,611 96.2 69.0
GFP+ 18 0.1 0.0
BFP+ 17 94.4 0.0
mCherry+ 0 0.0 0.0
d
Specimen_001-mCherry T
Tube: mCherry T
Population #Events %Parent %Total
All Events 120,000 #### 100.0
Cells 95,850 79.9 79.9
SSC Singlet 86,530 90.3 72.1
FSC Singlet 82,837 95.7 69.0
GFP+ 66 0.1 0.1
BFP+ 7 10.6 0.0
mCherry+ 58 87.9 0.0

Supplementary Fig. 11. Flow cytometry gating strategy for SONAR reciprocal target-swap enrichment. a. Negative control defining the gating hierarchy. **b.** Single-color BFP control. **c.** Target in BFP (BFP T): GFP^+ events map predominantly to the BFP^+ gate. **d.** Target in mCherry (mCherry T): GFP^+ events map predominantly to the $mCherry^+$ gate. Percentages shown are within the GFP^+ gate.

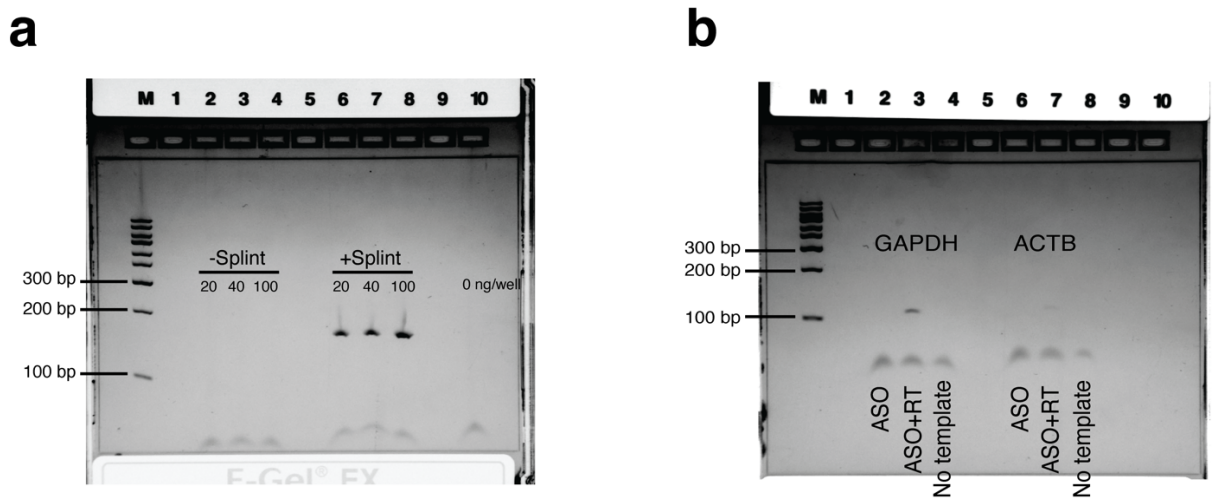


Supplementary Fig. 12. Exploratory evaluation of SONAR for viral ssDNA detection. **a.** AAV plasmids indicate the three AAV production plasmids needed for AAV replication, including transgenic gene (ITR) plasmid with target sequence, Cap/Rep plasmid, and helper plasmid. **b-c.** Quantification of sensor efficiency in detecting AAV genomic sequences in human cells. AAV plasmids minus Rep/Cap control in (**c**) are cells transfected with AAV plasmids, but without Rep/Cap plasmid, which is essential for AAV replication. Significance, in (**b**), determined by unpaired two-tailed *t*-test between targeting (T) and non-targeting (NT) sensor control, $^{**}P = 0.0031$. Data in **b and c** are plotted as mean ± s.d. of $n = 3$ independent biological replicates. RLU: Relative Luminescence Units. arb. units: arbitrary units. (b-c) Source data are provided as a Source Data file. Panel **a** was created in BioRender. Mahas, A. (2026) https://BioRender.com/7rx0nbb.



Supplementary Fig. 13. Design workflow and practical guidelines for SONAR sensor construction.

(1) Identify target regions on the transcript. Select 50–70 nt regions within the target transcript as SONAR recognition sites. Because these sequences form the 5′ UTR of the ssDNA sensor, it is advisable to avoid regions containing start codons (ATG) that could interfere with payload translation; if unavoidable, mutate a single base (e.g., T→G) to disrupt the start codon. Choose regions with balanced GC content (40–60%) and minimal predicted secondary structure. Screening a small set of candidates (2–4 per target) is typically sufficient to identify a robust site. **2)** Design the ssDNA sensor. The ssDNA sensor is designed as a linear, circularizable DNA construct in which the target region from Step 1 is incorporated as two sensing sequences (arms), shown in orange and blue, that flank the ligation junction. The arms retain the same 5′→3′ orientation as the target RNA, as they hybridize to the complementary RT-DNA generated during the ASO/RT reaction. The construct should include a 5′-phosphate modification on the end to enable ligation by endogenous ligases. Each SONAR sensor follows the 5′→3′ organization: 5′ sensing arm → payload → terminator (e.g., SV40 poly(A) signal) → DNA Targeting Sequence (DTS; SV40 72-bp enhancer) → promoter → 3′ sensing arm. This configuration preserves correct transcriptional orientation and ensures efficient hybridization and ligation. **3)** ASOs are designed to hybridize downstream of the target regions on the transcript. Each ASO is typically 18–24 nt in length and incorporates LNA (red nucleotides) and phosphorothioate (PS; asterisks) modifications for stability and binding affinity. The 3′ terminal four nucleotides should remain unmodified to function as a primer for reverse transcription. ASOs positioned closer to the sensing-arm hybridization region tend to perform better; therefore, we recommend selecting target sites within ~1–40 (up to 50) nt downstream of the sensing sequence. Screening 2–4 ASO candidates per target is usually sufficient to identify the most sensitive design with optimal signal-to-noise performance. Longer ASOs (22–24 nt) may increase efficiency but can also raise background, so initial testing with shorter designs (18–20 nt) is advisable. To minimize sequence-dependent structural effects such as self-dimerization or strong secondary structures, ASO candidates should be evaluated using oligo analysis tools (e.g., IDT OligoAnalyzer) to predict self-annealing and hairpin formation before synthesis.



Supplementary Fig. 14. Uncropped gel images. a. Uncropped gel electrophoresis of PCR-amplified sensor described in Fig. 1c. **b.** Uncropped gel electrophoresis of PCR product of GAPDH and ACTB experiments described in Fig. 1d.