Supplemental information

Supplementary Figures

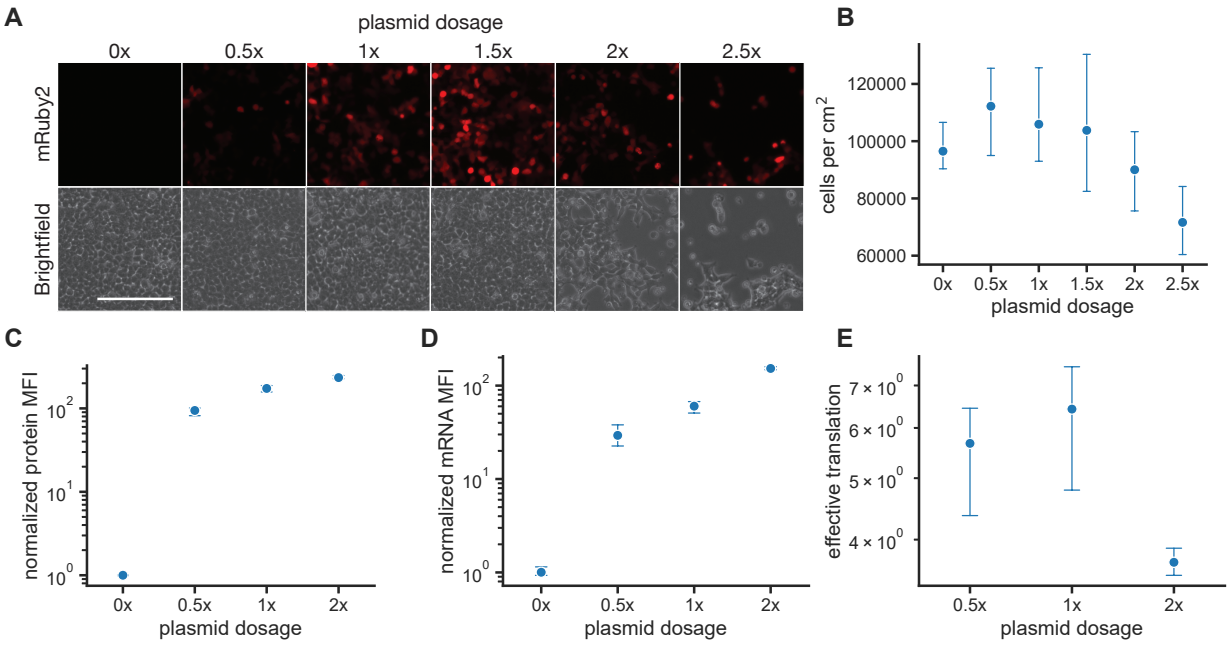

Figure S1. Effects of plasmid dosage on cell viability and expression.
A. Images of HEK293T cells transfected with varying amounts of a CAG-mRuby2-bGH plasmid. "1x" represents the plasmid dosage used in all other transfection experiments. Scale bar represents 200 µm.
B. Cell counts per cubic centimeter for varying plasmid transfection dosage as determined by flow cytometry.
C, **D**. Normalized geometric mean of protein (fig. S1C) and mRNA (fig. S1D) fluorescence for mRuby2 plasmid transfection with varying plasmid dosage.
E. Effective translation rate for a CAG-mRuby2-bGH plasmid transfected at varying dosages.
Normalized expression is calculated as the fold change of fluorescence intensity relative to a non-transfected sample. Points represent means of three biological replicates, and error bars represent the 95% confidence interval. All data is in arbitrary units from a flow cytometer.

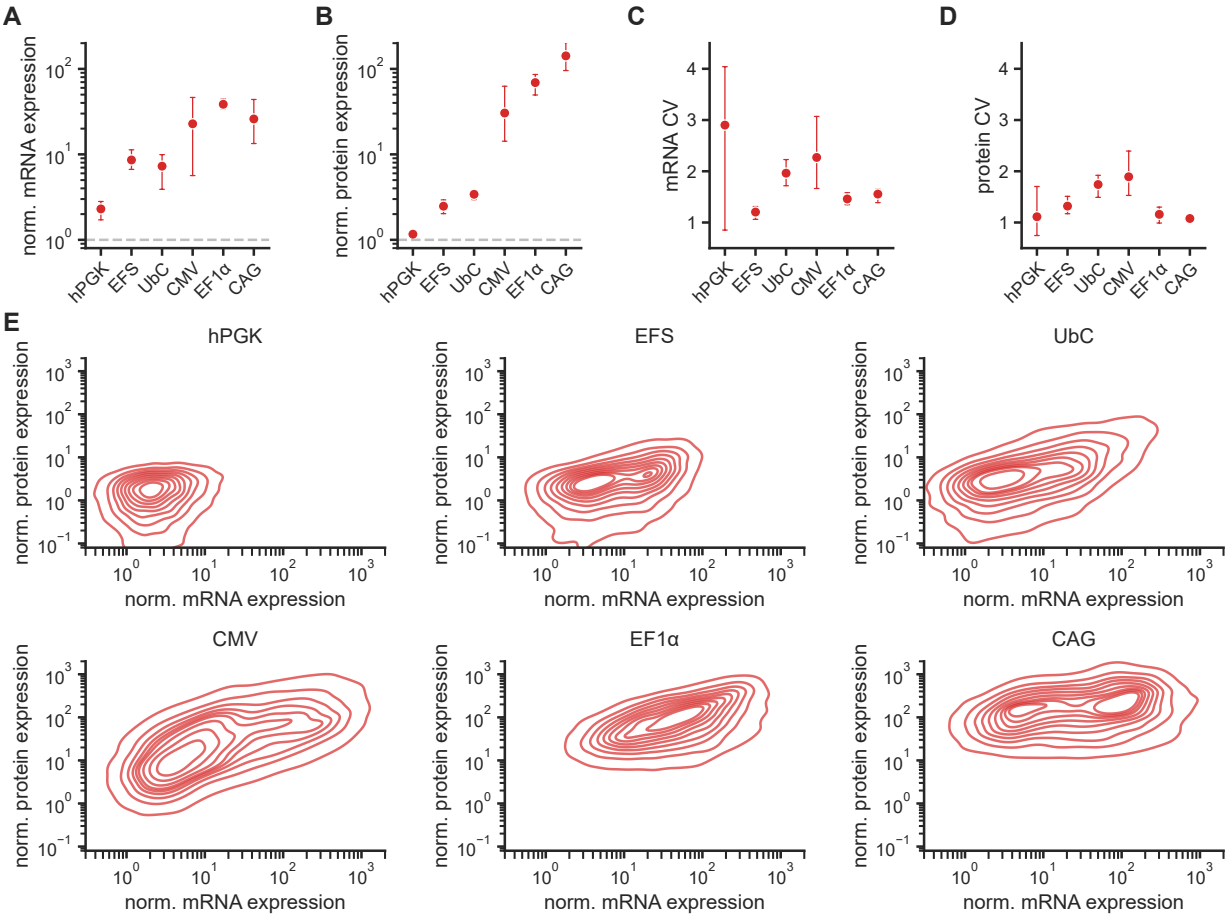

Figure S2. Additional mRuby2 transfection characterization results.

A. Normalized mRuby2 mRNA expression in HEK293T transfection with varying promoter sequences. Dashed line indicates background fluorescence level.

B. Normalized mRuby2 protein expression in HEK293T transfection with varying promoter sequences. Dashed line indicates background fluorescence level.

C. mRuby2 mRNA coefficient of variation in HEK293T transfection with varying promoter sequences.

D. mRuby2 protein coefficient of variation in HEK293T transfection with varying promoter sequences.

E. Joint distributions of mRuby2 mRNA and protein fluorescence in HEK293T transfection with varying promoter sequences. Data was randomly downsampled to 10,000 cells per condition for plotting.

Normalized expression is calculated as the fold change of fluorescence intensity relative to a non-transfected sample. Points represent means of three biological replicates, and error bars represent the 95% confidence interval. All data is in arbitrary units from a flow cytometer.

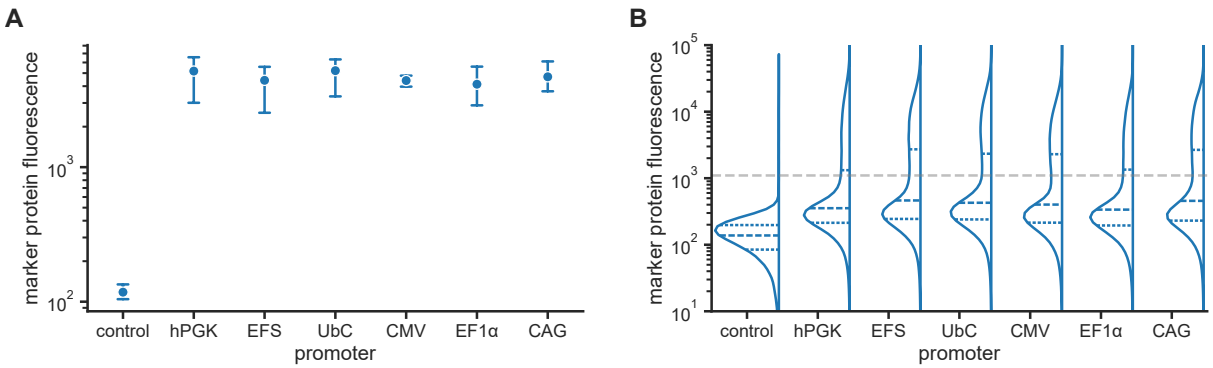

Figure S3. Marker gene expression data for experiments presented in fig. 2.
A. Geometric mean of marker gene expression (tagBFP) in transfected cells presented in fig. 2. Points represent means of three biological replicates, and error bars represent the 95% confidence interval.
B. Distribution of marker gene expression (tagBFP) across all cells. Grey dashed line indicates the marker expression gate used to identify transfected cells for further analysis.
All data is in arbitrary units from a flow cytometer.

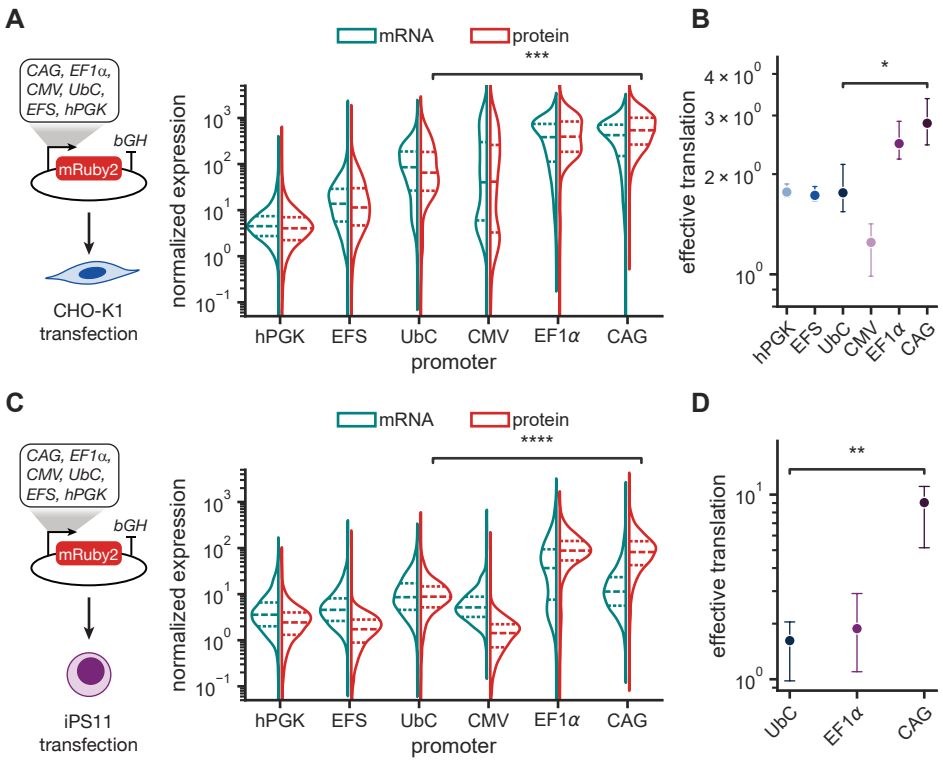

Figure S4. Generalizability of promoter strength and effective translation trends to other mammalian cell types.
A. Normalized expression distributions for mRNA (Alexa Fluor™ 488) and protein (mRuby2) in CHO-K1 cells with six constitutive promoters as measured by flow cytometry.
B. Effective translation rate of constitutive promoters transfected in CHO-K1 cells.
C. Normalized expression distributions for mRNA (Alexa Fluor™ 488) and protein (mRuby2) in iPS11 cells with six constitutive promoters as measured by flow cytometry.
D. Effective translation rate of constitutive promoters transfected in iPS11 cells.
Normalized expression is calculated as the fold change of fluorescence intensity relative to a non-transfected sample. Points represent means of three biological replicates, and error bars represent the 95% confidence interval. All data is in arbitrary units from a flow cytometer. *: $0.05 \geq p > 0.01$, **: $0.01 \geq p > 0.001$, ***: $0.001 \geq p > 0.0001$, ****: $0.00001 \geq p$, one-sided t-test.

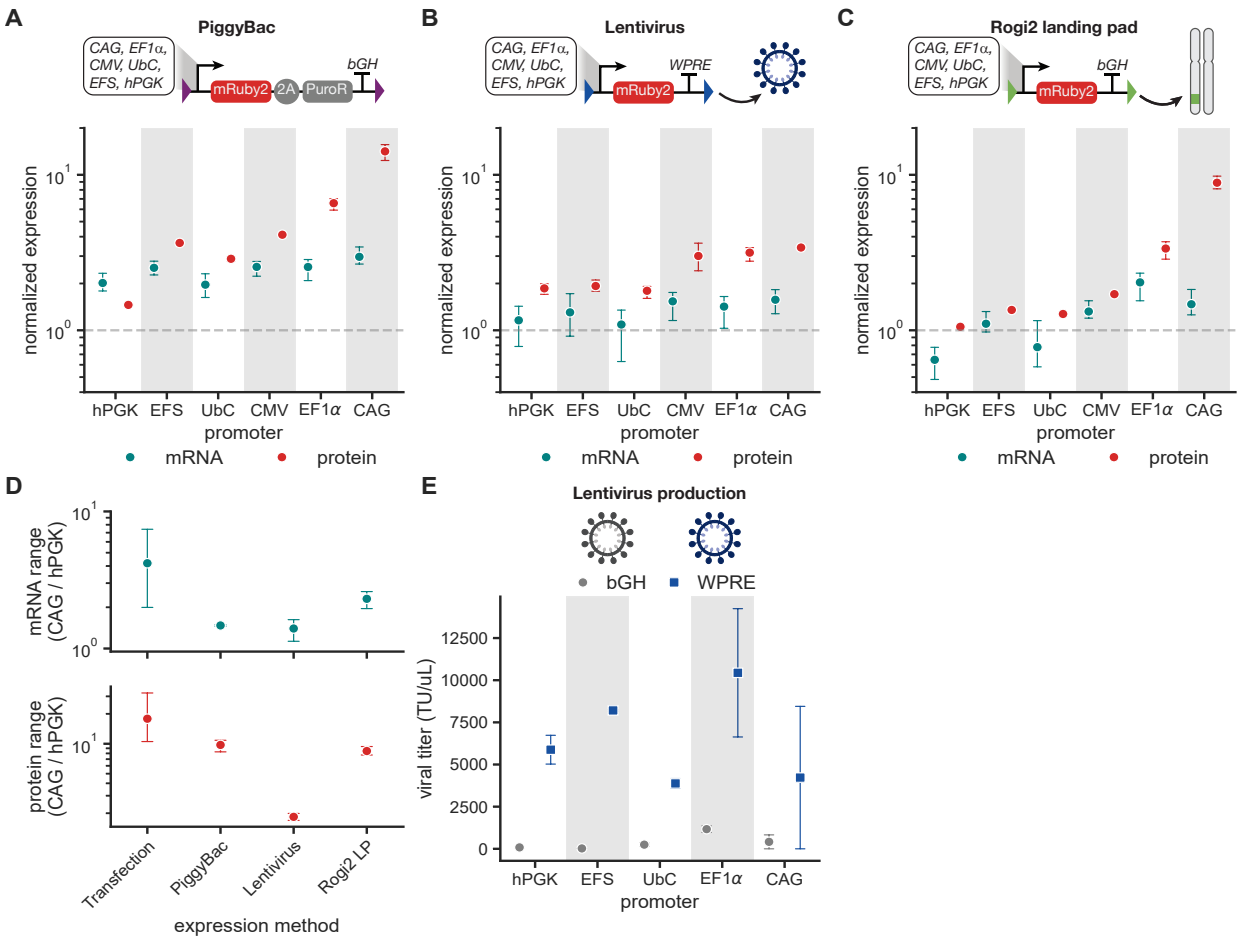

Figure S5. Generalizability of promoter strength across integration contexts.
A. Normalized protein and mRNA expression for HEK293T cells PiggyBac-integrated with mRuby2-2A-PuroR-bGH driven by six different constitutive promoters.
B. Normalized protein and mRNA expression for HEK293T cells Lentivirus-integrated with mRuby2-WPRE driven by six different constitutive promoters.
C. Normalized protein and mRNA expression for *Rogi2* LP HEK293T cells site-specifically integrated with mRuby2-bGH driven by six different constitutive promoters.
D. Range of mRNA and protein expression in different expression contexts as defined by the ratio between expression with the strong promoter CAG and the weak promoter hPGK.
E. Mean viral titer in transducing units (TU) per µL of lentivirus produced with varying promoter and 3' UTR sequence. Points represent the mean ± the 95% confidence interval for two batches of virus.
Normalized expression is calculated as the fold change of fluorescence intensity relative to a non-transfected sample. Dashed lines represent background fluorescence level of a non-transfected sample. Points represent the mean of three biological replicates ± the 95% confidence interval. All data is in arbitrary units from a flow cytometer.

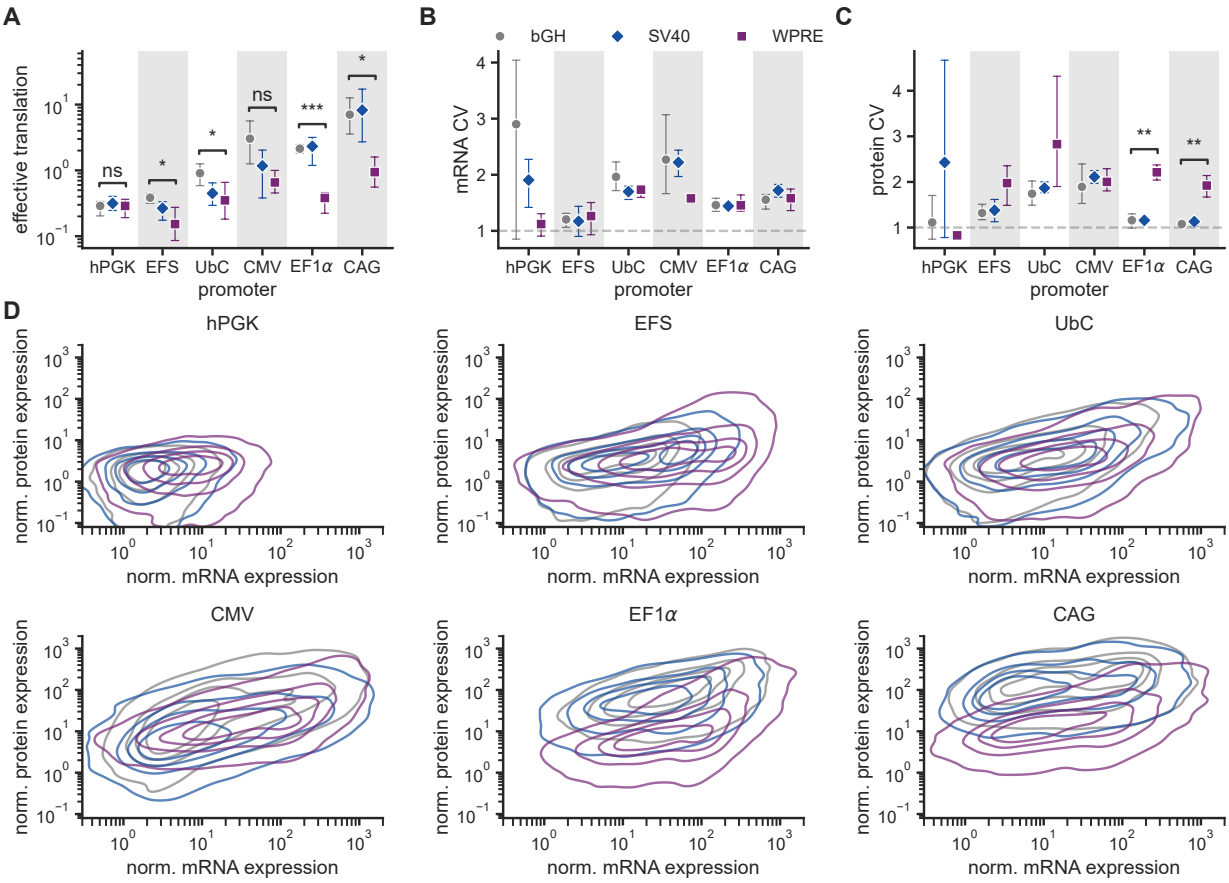

Figure S6. Additional 3' UTR transfection characterization results.

A. Effective translation rate as calculated by the slope of a line fitted the binned data in fig. 3D. *: $0.05 \geq p > 0.01$, ***: $0.001 \geq p > 0.0001$, one-sided t-test.

B. mRuby2 mRNA coefficient of variation in HEK293T transfection with varying promoter and PAS or 3' UTR sequences.

C. mRuby2 protein coefficient of variation in HEK293T transfection with varying promoter and PAS or 3' UTR sequences. **: $0.01 \geq p > 0.001$, two-sided t-test.

D. Joint distributions of mRuby2 mRNA and protein fluorescence in HEK293T transfection with varying promoter and PAS or 3' UTR sequences. Data was randomly downsampled to 10,000 cells per condition for plotting.

Normalized expression is calculated as the fold change of fluorescence intensity relative to a non-transfected sample. Points represent means of three biological replicates, and error bars represent the 95% confidence interval. All data is in arbitrary units from a flow cytometer.

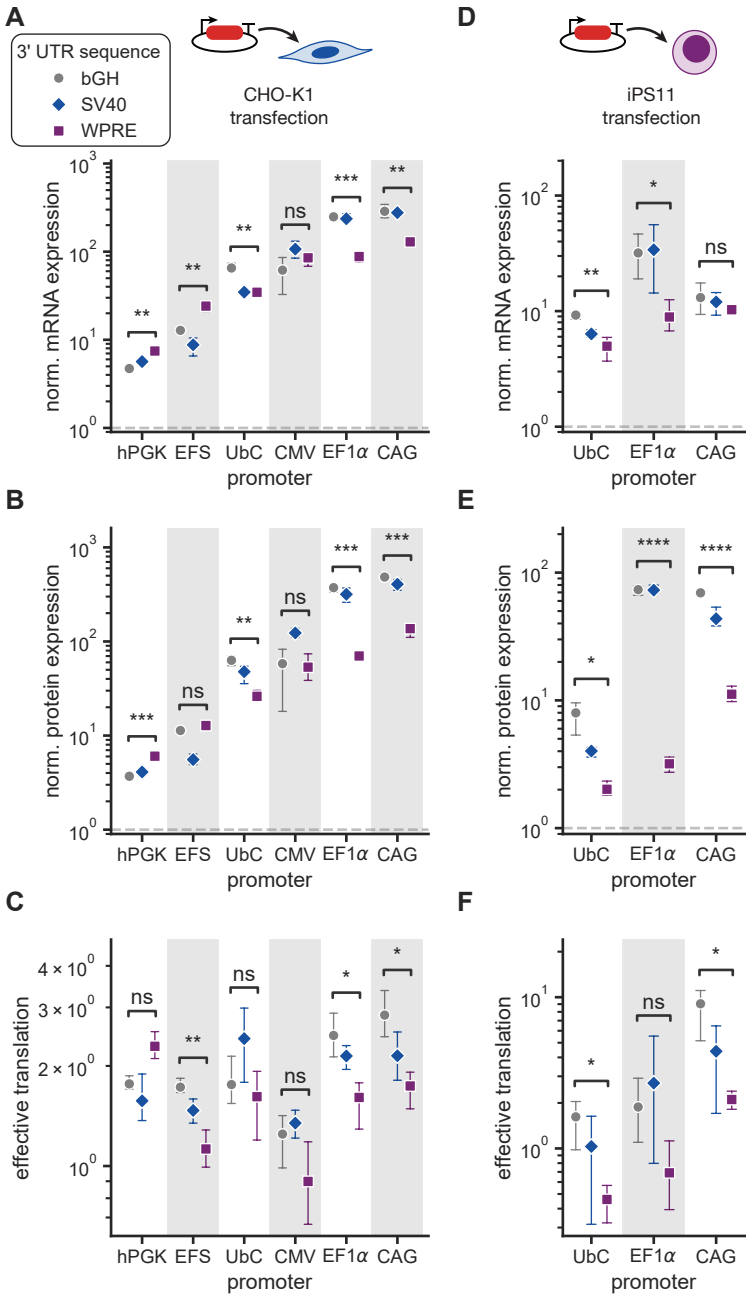

Figure S7. Generalizability of 3' UTR trends to other mammalian cell types.
A, **B**. Normalized geometric mean of mRNA (fig. S7A) and protein (fig. S7B) fluorescence for varying promoter and PAS or 3' UTR sequence in transfection of CHO-K1 cells. **: $0.01 \geq p > 0.001$, ***: $0.001 \geq p > 0.0001$, two-sided t-test.
C. Effective translation rate of constructs with varying 3' UTR sequence in CHO-K1 cell transfection. *: $0.05 \geq p > 0.01$, **: $0.01 \geq p > 0.001$, one-sided t-test.
D, **E**. Normalized geometric mean of mRNA (fig. S7D) and protein (fig. S7E) fluorescence for varying promoter and PAS or 3' UTR sequence in transfection of CHO-K1 cells. *: $0.05 \geq p > 0.01$, **: $0.01 \geq p > 0.001$, ****: $0.00001 \geq p$, two-sided t-test.
F. Effective translation rate of constructs with varying 3' UTR sequence in CHO-K1 cell transfection. *: $0.05 \geq p > 0.01$, one-sided t-test.
Normalized expression is calculated as the fold change of fluorescence intensity relative to a non-transfected sample. Points represent means of three biological replicates, and error bars represent the 95% confidence interval. All data is in arbitrary units from a flow cytometer.

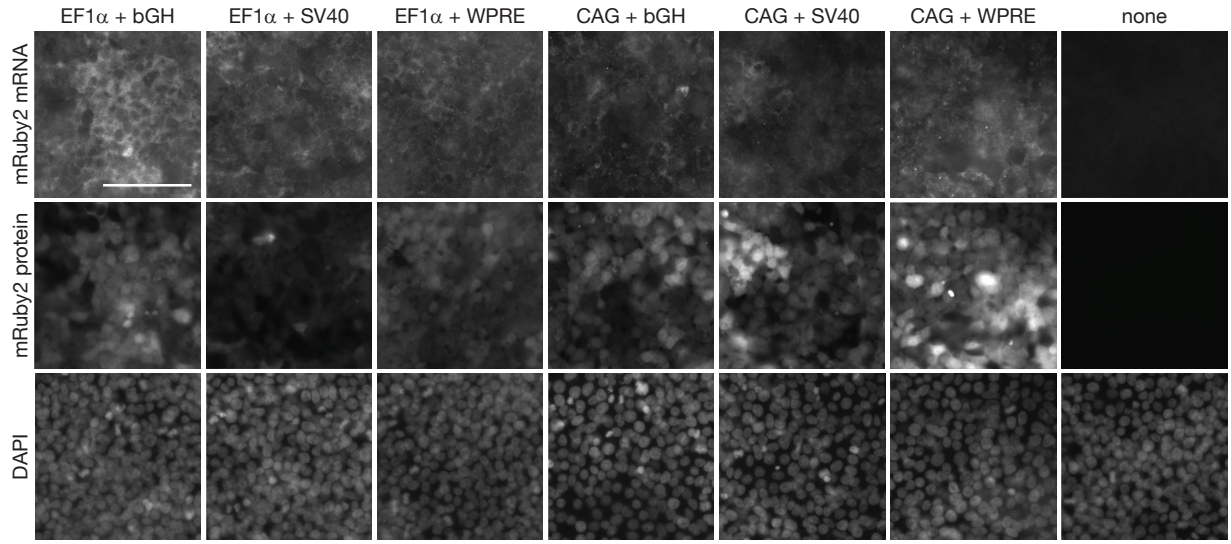

Figure S8. Imaging of *Rogi2* cell lines integrated with strong promoters and varying 5' UTR sequences. *Rogi2* cell lines were integrated with genes expressing mRuby2 from an EF1α or CAG promoter with a bGH, SV40, or WPRE 3' UTR sequence. Cells were labeled with Alexa Fluor™ 647 HCR amplifiers. Scale bar represents 100 μm.

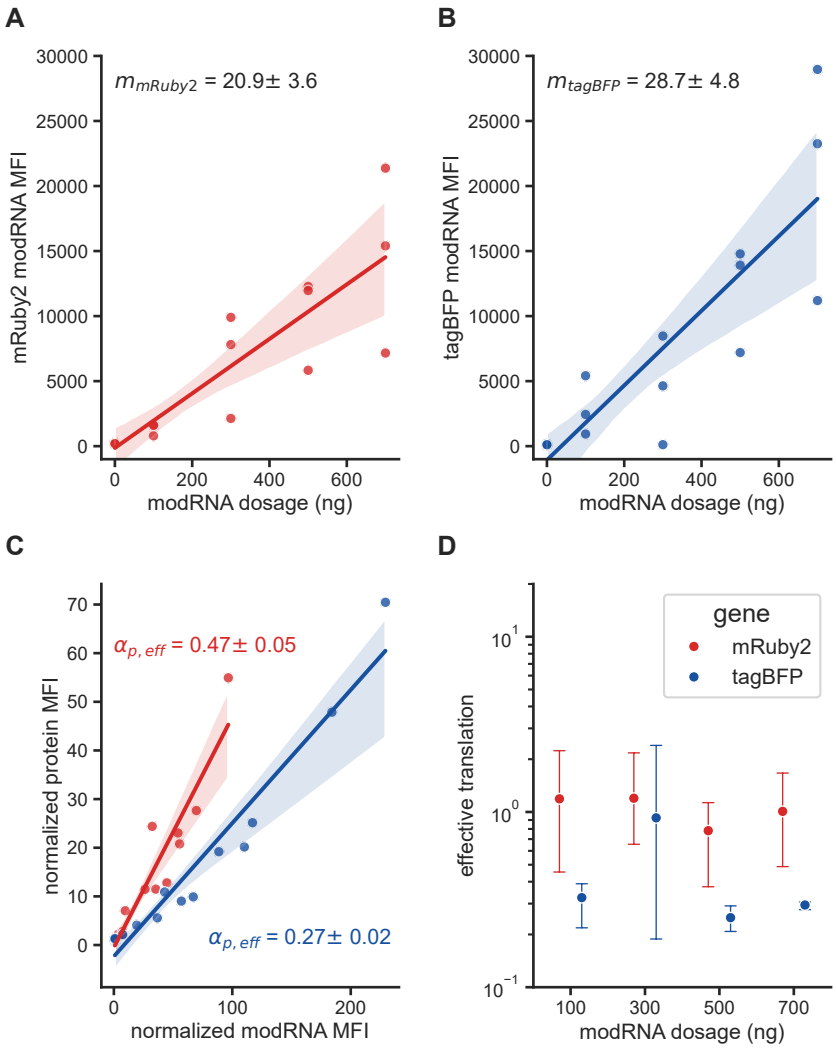

Figure S9. Comparison of modRNA and protein fluorescence between mRuby2 and tagBFP modRNA species.
A. Measurement of HCR Flow-FISH signal for HEK293T cells transfected with varying dosages of mRuby2 modRNA with the CMV promoter 5' UTR and bGH 3' UTR at 12 hours post-transfection.
B. Measurement of HCR Flow-FISH signal for HEK293T cells transfected with varying dosages of tagBFP modRNA with the CMV promoter 5' UTR and bGH 3' UTR at 12 hours post-transfection.
C. Normalized protein vs. normalized modRNA fluorescence for varying dosages of mRuby2 (red) or tagBFP (blue) modRNA with the CMV promoter 5' UTR and bGH 3' UTR at 12 hours post-transfection.
D. Effective translation rate as calculated by the slope of a line fitted the data binned by marker level.
Normalized fluorescence is calculated as the fold change of fluorescence intensity relative to a non-transfected sample. Points represent individual biological replicates, and shaded regions represent the 95% confidence interval of the regression estimate. Slopes of linear regressions are reported ± the standard error. All data is in arbitrary units from a flow cytometer.

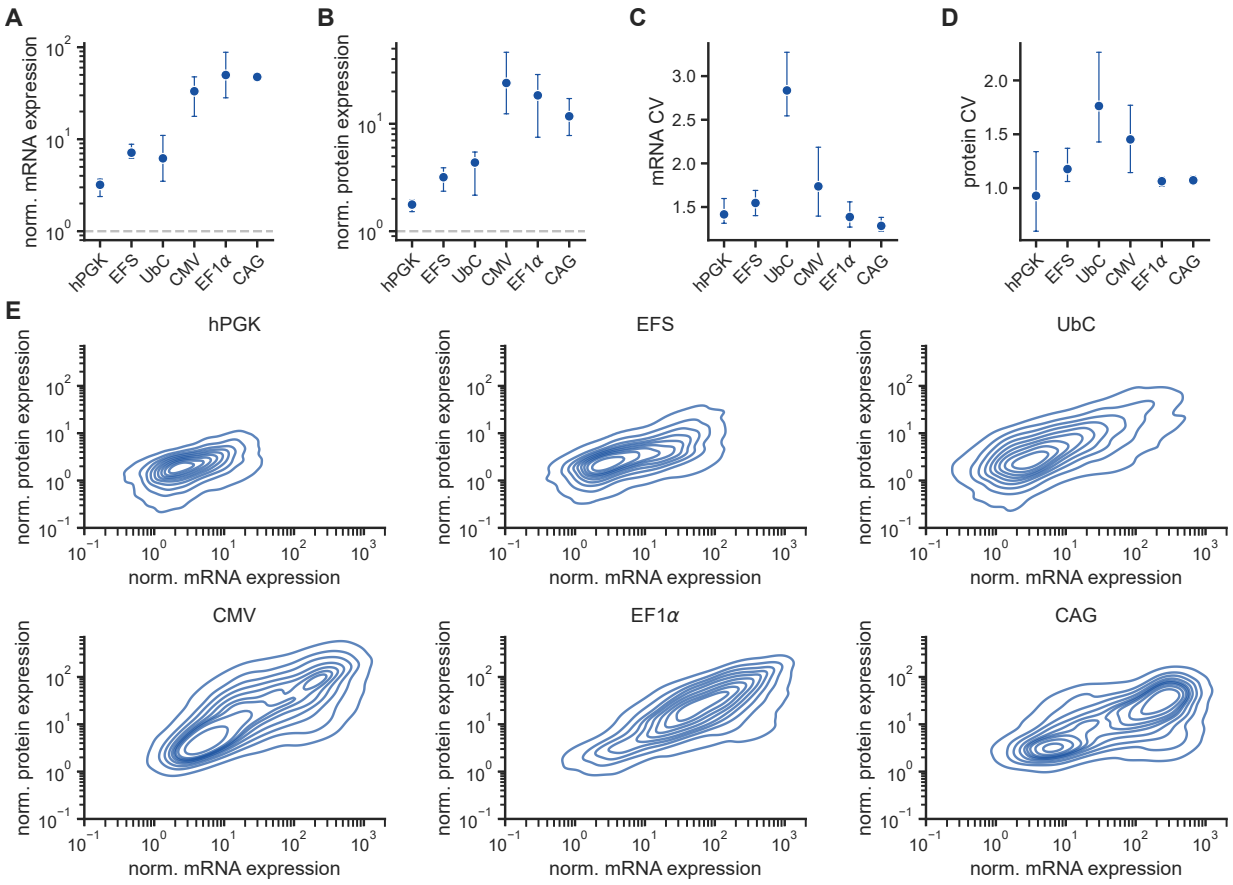

Figure S10. Additional tagBFP transfection characterization results.
A. Normalized tagBFP mRNA expression in HEK293T transfection with varying promoter sequences. Dashed line indicates background fluorescence level.
B. Normalized tagBFP protein expression in HEK293T transfection with varying promoter sequences. Dashed line indicates background fluorescence level.
C. TagBFP mRNA coefficient of variation in HEK293T transfection with varying promoter sequences.
D. TagBFP protein coefficient of variation in HEK293T transfection with varying promoter sequences.
E. Joint distributions of tagBFP mRNA and protein fluorescence in HEK293T transfection with varying promoter sequences. Data was randomly downsampled to 10,000 cells per condition for plotting.
Normalized expression is calculated as the fold change of fluorescence intensity relative to a non-transfected sample. Points represent means of three biological replicates, and error bars represent the 95% confidence interval. All data is in arbitrary units from a flow cytometer.

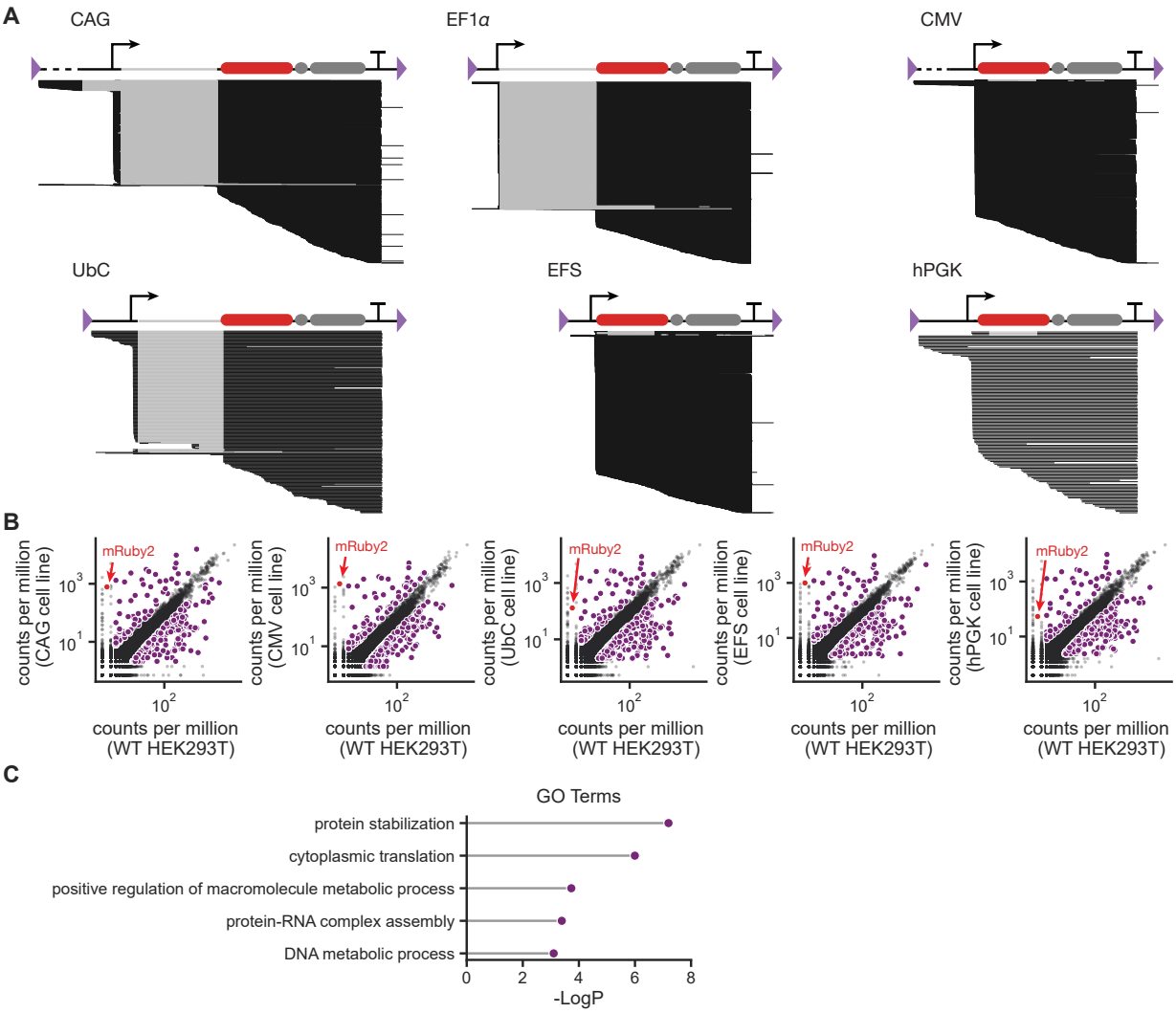

Figure S11. Additional long-read sequencing results.

A. RNA transcript maps for the six constitutive promoters tested without downsampling.

B. Comparison of endogenous gene expression levels between cell lines integrated with the CAG, CMV, UbC, EFS, or hPGK promoters relative to a WT HEK293T cell line. Genes are considered "differentially expressed" between the cell lines if the absolute value of the fold change is greater than 1.5. These genes are indicated in purple. mRuby2 expression is indicated in red.

C. Gene ontology analysis for genes differentially expressed in all PiggyBac-integrated cell lines.

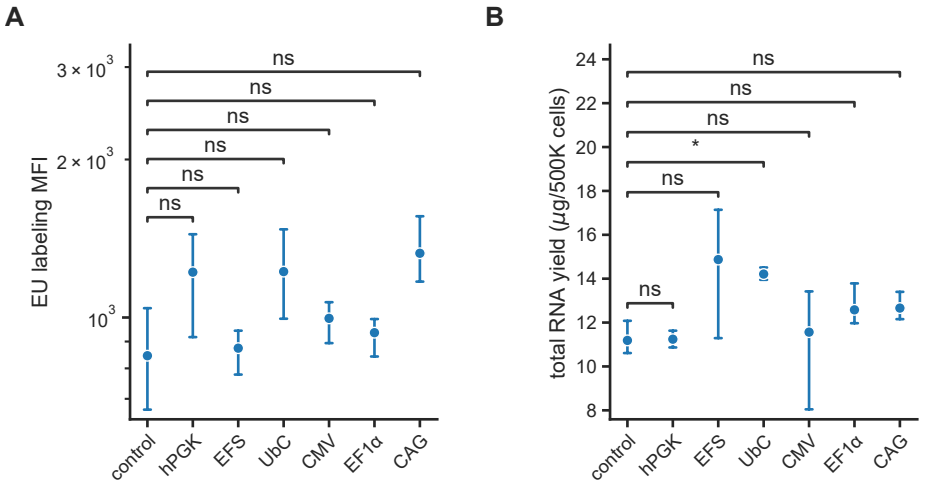

Figure S12. Analysis of total transcription in PiggyBac-integrated cell lines.
A. 5-ethynyluridine (EU) labeling with a Pacific Blue azide to assess levels of recently-transcribed RNA in cells. Data is in arbitrary units from a flow cytometer. Points represent means of three biological replicates, and error bars represent the 95% confidence interval.
B. Total RNA yield per 500,000 cells. Points represent means of three biological replicates, and error bars represent the 95% confidence interval.
*: $0.05 \geq p > 0.01$, two-sided t-test with Bonferroni correction.

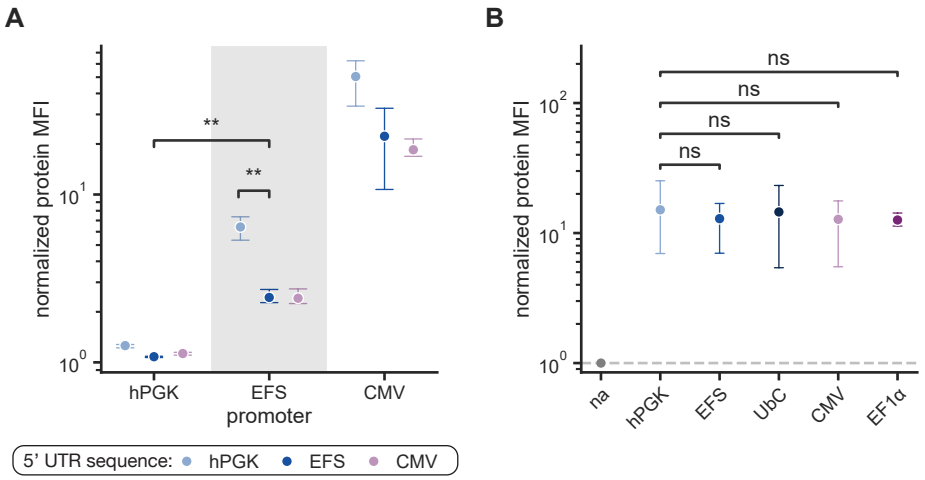

Figure S13. Protein expression data for plasmid transfection (A) and modRNA transfection (B) with varying sequences. Comparable conditions are presented in fig. 6.
A. Normalized geometric mean of protein fluorescence for mRuby2 plasmid transfection with varying 5' UTR and canonical promoter sequence.
B. Normalized geometric mean of protein fluorescence for mRuby2 modRNA transfection with varying 5' UTR sequence.
Normalized expression is calculated as the fold change of fluorescence intensity relative to a non-transfected sample. Points represent means of three biological replicates, and error bars represent the 95% confidence interval. All data is in arbitrary units from a flow cytometer.
**: $0.01 \geq p > 0.001$, two-sided t-test with Bonferroni correction.

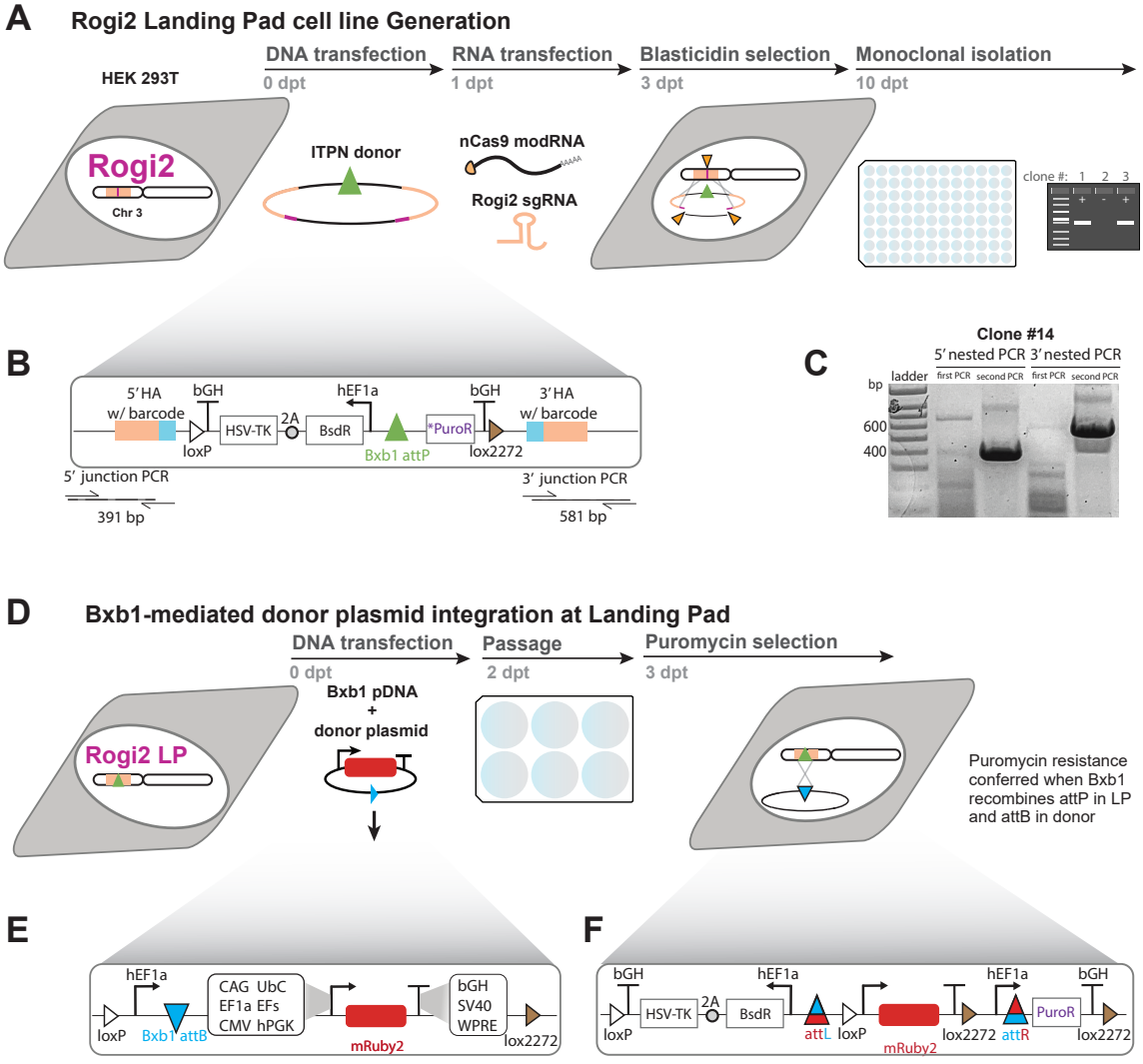

Figure S14. HEK293T *Rogi2* landing pad for evaluating genetic parts in site-specific integration.
A. Workflow for installation of landing pad at *Rogi2* via In Trans Paired Nicking (ITPN).
B. Architecture of landing pad encoded on *Rogi2*-specific ITPN donor. The barcodes directly internal to the homology arms are used in 5' and 3' junction PCRs when genotyping candidate clones.
C. Nested PCR results for 5' and 3' junctions in monoclone #14, verifying installation of donor DNA at *Rogi2*.
D. Workflow for integrating donor plasmids at *Rogi2* landing pad via Bxb1.
E. Depiction of genetic components encoded on Bxb1 donor plasmids to integrate the complete set of constitutive promoter-polyA mRuby2 transcriptional units.
F. Depiction of genomic sequence after Bxb1-mediated insertion of the donor plasmid at the Landing Pad. The insertion of the hEF1a promoter and start codon in-frame with the puromycin-resistance gene pre-installed at the landing pad confers resistance only to successfully-recombined cells.

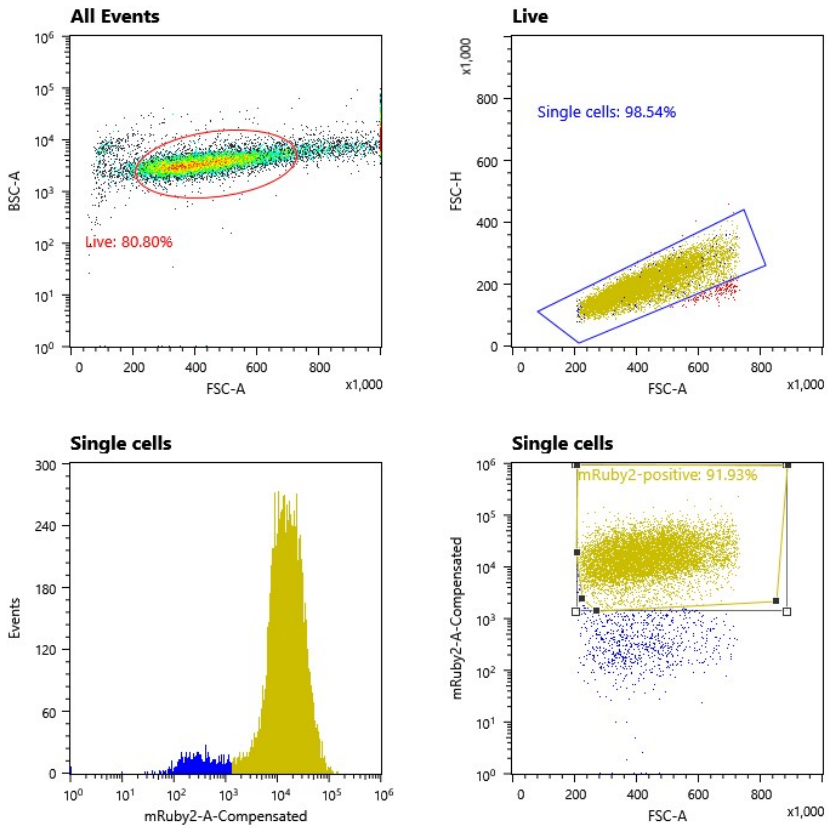

Figure S15. Gating strategy used for sorting PiggyBac-integrated cell lines on a Sony MA-900 flow sorter.

Supplementary Tables

Table S1. Constitutive promoter sequences annotated with the most common TSS (bold and underline) and intron (italic and lowercase) coordinates as determined by long-read sequencing

Promoter	Sequence
CAG	TTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGT TACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGAC GTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATG GGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAG TACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACAT GACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCAT GGTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCTCCCCACCCCC AATTTTGTATTTATTTATTTTTTAATTATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGG GGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGG TGCGGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCG GCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCGGCGGGCGGGAGTCGCTGCGCGCTG**C**CT TCGCCCCGTGCCCCGCTCCGCCGCCGCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGC GTTACTCCCACAG*gtgagcgggcgggacggcccttctcctccgggctgtaattagcgctt* *ggtttaatgacggcttgtttcttttctgtggctgcgtgaaagccttgaggggctccggga* *gggccctttgtgcgggggagcggctcgggggtgcgtgcgtgtgtgtgtgcgtggggag* *cgccgcgtgcggctccgcgctgcccggcggctgtgagcgctgcgggcgcggcgcggggct* *ttgtgcgctccgcagtgtgcgcgaggggagcgcggccgggggcggtgccccgcggtgcgg* *gggggctgcgaggggaacaaaggctgcgtgcggggtgtgtgcgtgggggggtgagcagg* *gggtgtgggcgcgtcggtcgggctgcaacccccccctgcacccccctccccgagttgctga* *gcacggcccggcttcgggtgcggggctccgtacggggcgtggcgcggggctcgccgtgcc* *gggcgggggtggcggcaggtgggggtgccgggcggggcggggccgcctcgggccgggga* *gggctcgggggaggggcgcggcggccccggagcgccggcggctgtcgaggcgcggcgag* *ccgcagccattgccttttatggtaatcgtgcgagagggcgcagggacttcctttgtccca* *aatctgtgcggagccgaaatctgggaggcgccgccgcacccccctctagcgggcgcggggc* *gaagcggtgcggcgccggcaggaaggaaatgggcggggagggccttcgtgcgtcgccgcg* *ccgccgtccccttctccctctccagcctcggggctgtccgcgggggacggctgccttcg* *gggggacggggcagggcggggttcggcttctggcgtgtgaccggcggctctagagcctc* *tgctaaccatgttcatgccttcttcttttcctacag*CTCCTGGGCAACGTGCTGGTTAT TGTGCTGTCTCATCATTTTGGCAAA

Continued on next page

Table S1 – continued from previous page

Promoter	Sequence
EF1α	GCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGG GAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTG ATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAG TAGTCGCCGTGAACGTTCTTTTTCG**C**AACGGGTTTGCCGCCAGAACACAG*gtaagtgccg* *tgtgtggttcccgcgggcctggcctctttacgggttatggcccttgcgtgccttgaatta* *cttccacgcccctggctgcagtacgtgattcttgatcccgagcttcgggttggaagtggg* *tgggagagttcgaggccttgcgcttaaggagccccttcgcctcgtgcttgagttgaggcc* *tggcttgggcgctggggccgccgcgtgcgaatctggtggcaccttcgcgcctgtctcgct* *gctttcgataagtctctagccatttaaaatttttgatgacctgctgcgacgcttttttc* *tggcaagatagtcttgtaaatgcgggccaagatctgcacactggtatttcggttttggg* *gccgcgggcggcgacggggcccgtgcgtcccagcgcacatgttcggcgaggcggggcctg* *cgagcgcggccaccgagaatcggacgggggtagtctcaagctggccggcctgctctggtg* *cctggcctcgcgccgccgtgtatcgccccgccctgggcggcaaggctggcccggtcggca* *ccagttgcgtgagcggaaagatggccgcttcccggccctgctgcagggagctcaaaatgg* *aggacgcggcgctcgggagagcgggcgggtgagtcacccacacaaaggaaaagggccttt* *ccgtcctcagccgtcgcttcatgtgactccacggagtaccgggcgccgtccaggcacctc* *gattagttctcgagcttttggagtacgtcgtctttaggttgggggagggtttttatgcg* *atggagtttccccacactgagtgggtggagactgaagttaggccagcttggcacttgatg* *taattctccttggaatttgccctttttgagtttggatcttggttcattctcaagcctcag* *acagtggttcaaagtttttttcttccatttcag*GTGTCGTGAG
CMV	CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATT GACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCA ATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCC AAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTA CATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC CATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGG ATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG GGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCTGAGCTCGTTTAGTGAACCGTC**A**GATCGCCTGG AGACGCCATCCACGCTGT

Continued on next page

Table S1 – continued from previous page

Promoter	Sequence
UbC	GGCCTCCGCGCCGGGTTTTGGCGCCTCCCGCGGGCGCCCCCCTCCTCACGGCGAGCGCTG CCACGTCAGACGAAGGGCGCAGCGAGCGTCCTGATCCTTCCGCCCGGACGCTCAGGACAG CGGCCCGCTGCTCATAAGACTCGGCCTTAGAACCCCAGTATCAGCAGAAGGACATTTTAG GACGGGACTTGGGTGACTCTAGGGCACTGGTTTTCTTTCCAGAGAGCGGAACAGGCGAGG AAAAGTAGTCCCTTCTCGGCGATTCTGCGGAGGGATCTCCGTGGGGCGGTGAACGCCGAT GATTATATAAGGACGCGCCGGGTGTGGCACAGCTAGTTCCGT**C**GCAGCCGGGATTTGGGT CGCGGTTCTTGTTTGTGGATCGCTGTGATCGTCACTTG*gtgagtagcgggctgctgggct* *ggccggggctttcgtggccgccgggccgctcggtgggacggaagcgtgtggagagatcgc* *caagggctgtagtctgggtccgcgagcaaggttgccctgaactgggggttgggggagcg* *cagcaaaatggcggctgttcccgagtcttgaatggaagacgcttgtgaggcgggctgtga* *ggtcgttgaaacaaggtgggggcatggtgggcggcaagaacccaaggtcttgaggcctt* *cgctaatgcgggaaagctcttattcgggtgagatgggctggggcaccatctggggaccct* *gacgtgaagtttgtcactgactggagaactcggtttgtcgtctgttgcgggggcggcagt* *tatggcggtgccgttgggcagtgcacccgtacctttgggagcgcgcgccctcgtcgtgtc* *gtgacgtcacccgttctgttggcttataatgcagggtggggccacctgtcggtaggtgtg* *cggtaggcttttctccgtcgcaggacgcagggttcgggcctagggtaggctctcctgaat* *cgacaggcgccggacctctggtgaggggagggataagtgaggcgtcagtttctttggtcg* *gttttatgtacctatcttcttaagtagctgaagctccggttttgaactatgcgctcgggg* *ttggcgagtgtgttttgtgaagtttttaggcacctttgaaatgtaatcatttgggtca* *atatgtaattttcagtgttagactagtaaattgtccgctaaattctggccgttttttggct* *ttttttgttag*AC
EFS	GGGCAGAGCGCACATCGCCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAA CCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCC GCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTC TTTTTCG**C**AACGGGTTTGCCGCCAGAACACAGT
hPGK	GGGGTTGGGGTTGCGCCTTTTCCAAGGCAGCCCTGGGTTTGCGCAGGGACGCGGCTGCTC TGGGCGTGGTTCCGGGAAACGCAGCGGCGCCGACCCTGGATCTCGCACATTCTTCACGTC CGTTCGCAGCGTCACCCGGATCTTCGCCGCTACCCTTGTGGGCCCCCCGGCGACGCTTCC TGCTCCGCCCCTAAGTCGGGAAGGTTCCTTGCGGTTCGCGGCGTGCCGGACGTGACAAAC GGAAGCCGCACGTCTCACTAGTACCCTCGCAGACGGACAGCGCCAGGGAGCAATGGCAGC GCGCCGACCGCGATGGGCTGTGGCCAATAGCGGCTGCTCAGCAGGGCGCGCCGAGAGCAG CGGCCGGGAAGGGGCGGTGCGGGAGGCGGGGTGTGGGGCGGTAGTGTGGGCCCTGTTCCT GCCCGCGCGGTGTTCCGCATTCTGCAAGC**C**TCCGGAGCGCACGTCGGCAGTCGGCTCCCT CGTTGACCGAATCACCGACCTCTCTCCCCAGA

Table S2. Differentially expressed genes common across all PiggyBac-integrated cell lines with the indicated promoters.

HGNC ID	Description	Category
AC019205.1	KHDC1 Antisense RNA 1	lncRNA
AC113935.1	Ribosomal Protein L17 (RPL17) Pseudogene 6	Pseudogene
ACTA2	Actin Alpha 2, Smooth Muscle	Protein Coding
ACTR3	Actin Related Protein 3	Protein Coding

Continued on next page

Table S2 – continued from previous page

HGNC ID	Description	Category
ANK2	Ankyrin 2	Protein Coding
ANP32B	Acidic Nuclear Phosphoprotein 32 Family Member B	Protein Coding
ATF3	Activating Transcription Factor 3	Protein Coding
ATP5F1C	ATP Synthase F1 Subunit Gamma	Protein Coding
BNIP3P1	BCL2 Interacting Protein 3 Pseudogene 1	Pseudogene
C11orf96	Chromosome 11 Open Reading Frame 96	Protein Coding
C6orf48	Small Nucleolar RNA Host Gene 32	lncRNA
CALM2	Calmodulin 2	Protein Coding
CCNB1IP1	Cyclin B1 Interacting Protein 1	Protein Coding
CCT5	Chaperonin Containing TCP1 Subunit 5	Protein Coding
CCT8	Chaperonin Containing TCP1 Subunit 8	Protein Coding
CHCHD3	Coiled-Coil-Helix-Coiled-Coil-Helix Domain Containing 3	Protein Coding
CLIP4	CAP-Gly Domain Containing Linker Protein Family Member 4	Protein Coding
CTH	Cystathionine Gamma-Lyase	Protein Coding
DHRS2	Dehydrogenase/Reductase 2	Protein Coding
EDARADD	EDAR Associated Via Death Domain	Protein Coding
EEF1A1	Eukaryotic Translation Elongation Factor 1 Alpha 1	Protein Coding
EEF1A1P5	Eukaryotic Translation Elongation Factor 1 Alpha 1 Pseudogene 5	Pseudogene
EIF3E	Eukaryotic Translation Initiation Factor 3 Subunit E	Protein Coding
EIF3L	Eukaryotic Translation Initiation Factor 3 Subunit L	Protein Coding
EIF4A2	Eukaryotic Translation Initiation Factor 4A2	Protein Coding
ENO1	Enolase 1	Protein Coding
FDPS	Farnesyl Diphosphate Synthase	Protein Coding
FHIT	Fragile Histidine Triad Diadenosine Triphosphatase	Protein Coding
FTH1P20	Ferritin Heavy Chain 1 Pseudogene 20	Pseudogene
GAPDHP1	Glyceraldehyde-3-Phosphate Dehydrogenase Pseudogene 1	Pseudogene
H1F0	H1.0 Linker Histone	Protein Coding
HMGB1P5	High Mobility Group Box 1 Pseudogene 5	Pseudogene
HMGN2	High Mobility Group Nucleosomal Binding Domain 2	Protein Coding
HMGN2P3	High Mobility Group Nucleosomal Binding Domain 2 Pseudogene 3	Pseudogene
HNRNPA1P10	Heterogeneous Nuclear Ribonucleoprotein A1 Pseudogene 10	Pseudogene
HNRNPA1P7	Heterogeneous Nuclear Ribonucleoprotein A1 Pseudogene 7	Pseudogene
HNRNPK	Heterogeneous Nuclear Ribonucleoprotein K	Protein Coding
HSP90AB1	Heat Shock Protein 90 Alpha Family Class B Member 1	Protein Coding
HSPD1	Heat Shock Protein Family D (Hsp60) Member 1	Protein Coding
ICA1	Islet Cell Autoantigen 1	Protein Coding
JUN	Jun Proto-Oncogene, AP-1 Transcription Factor Subunit	Protein Coding
KRTAP19-1	Keratin Associated Protein 19-1	Protein Coding
LDHB	Lactate Dehydrogenase B	Protein Coding
MMADHC	Metabolism Of Cobalamin Associated D	Protein Coding
MORF4L1P1	Mortality Factor 4 Like 1 Pseudogene 1	Pseudogene
NOP53	NOP53 Ribosome Biogenesis Factor	Protein Coding
NPM1	Nucleophosmin 1	Protein Coding
NPM1P27	Nucleophosmin 1 Pseudogene 27	Pseudogene

Continued on next page

Table S2 – continued from previous page

HGNC ID	Description	Category
NUPR1	Nuclear Protein 1, Transcriptional Regulator	Protein Coding
PABPC1	Poly(A) Binding Protein Cytoplasmic 1	Protein Coding
PCNA	Proliferating Cell Nuclear Antigen	Protein Coding
PGK1	Phosphoglycerate Kinase 1	Protein Coding
PHB	Prohibitin 1	Protein Coding
PLCB4	Phospholipase C Beta 4	Protein Coding
POLA1	DNA Polymerase Alpha 1, Catalytic Subunit	Protein Coding
PON2	Paraoxonase 2	Protein Coding
PPA1	Inorganic Pyrophosphatase 1	Protein Coding
PPP1R15A	Protein Phosphatase 1 Regulatory Subunit 15A	Protein Coding
PTGES3	Prostaglandin E Synthase 3	Protein Coding
PTMA	Prothymosin Alpha	Protein Coding
PTMAP2	Prothymosin Alpha Pseudogene 2	Pseudogene
RAN	Member Ras Oncogene Family	Protein Coding
RHOU	Ras Homolog Family Member U	Protein Coding
RPL13AP25	Ribosomal Protein L13a Pseudogene 25	Pseudogene
RPL15P3	Ribosomal Protein L15 Pseudogene 3	Pseudogene
RPL18AP3	Ribosomal Protein L18a Pseudogene 3	Pseudogene
RPL21P28	Ribosomal Protein L21 Pseudogene 28	Pseudogene
RPL23AP42	Ribosomal Protein L23a Pseudogene 42	Pseudogene
RPL26P30	Ribosomal Protein L26 Pseudogene 30	Pseudogene
RPL3	Ribosomal Protein L3	Protein Coding
RPL3P4	Ribosomal Protein L3 Pseudogene 4	Pseudogene
RPL4	Ribosomal Protein L4	Protein Coding
RPL4P4	Ribosomal Protein L4 Pseudogene 4	Pseudogene
RPL5	Ribosomal Protein L5	Protein Coding
RPS26P47	Ribosomal Protein S26 Pseudogene 47	Pseudogene
RPS27AP16	RPS27A Pseudogene 16	Pseudogene
RPS7P1	Ribosomal Protein S7 Pseudogene 1	Pseudogene
RPSA	Ribosomal Protein SA	Protein Coding
SCARNA1	Small Cajal Body-Specific RNA 1	scaRNA
SH3BGR	SH3 Domain Binding Glutamate Rich Protein	Protein Coding
SNORA19	Small Nucleolar RNA, H/ACA Box 19	snoRNA
TESK1	Testis Associated Actin Remodeling Kinase 1	Protein Coding
TPI1	Triosephosphate Isomerase 1	Protein Coding
TRIB3	Tribbles Pseudokinase 3	Protein Coding
UBE2S	Ubiquitin Conjugating Enzyme E2 S	Protein Coding
UBE2SP1	Ubiquitin Conjugating Enzyme E2 S Pseudogene 1	Pseudogene
YBX1	Y-Box Binding Protein 1	Protein Coding
YBX1P1	Y-Box Binding Protein 1 Pseudogene 1	Pseudogene
ZNF561-AS1	ZNF561 Antisense RNA 1	lncRNA

BsaI Golden Gate cloning			
Plasmid	Sequence type	Upstream overhang	Downstream overhang
pPV0	pShip backbone	CGCT	TACT
pPV1	Promoter sequence	TACT	AATG
pPV2	Coding sequence	AATG	CAAC
pPV3	Polyadenylation sequence or 3' UTR	CAAC	CGCT
PaqCI Golden Gate cloning			
Plasmid	Sequence type	Upstream overhang	Downstream overhang
pShip	Single transcriptional unit	TACT	CGCT
pHarbor	Backbone for genomic integration	CGCT	TACT

Table S3. Type IIS restriction enzyme overhangs used for golden gate cloning

Plasmid	Sequence of interest	Sequence source	Notes
pKG1117	pPV0-pShip backbone	Addgene #29652	
pKG2039	pPV1-CAG	Addgene #140534	Sequence has an internal PaqCI site
pKG2019	pPV1-EF1a	Addgene #138730	
pKG0618	pPV1-CMV	Addgene #40651	
pKG1988	pPV1-UbC	Addgene #25734	Sequence was domesticated by mutating internal BsaI sites
pKG1179	pPV1-EFS	Addgene #138730	EF1a sequence excluding intron
pKG1367	pPV1-hPGK	Addgene #41393	Sequence was domesticated by mutating internal BsaI sites
pKG2402	pPV1-TRE3G (Tet-On)	Addgene #63800	
pKG0743	pPV1-ZF43x6-C (COMET)	Addgene #138732	
pKG3657	pPV1-ZF10.BSx8 (synZiFTR)	Gift from the Khalil Lab	
pKG0586	pPV2-mRuby2	Addgene #90236	
pKG2387	pPV2-mRuby2-P2A-PuroR	Addgene #90236 and #1764	
pKG1055	pPV2-tagBFP	Addgene #70224	Sequence includes silent mutations to remove BsaI sites
pKG3659	pPV2-rtTA-P2A-iRFP720 (Tet-On)	Addgene #105840 and gift from the Khalil Lab	
pKG3658	pPV2-NLS-FKBP-ZF43-P2A-NES-VP64-FRB-P2A-iRFP720 (COMET)	Addgene #138844 and #138852	
pKG2197	pPV2-ZF10-NS3-p65 (synZiFTR)	Addgene #195468	
pKG0587	pPV3-bGH	Addgene #105841	
pKG0588	pPV3-SV40	Addgene #25734	
pKG1313	pPV3-WPRE	Addgene #25734	
pKG0893	pHarbor-LentiX1 backbone	Addgene #17297	
pKG1334	pHarbor-PiggyBac backbone	Addgene #63800	
pKG3560	pHarbor-Rogi2 backbone	Addgene #198040	

Table S4. List of pPV and pHarbor plasmids used in Golden Gate cloning

Plasmid ID	Name	Description
pKG1970	pGEEC501	pShip-EFS-mRuby2-bGH
pKG1971	pGEEC502	pShip-EFS-mRuby2-SV40
pKG1972	pGEEC503	pShip-EFS-mRuby2-WPRE
pKG1973	pGEEC504	pShip-hPGK-mRuby2-bGH
pKG1974	pGEEC505	pShip-hPGK-mRuby2-SV40
pKG1975	pGEEC506	pShip-hPGK-mRuby2-WPRE
pKG2021	pGEEC507	pShip-UbC-mRuby2-bGH
pKG2022	pGEEC508	pShip-UbC-mRuby2-SV40
pKG2023	pGEEC509	pShip-UbC-mRuby2-WPRE
pKG2116	pGEEC510	pShip-EF1α-mRuby2-bGH
pKG2117	pGEEC511	pShip-EF1α-mRuby2-SV40
pKG2118	pGEEC512	pShip-EF1α-mRuby2-WPRE
pKG2075	pGEEC516	pShip-CAG-mRuby2-bGH
pKG2076	pGEEC517	pShip-CAG-mRuby2-SV40
pKG2077	pGEEC518	pShip-CAG-mRuby2-WPRE
pKG2309	pGEEC524	pShip-EFS-ZF10-NS3-p65-bGH
pKG2215	pGEEC525	pShip-ZF43x6-C-mRuby2-bGH
pKG2410	pGEEC526	pShip-TRE3G-mRuby2-bGH
pKG2497	pGEEC529	pShip-CMV-mRuby2-bGH
pKG2498	pGEEC530	pShip-CMV-mRuby2-SV40
pKG2499	pGEEC531	pShip-CMV-mRuby2-WPRE
pKG2436	pGEEC532	pShip-EF1α-mRuby2-P2A-PuroR-bGH
pKG2437	pGEEC533	pShip-EFS-mRuby2-P2A-PuroR-bGH
pKG2438	pGEEC534	pShip-CMV-mRuby2-P2A-PuroR-bGH
pKG2439	pGEEC535	pShip-hPGK-mRuby2-P2A-PuroR-bGH
pKG2440	pGEEC536	pShip-UbC-mRuby2-P2A-PuroR-bGH
pKG2441	pGEEC537	pShip-CAG-mRuby2-P2A-PuroR-bGH
pKG2906	pGEEC546	pShip-EFS-tagBFP-bGH
pKG2907	pGEEC547	pShip-hPGK-tagBFP-bGH
pKG2897	pGEEC548	pShip-UbC-tagBFP-bGH
pKG2898	pGEEC549	pShip-EF1α-tagBFP-bGH
pKG2899	pGEEC550	pShip-CMV-tagBFP-bGH
pKG2900	pGEEC551	pShip-CAG-tagBFP-bGH
pKG3660	pGEEC570	pShip-ZF10.BSx8-ybTATA-spacer-mRuby2-bGH
pKG3666	pGEEC576	pShip-EFS-ZF43a(inducible)-P2A-iRFP720-bGH
pKG3668	pGEEC578	pShip-EFS-rtTA-P2A-iRFP720-bGH
pKG3831	pGEEC579	pShip-UbC(no intron)-mRuby2-bGH
pKG3774	pGEEC580	pShip-EF1α(no intron)-mRuby2-bGH
pKG3775	pGEEC582	pShip-hPGK(promoter)-CMV(UTR)-mRuby2-bGH
pKG3776	pGEEC583	pShip-hPGK(promoter)-EFS(UTR)-mRuby2-bGH
pKG3777	pGEEC584	pShip-EFS(promoter)-CMV(UTR)-mRuby2-bGH
pKG3778	pGEEC585	pShip-EFS(promoter)-hPGK(UTR)-mRuby2-bGH
pKG3779	pGEEC586	pShip-CMV(promoter)-EFS(UTR)-mRuby2-bGH
pKG3780	pGEEC587	pShip-CMV(promoter)-hPGK(UTR)-mRuby2-bGH

Table S5. List of pShip plasmids used in transfection experiments and cloning.

modRNA species	Plasmid template	Forward primer	Reverse primer
β-globin-mRuby2	pKG2707	5'-AGCTATAATACGACT CACTATAAGctcctgggc aacgtgctg-3'	5'-poly(T)116-GCAA TGAAAATAAATGTTTTTT ATTAGGCAGAAT-3'
β-globin-tagBFP	pKG3832	5'-AGCTATAATACGACT CACTATAAGctcctgggc aacgtgctg-3'	5'-poly(T)116-GCAA TGAAAATAAATGTTTTTT ATTAGGCAGAAT-3'
hPGK-mRuby2	pKG1973	5'-TAATACGACTCACTA TAAGgctccggagcgca c-3'	5'-poly(T)124-TGCA ATTTCCTCATTTTATTAG GAAA-3'
EFS-mRuby2	pKG1970	5'-TAATACGACTCACTA TAAGgcaacgggtttgcc g-3'	5'-poly(T)124-TGCA ATTTCCTCATTTTATTAG GAAA-3'
UbC-mRuby2	pKG3831	5'-TAATACGACTCACTA TAAGgcgcagccgggatt tg-3'	5'-poly(T)124-TGCA ATTTCCTCATTTTATTAG GAAA-3'
CMV-mRuby2	pKG2497	5'-TAATACGACTCACTA TAAGgagatcgcctggag acg-3'	5'-poly(T)124-TGCA ATTTCCTCATTTTATTAG GAAA-3'
EF1α-mRuby2	pKG3774	5'-TAATACGACTCACTA TAAGgcaacgggtttgcc g-3'	5'-poly(T)124-TGCA ATTTCCTCATTTTATTAG GAAA-3'
CMV-tagBFP	pKG2899	5'-TAATACGACTCACTA TAAGgagatcgcctggag acg-3'	5'-poly(T)124-TGCA ATTTCCTCATTTTATTAG GAAA-3'

Table S6. Plasmid templates and primer sequences used in modRNA synthesis. In the forward primers, the T7 promoter sequence is indicated in capitalize letters, and the region complementary to the 5' UTR is indicated in lowercase letters.

Hybridization buffer (make fresh for every use)		
Component	Source	Volume
30% formamide	Fisher Scientific, BP227	3 mL
20X SSC	Fisher Scientific, S2713 and BP3271	2.5 mL
1M citric acid, pH 6	Fisher Scientific, BP327 and A142-212	90 µL
10% Tween-20	Sigma-Aldrich, P2287	100 µL
50X Denhardt's solution	Fisher Scientific, AAJ63135AD	200 µL
50% dextran sulfate	Fisher Scientific, BP1585	1.5 mL
20 mg/mL BSA	Sigma-Aldrich, A2058	50 µL
nuclease-free water		to 10 mL final volume
Wash buffer (store at -20°C for up to one week)		
Component	Source	Volume
30% formamide	Fisher Scientific, BP227	3 mL
20X SSC	Fisher Scientific, S2713 and BP3271	2.5 mL
1M citric acid, pH 6	Fisher Scientific, BP327 and A142-212	90 µL
10% Tween-20	Sigma-Aldrich, P2287	100 µL
nuclease-free water		to 10 mL final volume
5X SSCT (store at -4°C for up to one month)		
Component	Source	Volume
20X SSC	Fisher Scientific, S2713 and BP3271	2.5 mL
10% Tween-20	Sigma-Aldrich, P2287	100 µL
DEPC-treated water	Genesee Scientific, 20-138	to 10 mL final volume
Amplification buffer (pH 6.8, store at room temperature)		
Component	Source	Amount
sodium chloride	Fisher Scientific, S2713	292 mg
sodium phosphate dibasic	Mallinckrodt, 7917	71 mg
nuclease-free water		to 10 mL final volume

Table S7. Buffer recipes used in RNA-FISH. Volumes are given to make the buffers at 10 mL scale.

Channel	Laser (nm)	Filter	PMT Voltage
all experiments unless otherwise noted			
tagBFP	405	440 / 50	220
Alexa Fluor™ 514	488	530 / 30	260
mRuby2	561	620 / 15	260
iRFP720	637	720 / 30	340
FSC			100
SSC			360
figs. S4 and S7 CHO-K1 transfection			
tagBFP	405	440 / 50	220
Alexa Fluor™ 488	488	530 / 30	240
mRuby2	561	620 / 15	260
FSC			100
SSC			360
figs. S4 and S7 iPS11 transfection			
tagBFP	405	440 / 50	260
Alexa Fluor™ 488	488	530 / 30	300
mRuby2	561	620 / 15	320
FSC			100
SSC			360
fig. S12 EU labeling			
Pacific Blue	405	440 / 50	260
mRuby2	561	620 / 15	340
FSC			100
SSC			360

Table S8. Attune NxT flow cytometer channels and voltages

	Alexa Fluor™ 514-A	mRuby2-A	tagBFP-A	iRFP720-A
Alexa Fluor™ 514-A	1	0	0	0
mRuby2-A	0.01	1	0.015	0
tagBFP-A	0	0	1	0
iRFP720-A	0	0	0	1

Table S9. HCR Flow-FISH compensation matrix for experiments using Alexa Fluor™ 514

Gene	Forward primer	Reverse primer
mRuby2	CCTTGAGGATGGCTGTCTCG	ATGGCCACCACCATCAACTT
GAPDH	GTATCGTGGAAGGACTCATGAC	ACCACCTTCTTGATGTCATCAT

Table S10. Primer sequences used in RT-qPCR experiments