Supporting Information

Multiplexed Gene Expression Tuning with Orthogonal Synthetic Gene Circuits

by Mariola Szenk, Terrence Yim, and Gábor Balázsi

Supporting Methods

Vectors used in this study

phlFs-mLin: The DNA segment GS-CMV-1xphLFO-PUC57mini (GS-1x) was custom-synthesized by GenScript to contain one phlF operator site flanking the initiator sequence (Figure S1) and to contain the restriction sites SnaBI and HindIII. The TetR-mLin sequence was digested from pDN-D2ir-TNG4kwh and cloned into pcDNA5/FRT (Thermo Fisher Scientific, V652020) using Spe and Sph restriction enzyme sites. This produced D2irTN2aG5kwh. GS-1x and D2irTN2aG5kwh were digested with SnaBI and HindIII. The CMV promoter fragment bearing the TATA box, operator site and the Initiator element was inserted into D2irTN2aG5kwh to form an intermediate plasmid (IM1). The vector 12AAQHGC-PhlF-Repressor (pPhlF) was a gift from the Voigt lab. The PhlF:NLS sequence was amplified from pPhlF using primers P1 and P2 (Table S1) and restriction cloned into the IM1 using SbfI and NotI restriction sites, replacing hTetR:2A:EGFP (IM2). The T2A:mCherry sequence was amplified from pDC-BACH1-mCherry-PUC19 (gift of Kevin Farquhar) with primers P3 and P4 (Table S1) and restriction-cloned into IM2 with NheI and NotI. This generated phlFs-mLin.

phlFd-mLin: As above, the DNA segment GS-CMV-2xphLFO-PUC57mini (GS-2x) was custom-synthesized by GenScript to contain two phlF operator sites flanking the Initiator sequence as shown in Figure S1 and to contain the restriction sites SnaBI and HindIII. GS-2x and D2irTN2aG5kwh were digested with SnaBI and HindIII. The CMV promoter fragment bearing the TATA box, operator site and the Initiator element was inserted into D2irTN2aG5kwh to form an intermediate plasmid (IM3). The PhlF:NLS sequence was amplified from pPhlF using primers P1 and P2 (Table S1) and restriction cloned into the IM3 using SbfI and NotI restriction sites, replacing hTetR:2A:EGFP (IM4). The T2A:mCherry sequence was amplified from pDC-BACH1-mCherry-PUC19 with primers P3 and P4 (Table S1) and restriction-cloned into IM4 with NheI and NotI. This generated phlFd-mLin.

Data Analysis

Flow cytometry .FCS files were gated and analyzed using custom MATLAB scripts based on fca_readfcs (Mathworks File Exchange). Cells were adaptively gated with a density-threshold fit of log-transformed forward and side scatter (SSC and FSC) values per sample, to exclude debris and cell clumps. In this method, raw FSC and SSC values were log-transformed, and then plotted as a 2-dimensional histogram using the *hist3* command. The 2-dimensional histogram counts were contoured using the MATLAB *contour* command. We chose the widest contour level on average, which outlined SSC and FSC regions with increasing number of cellular events. We extracted events that

fell within our density-based FSC and SSC contoured gate by using the *inpolygon* MATLAB function. To correct the data for auto-fluorescence, we measured and subtracted the mean fluorescence of the Flp-In-293 parental cell line from the appropriate channel from each sample flow event.

To determine the range of linearity for each gene circuit, L1-norm analysis, linear regression fits and R^2 were computed for a moving dose window starting from uninduced (0 µM DAPG or 0 ng/mL Dox) to maximal induction dose used experimentally per circuit. To compute R^2, we performed parametric linear regression across this moving dose window. For each fit, we calculated $R^2 = 1 – (SS_{res}/SS_{tot})$ from the sum of squares of the residuals and the total sum of squares from the data and the fit. Alternatively, we computed L1-norm, a more accurate a measure of linearity which ranges from 0 (linear) to .5 (non-linear). We calculated L1-norm as previously defined by Nevozhay, Zal & Balazsi, Nat. Commun. 4:1451 (2013). Briefly, we rescaled each dose response curve to the [0,1] range, interpolated the curve through the MATLAB function *interp1*, we calculated L1-norm as the area between dose response curve and the straight line $y=x$ across each dose range. We calculated the maximum inducer dose at which the dose response was still linear using an L1-norm threshold of 0.05.

All plots were generated in MATLAB.

Supporting Tables

	Primer Name	Sequence	Note
P1	SbfI-PhlF-fwd	TGATTTCCTGCAGGAATAGG ACCATGGCCCGGAC	Amplify PhlF repressor gene from 12AAQHGC-PhlF-Repressor (Voigt)
P2	PhlF-NheI-NotI-rev	TCGAATGCGGCCGCTCTCTG CTAGCCACTTTCCGCTTTTTC TTGGGGG	Amplify PhlF repressor gene from 12AAQHGC-PhlF-Repressor (Voigt)
P3	NheI-T2A-mCherry-fwd	GTATATGCTAGCgagggcagagg aagtcttctaa	Amplify T2A-mCherry from pDC-BACH1-T2A-mCherry
P4	mCherry-NotI-rev	TATATGCGGCCGCGTGAATG CAAGTTTActtgtacagctc	Amplify T2A-mCherry from pDC-BACH1-T2A-mCherry
P5	GA-2xophlF-mlin-Frag-fwd	CCCCTGATTCTGTATACGCGT TGACATTGATTATTGACTAGT TATTAATAGTAATCAATT	Gibson assembly primers for Tandem Linearizer Plasmid
P6	GA-2xophlF-mlin-Frag-rev	CCGATTTAGTGCTAATCGAAA TCTCGTAGTACGTGCTAT	Gibson assembly primers for Tandem Linearizer Plasmid
P7	GA-D2iRtetR-mlin-Vec-fwd	TGTCAACGCGTATACAGAATC AGGGGATAACGCAG	Gibson assembly primers for Tandem Linearizer Plasmid
P8	GA-D2iRtetR-mlin-Vec-rev	CGAGATTTCGATTAGCACTAA ATCGGAACCCTAAAGG	Gibson assembly primers for Tandem Linearizer Plasmid
P9	CMV-phlFo-seq-fwd	CATCAAGTGTATCATATGCC	Sequencing primer for promoter variants of phlF linearizer
P10	CMV-phlFo-seq-rev	TCTTCACTACTTCTGGAATAG	Sequencing primer for promoter variants of phlF linearizer
P11	PhlF-rep-seq-rev	TGTTTGTCCACCACCGGTAG AT	Sequencing primer for phlF repressor protein
P12	PhlF-rep-seq-fwd	ACAGACTGCTGACCGAGCAG	Sequencing primer for phlF repressor protein
P13	B-Glob-Intron-Seq-f	GTGAGTTTGGGGACCCTTGA T	Sequencing primer for intron
P14	AmpR-split-rev	TTGATCGTTGGGAACCGGAG	Sequencing primer for Tandem Linearizer Plasmid
P15	AmpR-split-fwd	CTCCGGTTCCCAACGATCAA	Sequencing primer for Tandem Linearizer Plasmid
P16	HygR-fwd	TCTCGTGCTTTCAGCTTCGA	Sequencing primer for Tandem Linearizer Plasmid
P17	FRT-seq-rev	TAGGAACTTCTAGGTACGTG AACC	Sequencing primer for Tandem Linearizer Plasmid

Table S1: Primers used in this study.
The table below lists all oligonucleotide primers and their 5’ to 3’ sequence.

DAPG Dose	P-value
0 µM	0.47
0.5 µM	0.33
1 µM	0.05
1.5 µM	0.37
2.5 µM	0.49
3 µM	0.15
3.5 µM	0.83

Table S2: One-way ANOVA results for phlFd-mLin cell line treated with pairwise DAPG-Dox induction, assessed across Dox dose groups for every DAPG dose.

Dox Dose	P-value
0 ng/mL	0.68
2 ng/mL	0.65
4 ng/mL	0.46
6 ng/mL	0.75

Table S3: One-way ANOVA results for TetR-mLin cell line treated with pairwise DAPG-Dox induction, assessed across DAPG dose groups for every Dox dose.

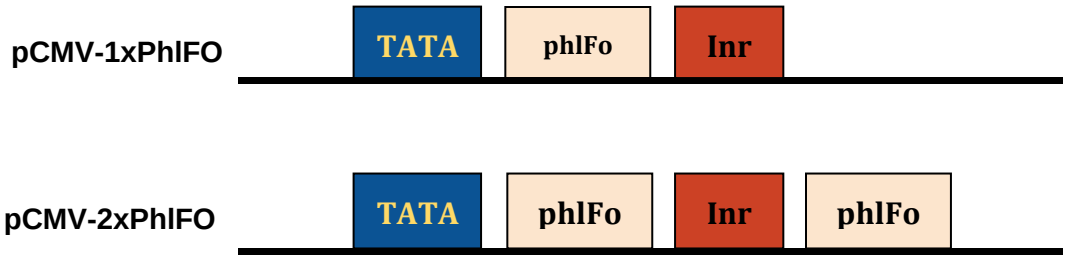
Dox Dose	P-value
0 ng/mL	0.50
2 ng/mL	0.28
4 ng/mL	0.77
6 ng/mL	0.83

Table S4: One-way ANOVA results for phlFd-mLin and TetR-mLin double integrant cell line treated with pairwise DAPG-Dox induction. The ANOVA test was conducted on eGFP fluorescence across Dox dose groups for every DAPG dose.

DAPG Dose	P-value
0 µM	0.38
0.5 µM	0.05
1 µM	0.96
1.5 µM	0.15
2.5 µM	0.41
3 µM	0.35
3.5 µM	0.59

Table S5: One-way ANOVA results for phlFd-mLin and TetR-mLin double integrant cell line treated with pairwise DAPG-Dox induction. The ANOVA test was conducted on mCherry fluorescence across Dox dose groups for every DAPG dose.

Supporting Figures



Original: CMV-TetO

```
...ACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGG
TGGGAGGTCTATATAAGCATCCCTATCAGTGATAGAGATCAGATCTCCCTATCAGTGATAGAGAGCTGTT
TTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGGACTCTAGCGTTTAAACTTAAGCTT
```

Variant 1: 2 phlf operator sites

```
...ACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGG
TGGGAGGTCTATATAAGCAATGATACGAAACGTACCGTATCGTTAAGGTTCAGATCATGATACGAAACGT
ACCGTATCGTTAAGGTGCTCCGGGACCGATCCAGCCTCCGGACTCTAGCGTTTAAACTTAAGCTT
```

Variant 2: 1 phlf operator site before initiator

```
...ACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGG
TGGGAGGTCTATATAAGCAATGATACGAAACGTACCGTATCGTTAAGGTTCAGATCGCTGTTTTGACCTC
CATAGAAGACACCGGGACCGATCCAGCCTCCGGACTCTAGCGTTTAAACTTAAGCTT
```

Figure S1: Promoter variants of the PhlF-mLinearizer gene circuits compared to the original TetR-mLin gene circuit.

Figure S2: Plasmid map of the TetR-mLinearizer.

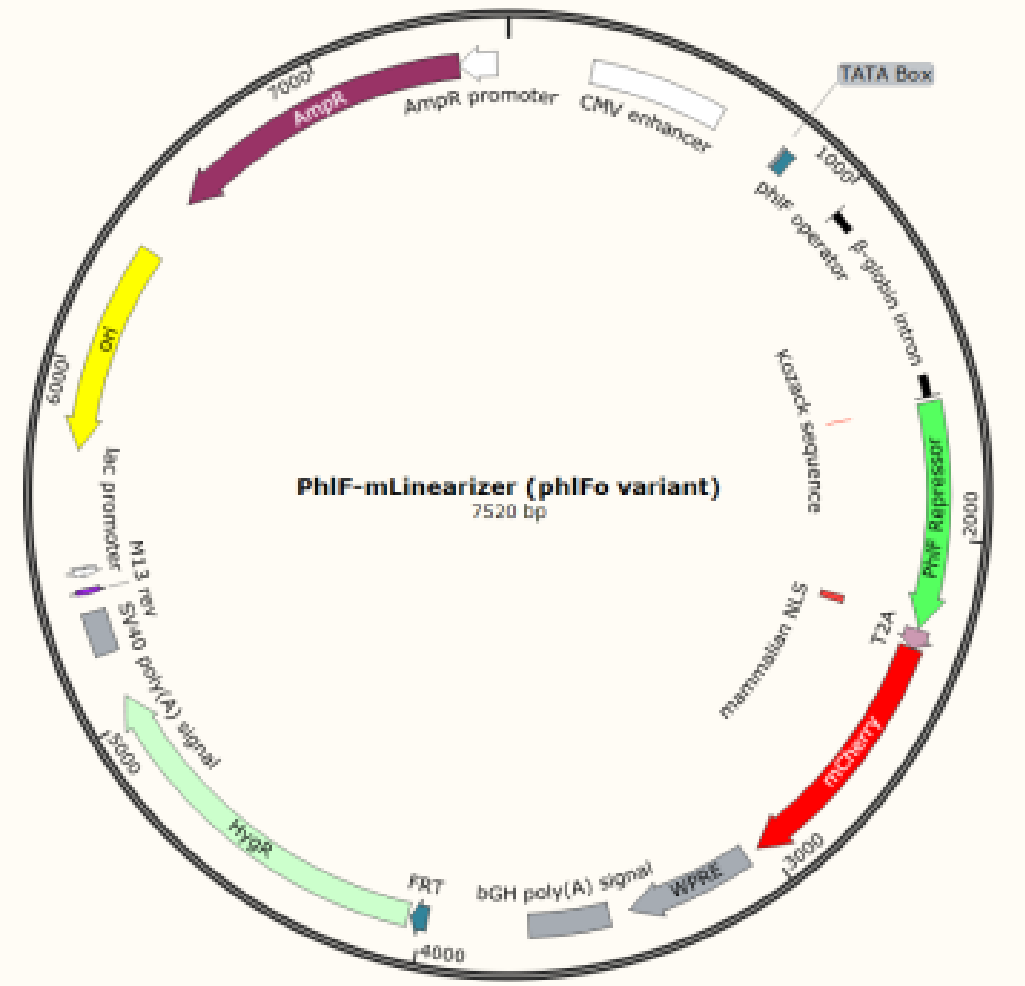


Figure S3: Plasmid map of the phlFs variant of PhlF-mLinearizer.

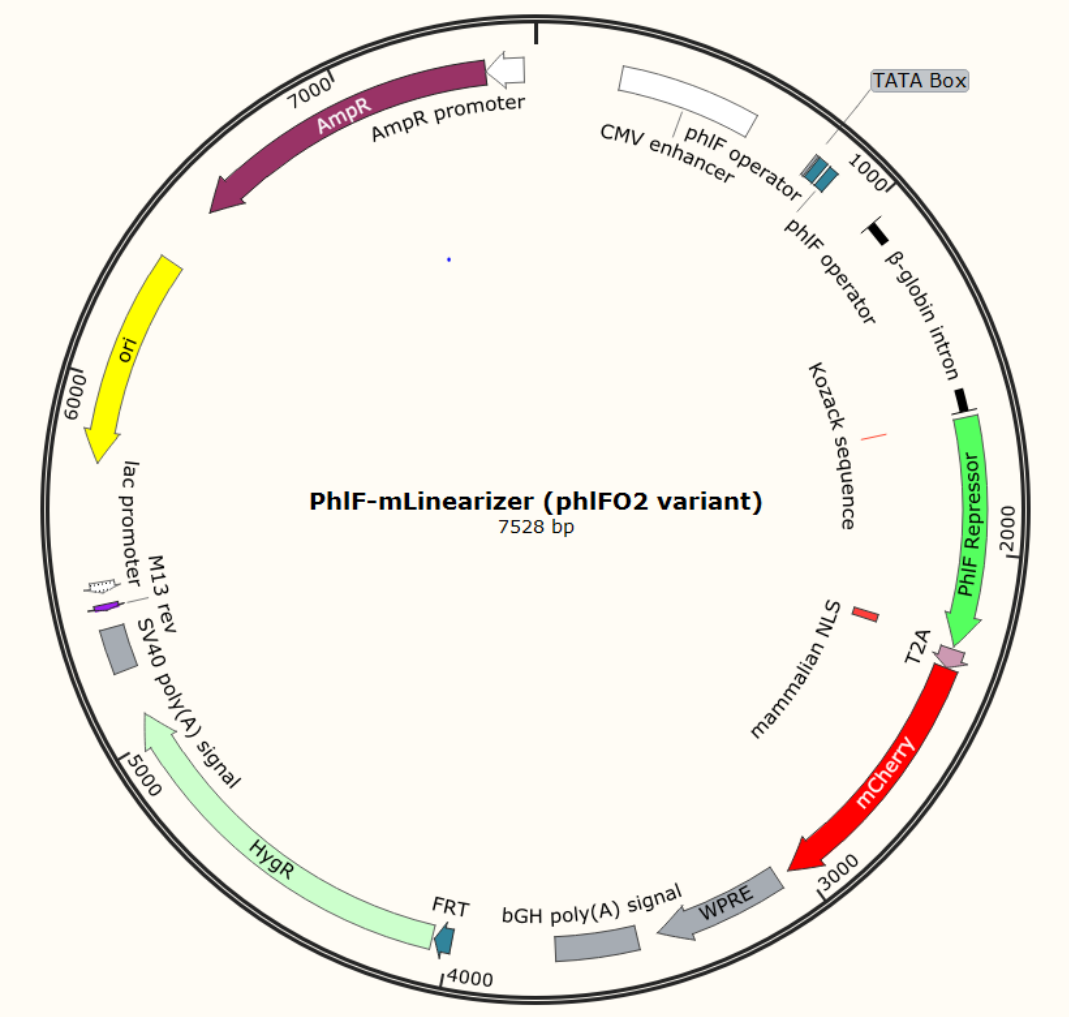


Figure S4: Plasmid map of the phlFd variant of the PhlF-mLinearizer.

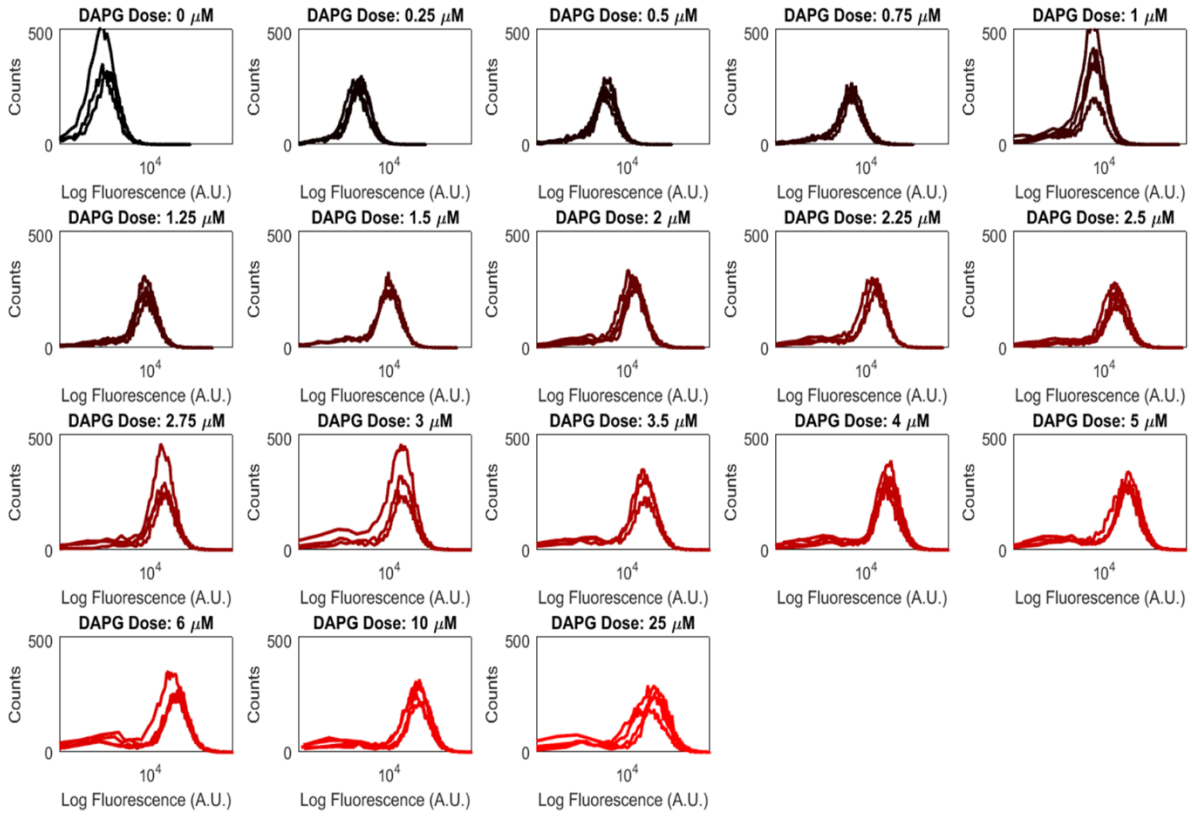


Figure S5: Gene expression distributions of clonal Flp-In-293 cells transfected with phlFs promoter variant of the PhlF-mLin induced for 48 hours with corresponding doses of DAPG. FSC and SSC gated and auto-fluorescence normalized replicate curves for each dose are shown.

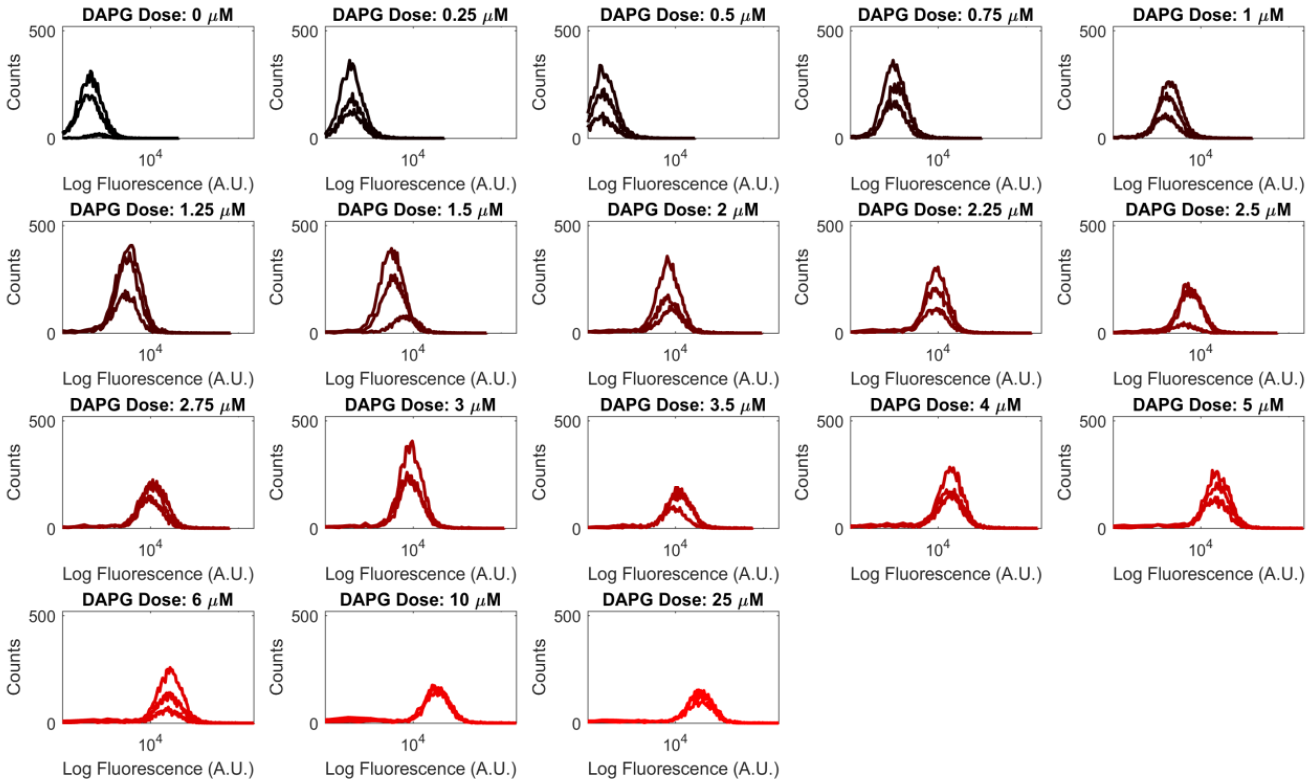


Figure S6: Gene expression distributions of clonal Flp-In-293 cells transfected with phlFd promoter variant of the PhlF-mLin induced for 48 hours with corresponding doses of DAPG. FSC and SSC gated and auto-fluorescence normalized replicate curves for each dose are shown.

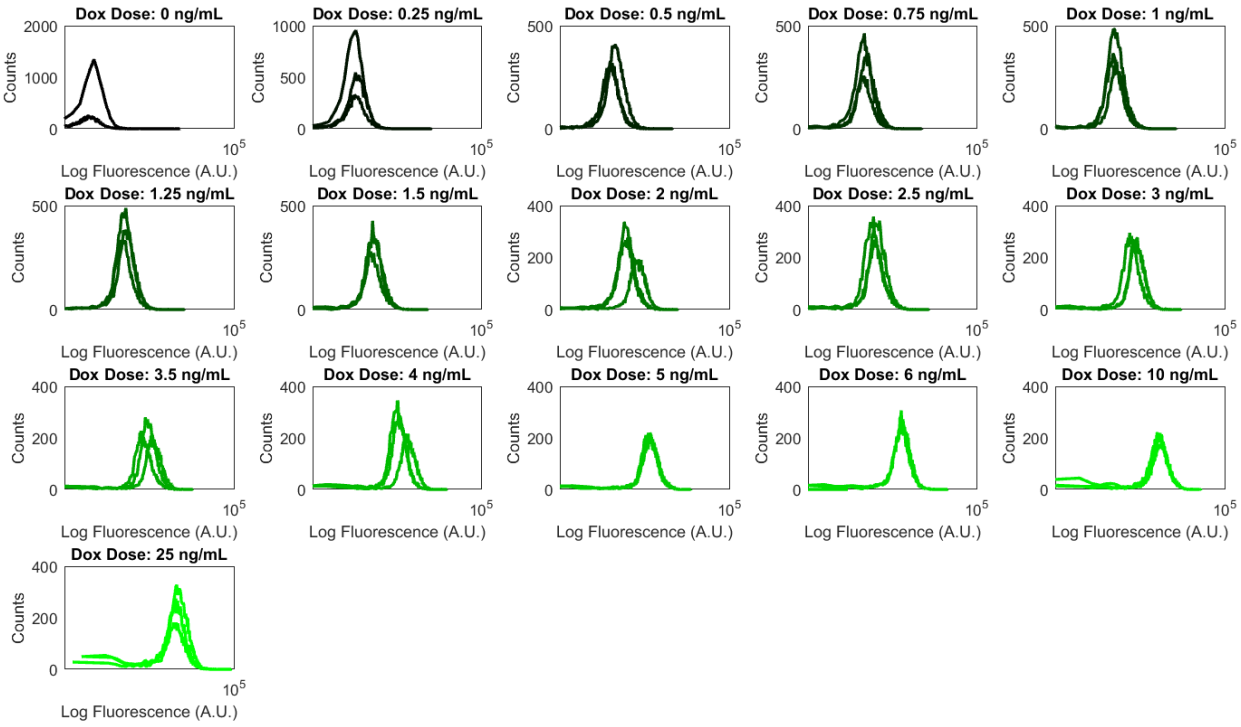


Figure S7: Gene expression distributions of clonal Flp-In-293 cells stably expressing the D2iR promoter variant of the TetR-mLin induced for 48 hours with corresponding doses of Doxycycline. FSC and SSC gated and auto-fluorescence normalized replicate curves for each dose are shown.

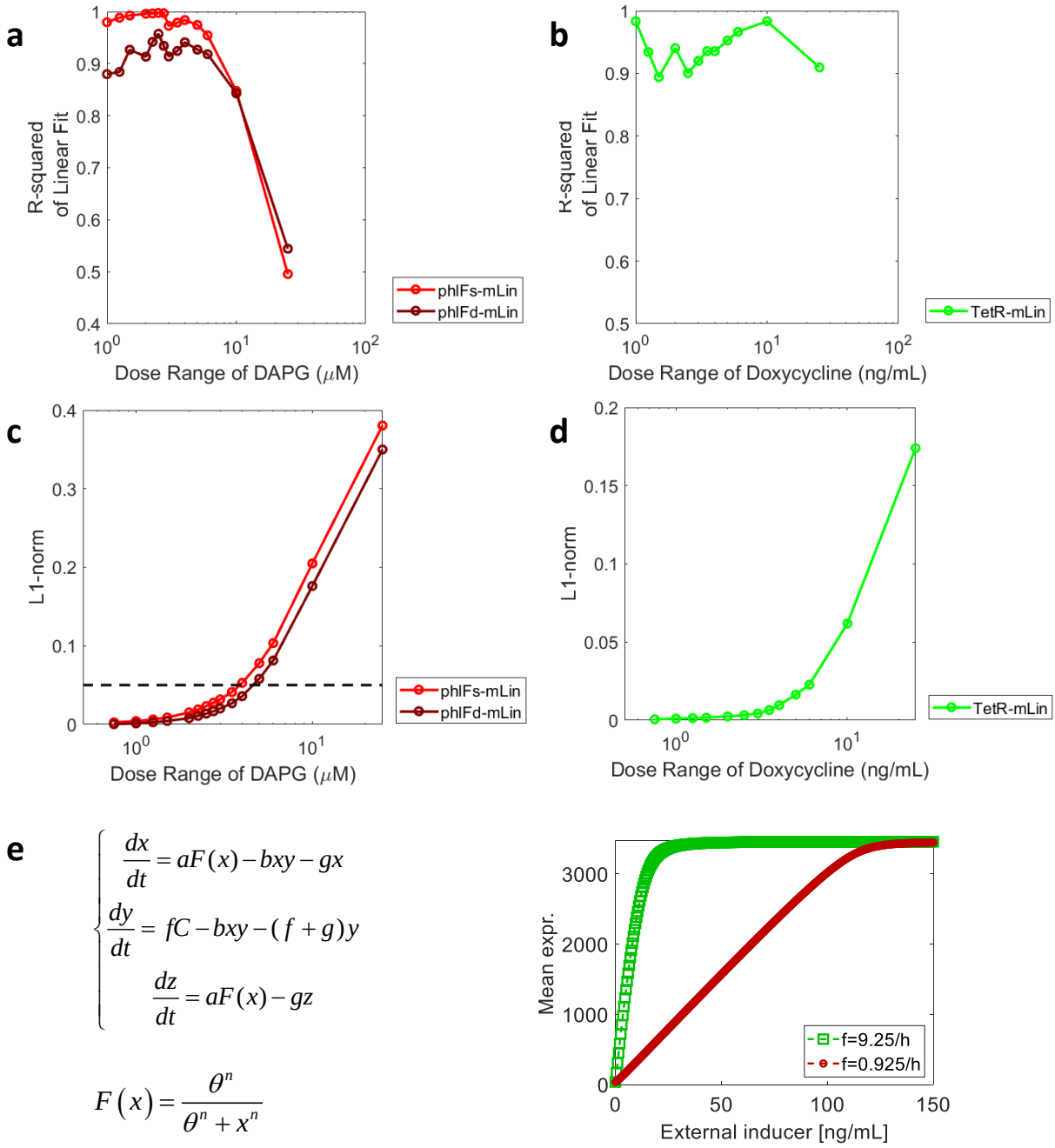


Figure S8: Assessment of linearity and inducer-sensitivity for linearizer gene circuits. (a) R^2 values computed for line fits across a moving dose range window for PhlFs-mLin (red) and PhlFd-mLin (dark red) clonal cell lines. (b) R^2 values computed for TetR-mLin. (c) L1-norm statistic computed across a moving dose range window for phlFs-mLin (red) and phlFd-mLin (dark red) PhlF. Black dashed line indicates threshold used to calculate range of linearity for each gene circuit. (d) L1-norm calculated for TetR-mLin. (e) Mathematical model (ordinary differential equations) for the linearizer system. Variables: x=free TetR, y=internal inducer, z=reporter. Parameters: g=1/24 h^{-1}, growth/dilution rate;

f=inducer diffusion rate; *b*=10, inducer-repressor binding rate; *a*=100 h^{-1}, max. protein synthesis rate; θ=4, *n*=2, Hill parameters. (f) The model indicates that the lower inducer-sensitivity of PhlF-mLin most likely originates from slower DAPG diffusion across the cell membrane compared to Dox.

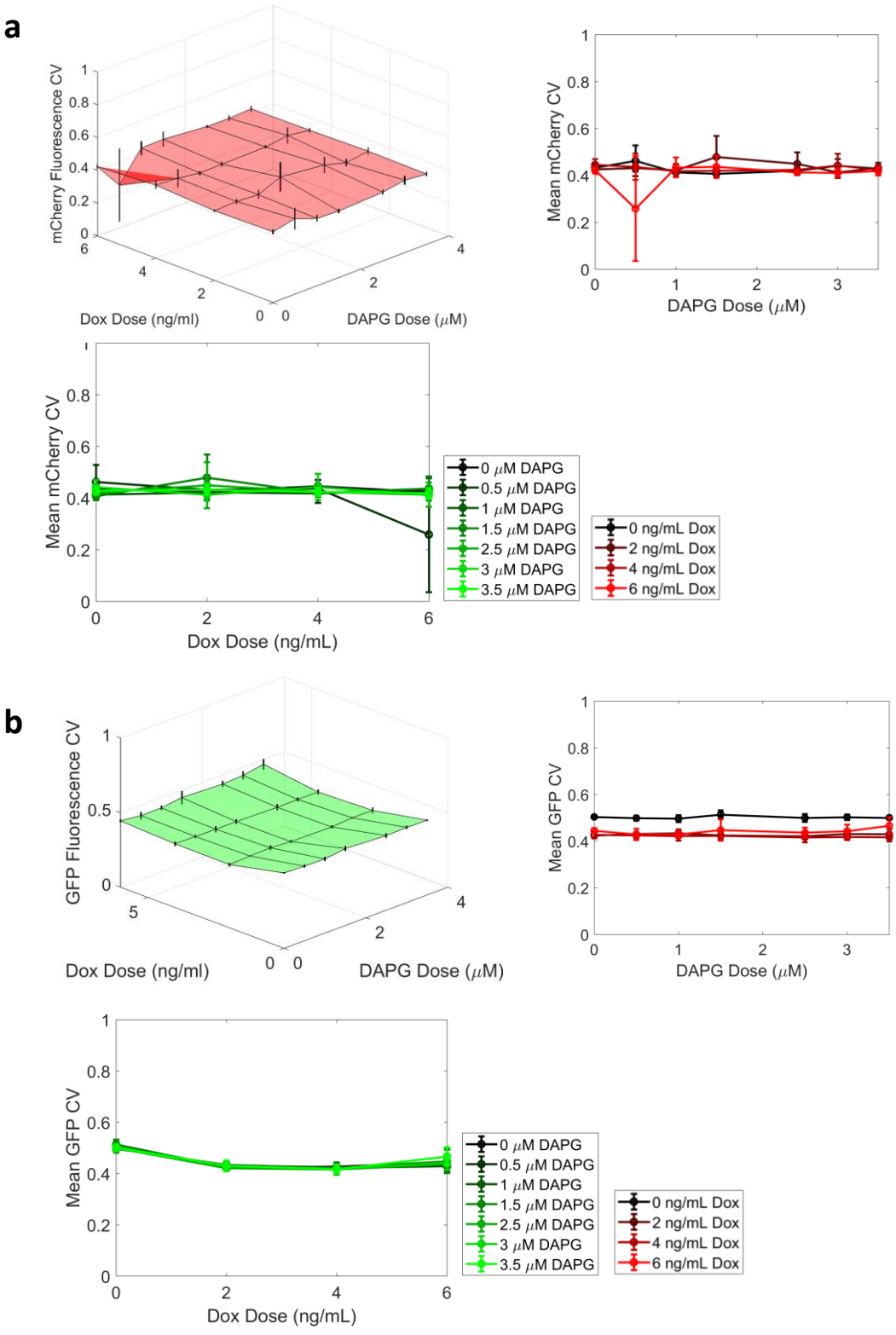
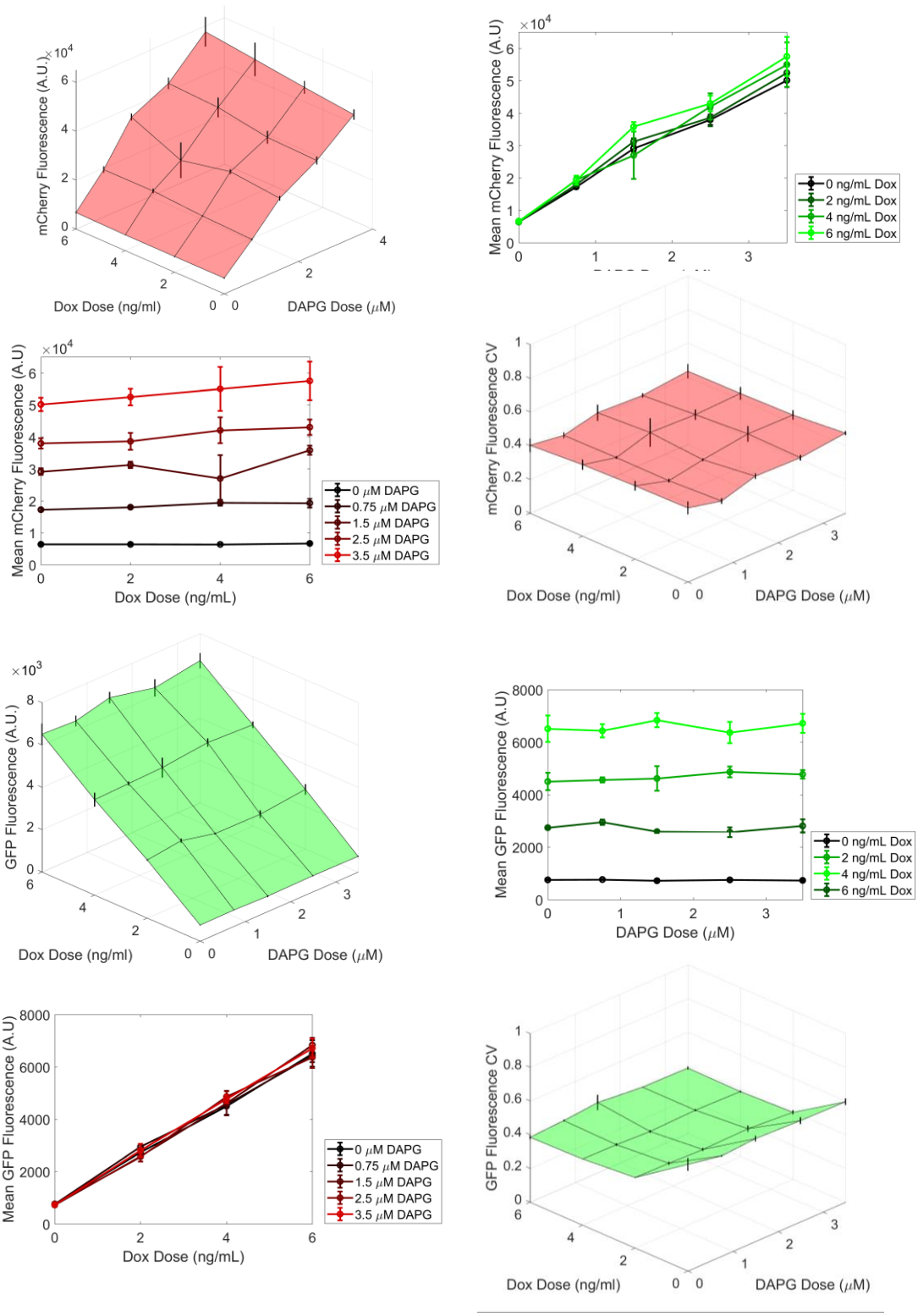


Figure S9: Assessing cross induction CV. (a) 2-dimensional CV dose response of clonal HEK 293 Flp-In cells harboring stably integrated phlFd variant of the PhlF-mLin gene circuit. Plane representation of CV of mCherry fluorescence expression averaged from 3 replicates across pairwise DAPG and Doxycycline induction doses. (b) 2-dimensional CV dose response of clonal HEK 293 Flp-In cells harboring stably integrated TetR-mLin gene circuit. Plane representation of CV of GFP fluorescence expression averaged from 3 replicates across pairwise DAPG and Doxycycline induction doses.



Figures S10: Assessing orthogonality of multiplexed linearizer gene circuits. 2-dimensional dose response of clonal HEK 293 Flp-In cells harboring stably integrated phlFd-mLin and TetR-mLin circuits. Plane representations of mean mCherry and GFP fluorescence expression averaged from 3 replicates across pairwise DAPG and Doxycycline induction doses. Plane representations of mean mCherry and GFP coefficient of variation (CV) (n=3) across pairwise DAPG and Doxycycline induction doses.

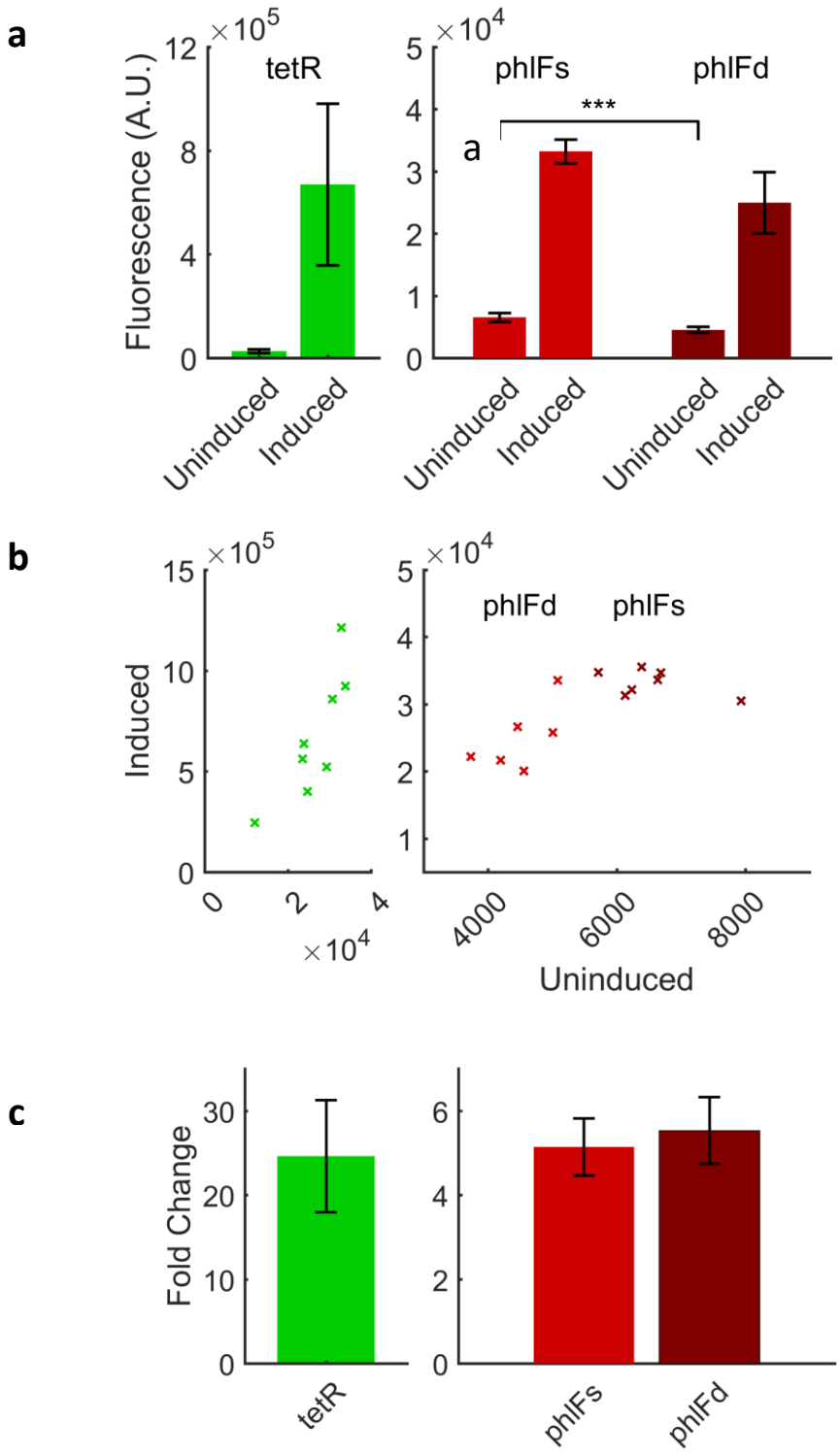


Figure S11: Assessing variation between clones bearing linearizer gene circuits. (a) GFP fluorescence expression differences between uninduced and induced (0 ng/mL, 10 ng/mL) HEK 293 Flp-In TetR-mLin clones and mCherry fluorescence expression of uninduced and induced (0 μM, 2 μM) HEK 293 Flp-In PhlF-mLin clones. PhlFd variants, on average, have significantly lower basal expression ($p<.001$). (b) Phase plots of uninduced and induced expression for individual clones. Correlation coefficients r_{tetR}=0.827, r_{phlFs}=-.493 , r_{phlFd}=.661. (c) Fold changes between uninduced and induced doses for tetR-mLin, phlFs-mLin and phlFd-mlin clones.

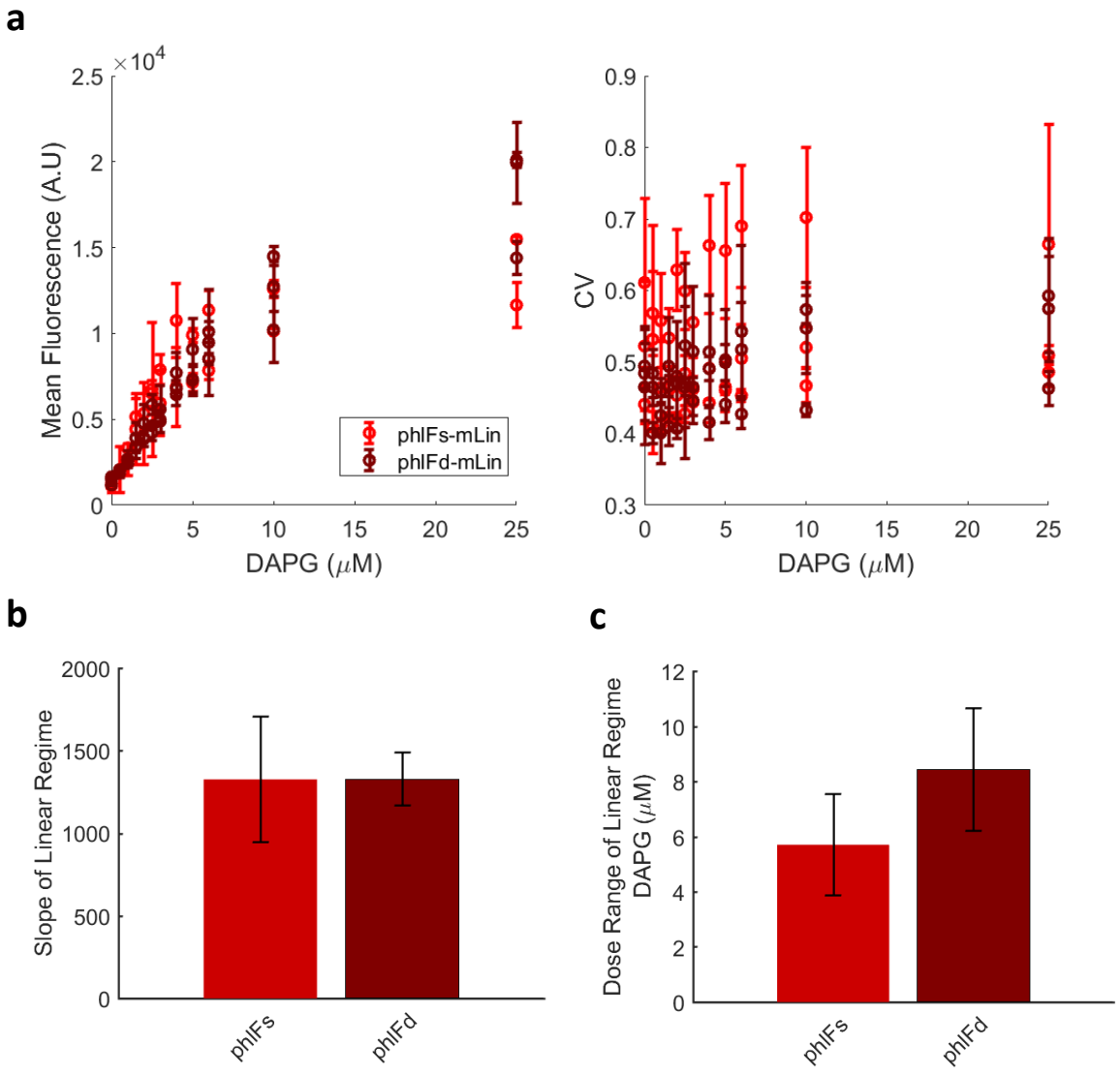


Figure S12: Assessing dose response functions of phlF-mLin clones. (a) Inducer dose response profiles of mean fluorescence and CV for clones integrated with phlFs-mLin or phlFd-mLin (n=3 clones per variant). PhlFd-mLin clones had on average lower CV than phlFs-mLin clones (*p=.0059*). (b) Slope of the linear regime between clonal variants. (c) Dose range of the linear regime as assessed by L1-norm analysis on clonal inducer dose response.