

## Split Fluorescent Proteins for *C. elegans*

[View On GitHub](#)

[Download PDF](#)

This project is maintained by [Maria Ingaramo](#) in the [York lab](#), and was funded by [Calico Life Sciences LLC](#)

## Supplementary Text and Figures

Figure S1. Split-wrmScarlet sequence comparison to mScarlet.

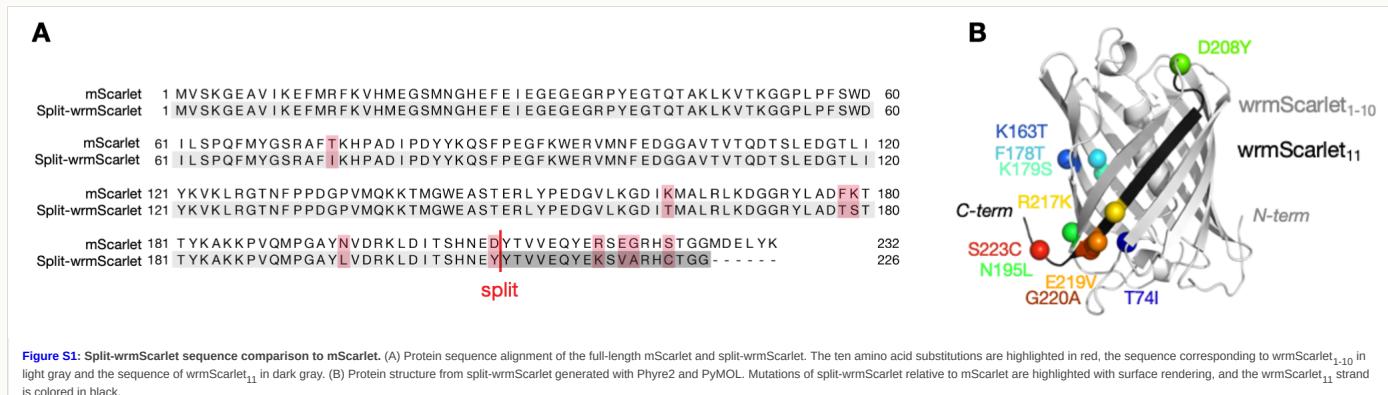
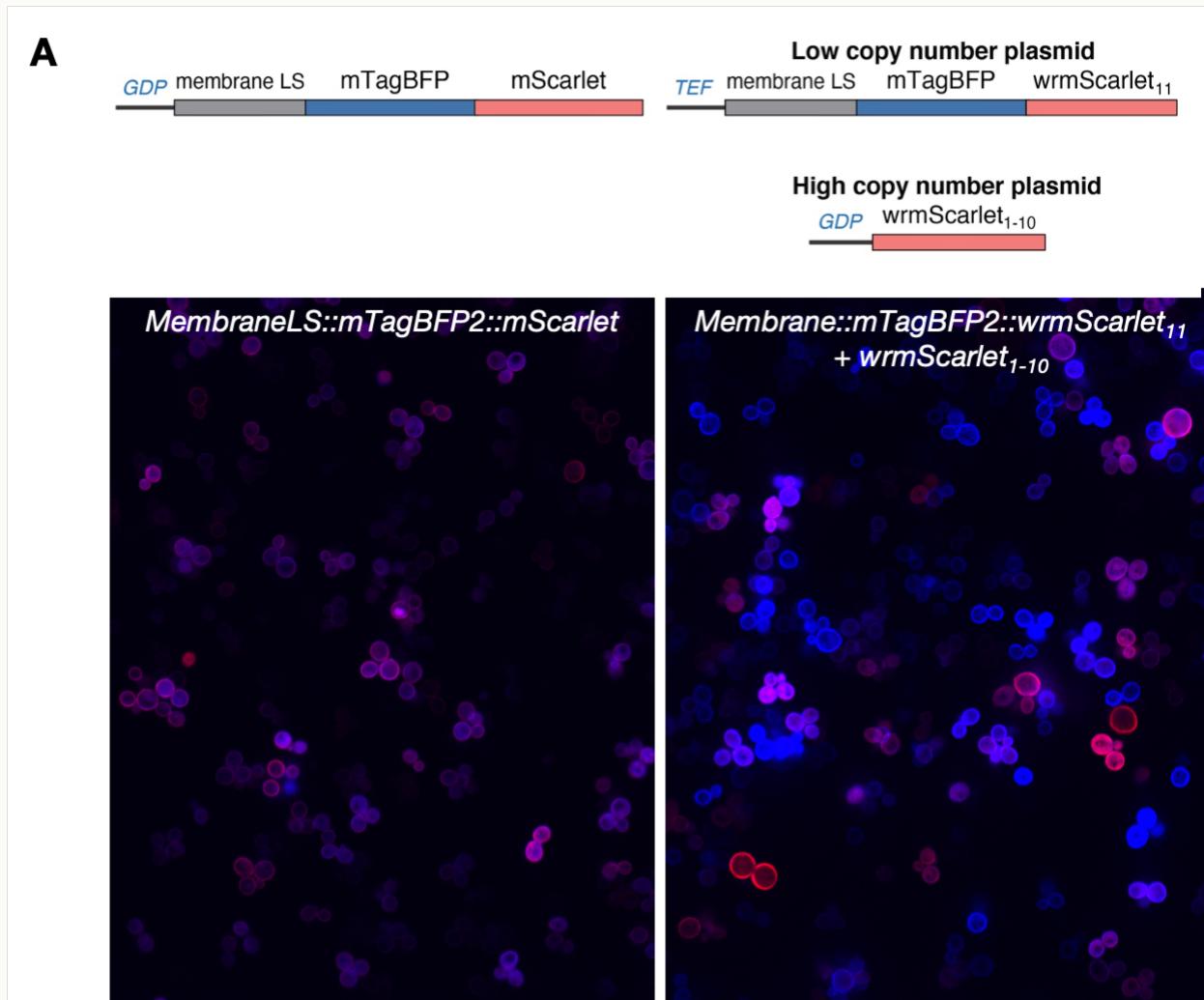
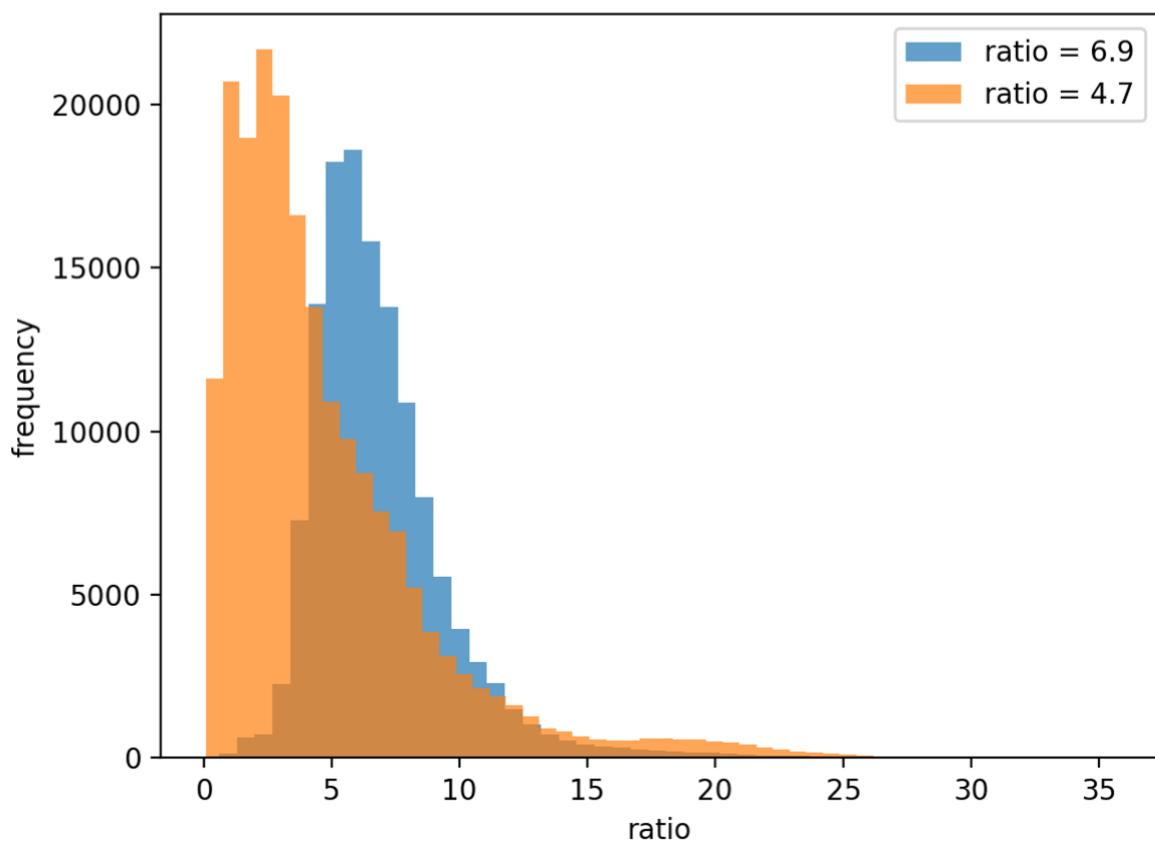


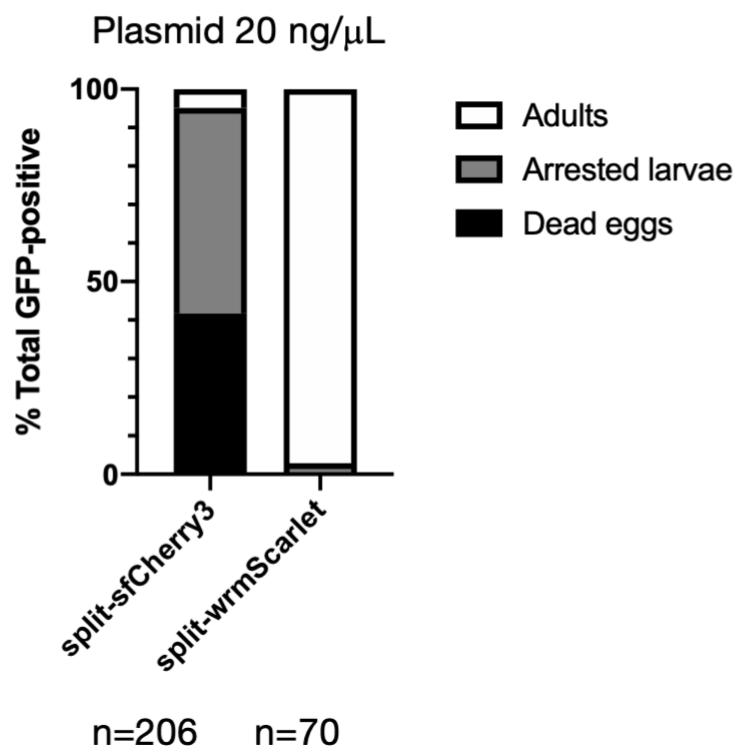
Figure S2. Split-wrmScarlet brightness in *S. cerevisiae*.



**B**

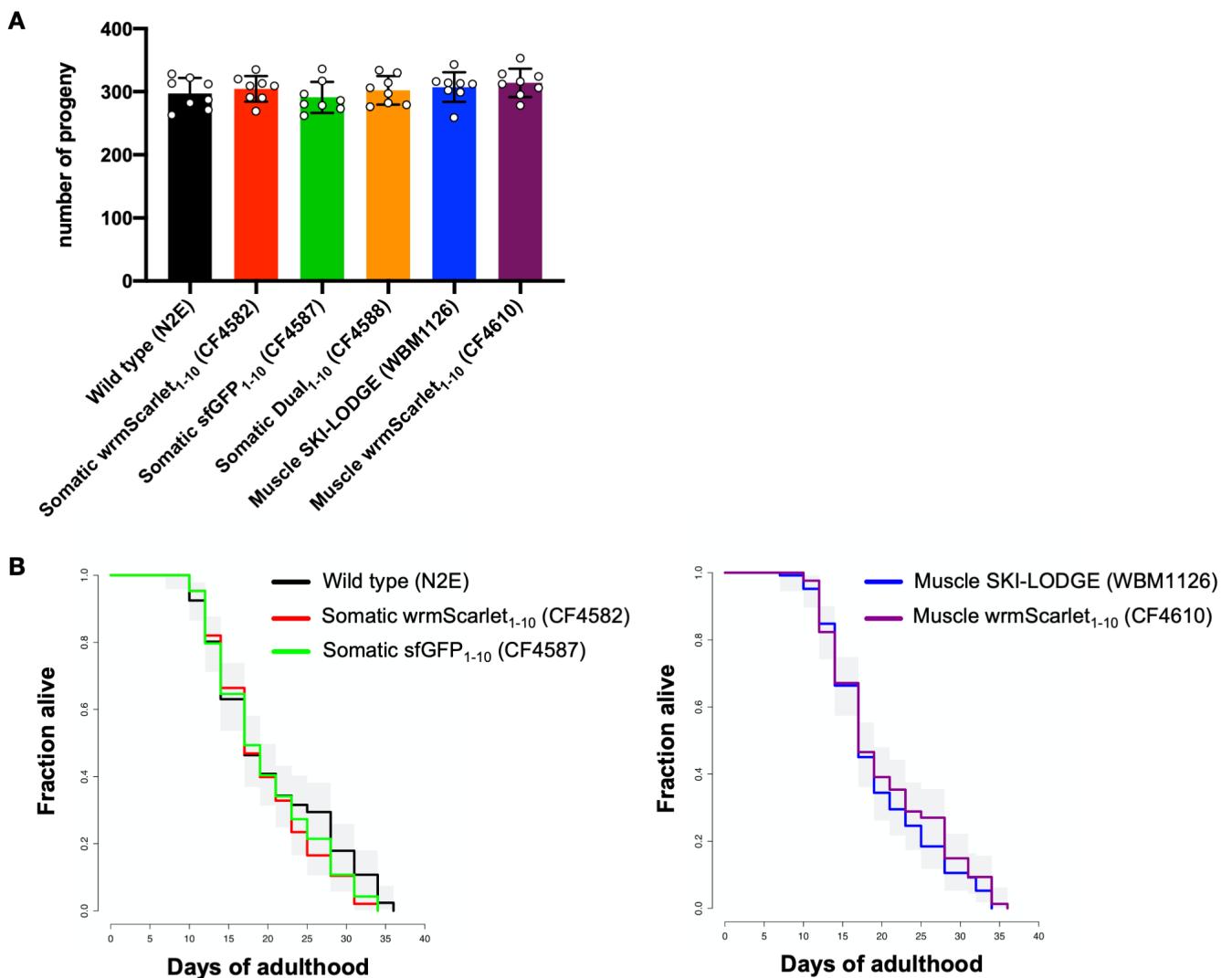
**Figure S2:** Split-wrmScarlet brightness in *S. cerevisiae*. (A) Composite display of red and blue channels for membrane-localized mTagBFP-mScarlet (wild-type) fusion or wrmScarlet<sub>1-10</sub> plus membrane localized mTagBFP-wrmScarlet<sub>11</sub> in yeast. Images were acquired and are displayed under identical conditions. Note that the heterogeneity inherent to expression from plasmids is large, but split-wrmScarlet is capable of brightness levels similar to the parent protein. A schematic of the plasmids transformed is presented above each image. (B) Histograms displaying the pixel-per-pixel ratio of red to blue fluorescence for background corrected, masked images. mTagBFP/full-length Scarlet ratios are displayed in blue, and split-wrmScarlet in orange. The inset displays the average red/blue ratio.

Figure S3. Developmental toxicity in worms expressing split-sfCherry3 in somatic nuclei.

**A****B**

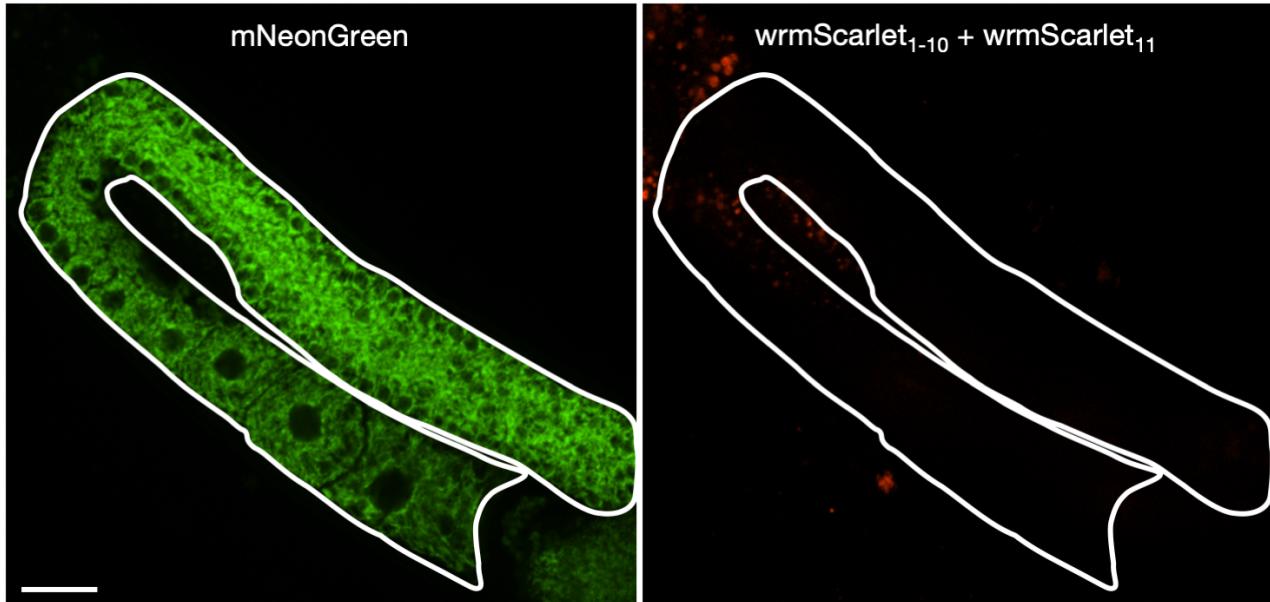
**Figure S3:** Developmental toxicity in worms expressing split-sfCherry3 in somatic nuclei. (A) Schematic of the plasmids encoding split-wrmScarlet and split-sfCherry3 used for comparison. Each plasmid consists of a large FP<sub>1-10</sub> fused to mNeonGreen, and the corresponding small FP<sub>11</sub> fused to the blue fluorescent protein 2 (mTagBFP2), preceded with three SV40 nuclear localization sequences (NLS). The T2A sequence ensures the separation of NLS::mTagBFP2::FP<sub>11</sub> and the corresponding mNeonGreen::FP<sub>1-10</sub>. (B) Quantification of mNeonGreen-positive animals into one of three classes, dead eggs, larvae or adults.

Figure S4. Brood size and lifespan of wrmScarlet<sub>1-10</sub> and sfGFP<sub>1-10</sub> lines.



**Figure S4:** Brood size and lifespan of wrmScarlet<sub>1-10</sub> and sfGFP<sub>1-10</sub> lines. Split-wrmScarlet<sub>1-10</sub> and split-sfGFP<sub>1-10</sub> lines produced wild-type numbers of progeny (A) and a wild-type lifespan (B). Genotypes: N2E (wild-type), CF4582 (muls252[Peft-3:::wmrScarlet<sub>1-10</sub>::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III), CF4587 (muls253[(Peft-3:::sfGFP<sub>1-10</sub>::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III), CF4588 (muls253[Peft-3:::sfGFP1-10::unc-54 3'UTR, Cbr-unc-119(+)], muls252[Peft-3:::wmrScarlet<sub>1-10</sub>::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III), CF4610 (muls257[pmyo-3:::wmrScarlet<sub>1-10</sub>::unc-54 3'UTR] I) and WBM1126 (wbm1s61[myo-3p::3XFLAG::dpy-10 crRNA::unc-54 3'UTR] I). Supplementary tables show survival statistics for all lifespan experiments.

**Figure S5.** Tissue-specific split-wrmScarlet fluorescence in the germline is undetectable.

**A****B**

**Figure S5:** Tissue-specific split-wrmScarlet fluorescence in the germline is undetectable. (A) Schematic of the plasmid encoding *Psun-1::mNeonGreen::linker::wmrScarlet<sub>11</sub>::tbb-2 3'UTR* (left), which was injected into the (MosSCI) strain PHX1797 carrying a single, integrated copy of *Psun-1::wmrScarlet<sub>1-10</sub>::sun-1 3'UTR* (right). (B) Images of animal expressing *mNeonGreen::linker::wmrScarlet<sub>11</sub>* and *wmrScarlet<sub>1-10</sub>* in the germline. Scale bar, 20  $\mu$ m.

## Supplementary Materials and Methods

### Mammalian cell culture.

HEK293T cells (ATCC # CRL-3216) were cultured in high-glucose DMEM supplemented with 10% FBS, 1 mM glutamine and 100  $\mu$ g/mL penicillin/streptomycin (Gibco). An mScarlet<sub>1-10</sub> cDNA codon-optimized for mammalian expression was fused to the C-terminus of eGFP and cloned into a pCDH lentiviral expression vector (SFFV GFP-mScarlet<sub>1-10</sub>). Lentivirus was prepared using standard protocols [Kamiyama 2016] and used to infect HEK293T cells. A polyclonal population of GFP-mScarlet<sub>1-10</sub> positive cells was isolated by FACS (using GFP fluorescence) and served as parental cell line for further experiments. For CLTA-N CRISPR engineering, *S. pyogenes* Cas9/sgRNA ribonucleoprotein complexes were prepared as in [Leonetti 2016], mixed with HDR donor templates and electroporated into of GFP-mScarlet<sub>1-10</sub> cells by nucleofection.

### CLTA-N mScarlet<sub>11</sub> donor library.

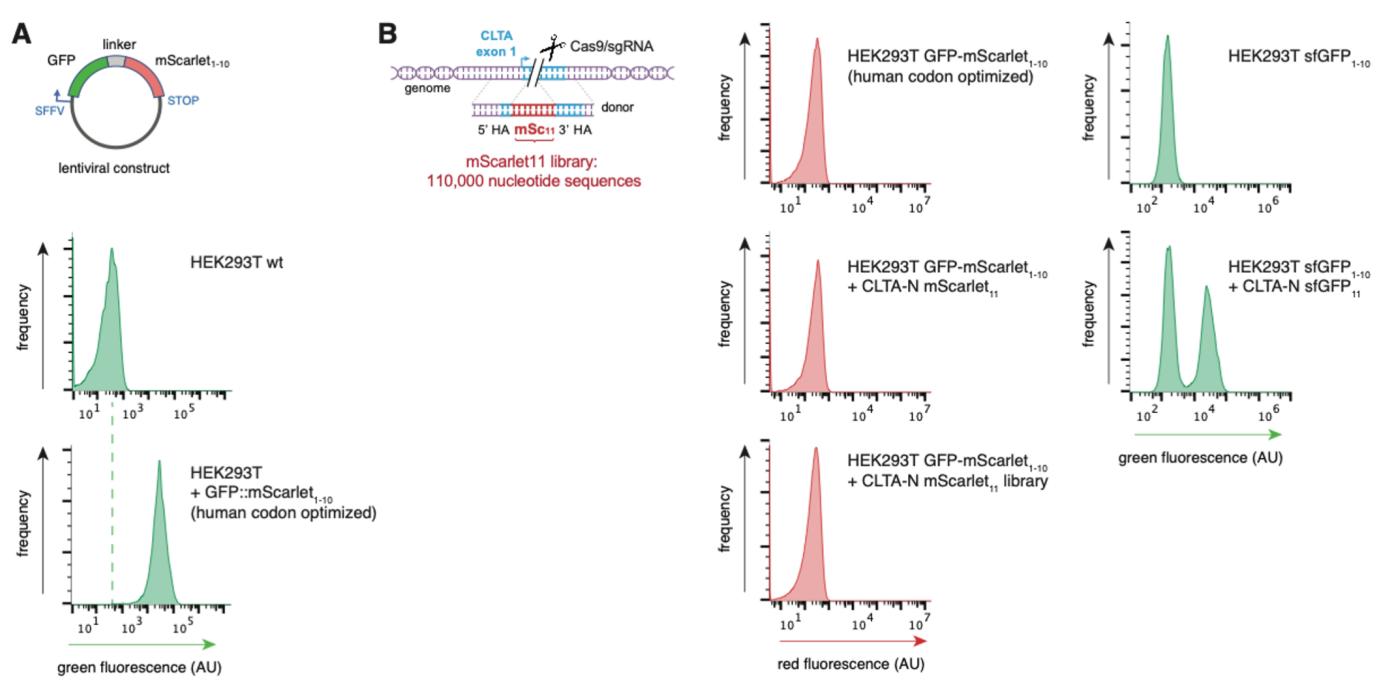
A cDNA pool of degenerate mScarlet<sub>11</sub> sequences was generated by oligonucleotide synthesis (GeneScript) and homology arms for HDR-mediated insertion at CLTA N-terminus were appended by PCR (Supplementary Material – Tables for full sequences). Library diversity was verified by Illumina MiSeq deep-sequencing.

## Supplementary Results

### Split mScarlet screening in mammalian cells

We tested the applicability of the wrmScarlet<sub>1-10</sub> system for mammalian cell engineering but were surprisingly unsuccessful at detecting fluorescence. We designed a human codon-optimized mScarlet<sub>1-10</sub> cDNA and expressed it as a C-terminal GFP fusion in HEK293T cells by lentiviral transduction. Expression of GFP verified the successful expression of the fusion protein (Figure S6A). However, subsequent expression of mScarlet<sub>11</sub> fragments did not give rise to detectable red fluorescence despite numerous attempts. We reasoned that the mScarlet<sub>11</sub> amino-acid sequence might be sub-optimal for complementation in human cells and synthesized a library of degenerate mScarlet<sub>11</sub> sequences covering any possible single and double amino-acid mutants. Using an established assay for CRISPR-based knock-in of sequences at the CLTA N-terminus (a highly expressed gene in HEK293T cells [Leonetti 2016]), neither our original wrmScarlet<sub>11</sub> sequence nor its mutant library enabled detectable complementation (Figure S6B, left panels). By contrast, a control experiment using the GFP<sub>1-10</sub>/GFP<sub>11</sub> system showed a high level of knock-in and complementation in HEK293T (Figure S6B, right panels). It is possible that mScarlet<sub>1-10</sub> is expressed in a non-functional form in human cells, or that its binding to mScarlet<sub>11</sub> is occluded by competing interactions (with cellular chaperones, for example). In addition, we did not attempt complementation on primary non-transformed cell lines, like WI-38 cells, whose different proteostasis network and chaperones could aid split mScarlet folding. At this point, more experiments will be required to fully test the portability of split wrmScarlet to mammalian systems.

Figure S6. Screen for split-mScarlet fluorescence in mammalian cells.



**Figure S6:** Screen for split-mScarlet fluorescence in mammalian cells. (A) FACS histograms of human codon-optimized mScarlet<sub>1–10</sub> expressed as a C-terminal GFP fusion. GFP expression verifies successful expression of the fusion protein in HEK293T cells by lentiviral transduction. (B) Schematic of the CRISPR-based knock-in design for screening single and double mutants of mScarlet<sub>11</sub>. Left panel shows that neither our original mScarlet<sub>11</sub> sequence nor its mutant library enabled detectable complementation as detected by FACS. Right panel shows that the control experiment using the sGFP<sub>11</sub>-sGFP<sub>11</sub> system displays high levels of knock-in and complementation in HEK293T cells.

## Supplementary Tables

## DNA Sequences of wrmScarlet<sub>1-10</sub>, wrmScarlet<sub>11</sub>, sfGFP<sub>1-10</sub> and sfGFP<sub>11</sub>

*C. elegans* lines expressing single-copy of wrmScarlet<sub>1-10</sub> and/or sfGFP<sub>1-10</sub>

### C. elegans strains, genotypes and sources

Strain	Genotype	Source
N2E	wild type	Kenyon Lab
CF4582	muls252[Peft-3::wrmScarlet1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III	This study
CF4586	muls252[Peft-3::wrmScarlet1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III; vha-13(muls262[wrmScarlet11::vha-13]) V	This study
CF4587	muls253[Peft-3::IGFP1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III	This study
CF4588	muls253[Peft-3::IGFP1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; muls252[Peft-3::wrmScarlet1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III	This study
CF4589	muls253[Peft-3::IGFP1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III; vha-13(muls268[sfGFP11::vha-13]) V	This study
CF4592	muls253[Peft-3::IGFP1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III; his-3(muls255[his-3::sfGFP11]) V	This study
CF4594	muls252[Peft-3::wrmScarlet1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III; his-3(muls258[his-3::wrmScarlet11]) V	This study
CF4601	muls252[Peft-3::wrmScarlet1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III; fib-1(muls254[wrmScarlet11::fib-1]) V	This study
CF4602	muls253[Peft-3::IGFP1-10::unc-54 3'UTR, Cbr-unc-119(+)], muls252[Peft-3::wrmScarlet1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III; fib-1(muls244[wrmScarlet11::fib-1]), his-3(muls255[his-3::sfGFP11]) V	This study
CF4603	muls252[Peft-3::wrmScarlet1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III; eat-6(muls269[eat-6::wrmScarlet11]) V	This study
CF4608	muls252[Peft-3::wrmScarlet1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III; his-3(muls267[his-3::wrmScarlet11(x3)]) V	This study
CF4610	muls257[myo-3::wrmScarlet1-10::unc-54 3'UTR]	This study
CF4611	muls257[myo-3::wrmScarlet1-10::unc-54 3'UTR] I; fib-1(muls254[wrmScarlet11::fib-1]) V	This study
CF4612	muEx690[myo-3::tagBFP2::stCherry3_11::T2A::NeonGreen::sfCherry3_1_10::fib-1 3'UTR]	This study
CF4613	muEx691[myo-3::tagBFP2::wrmScarle11-10::T2A::NeonGreen::wrmScarlet1-10::fib-1 3'UTR]	This study
CF4614	muls252[Peft-3::wrmScarlet1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; tbb-2(muls260[wrmScarlet11::tbb-2]), unc-119(ed3) III	This study
CF4615	muls252[Peft-3::wrmScarlet1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III; tomn-20(muls261[tomn-20::wrmScarlet11]) V	This study
CF4616	muls252[Peft-3::wrmScarlet1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III; vha-13(muls264[wrmScarlet11(x2)::vha-13]) V	This study
COP1795	knulS1785 [pNU1687 (Plef-858::sfGFP1-10::unc-54 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III	Nemametrix
PHX731	vha-13(syb731[wrmScarlet::vha-13]) V	SunyBiotech
PHX1797	sybsI66[Psn-1::wrmScarlet1-10::sun-1 3'UTR, Cbr-unc-119(+)] II; unc-119(ed3) III	SunyBiotech
CA1200	iEsI57[eft-3p::TIR-1::mRuby::unc-54 3'UTR + Cbr-unc-119(+)] II; unc-119(ed3) III	CGC
WBM1126	wbmls61[myo-3p::3XFLAG::dp-10 crRNA::unc-54 3'UTR] I	CGC

## crRNAs, HDR templates and oligonucleotide sequences

#### A. Sequences of crRNA and HDR template used for wrmScarlet<sub>11</sub> knock-in experiments

#### A'. Sequences of crRNA and HDR template used for sfGFP<sub>11</sub> knock-in experiments

#### B. DNA template for wrmScarlet tandems HDR donor sequence - plasmids

Template name Tagged term DNA template for wrmScarlet tandems - dsDNA (lower case: homology arms; red: wrmScarlet11; blue: linkers)

#### C. Primers used to PCR wrmScarlet tandems HDR donor sequence from plasmid

Primer Name	Sequence of DNA oligo
his-3_F	CCAAAGAGGAGTTCTTCAAATATC
his-3_R	GATTAAATATTGTGCCCTAAAG
vha-13_F	GGTTTATTTGATTTCTTTGATTTCC
vha-13_R	CCATCTTCGGCTGTGAC

D. Sequences of crRNA and HDR template used to generate wrmScarlet<sub>1-10</sub> and sfGFP<sub>1-10</sub> strains

Primers long and short used to amplify wrmScarlet<sub>1-10</sub> to generate CF4582

## Primers long and short used to amplify Peft-3 to generate CF4587

Primers used to amplify wrmScarlet<sub>1-10</sub> to generate CF4610

## Plasmids

Pmyo-3::mTagBFP2::wrnScarlet11::T2A::mNeonGreen::wrnScarlet1-10::fib-1 UTR (*C. elegans*)

Pmyo-3::mTagBFP2::sfCherry311::T2A::mNeonGreen::sfCherry31-10::fib-1 UTR (*C. elegans*)

### Adult lifespans of strains in this study

Strain	Events / n initial	Mean lifespan +/- SEM (Days)	Median lifespan	% mean lifespan change vs. N2	P-value (log-rank) vs. N2
N2E	105 / 134	20.20 +/- 0.74	17		
CF4582	118 / 136	19.27 +/- 0.57	17	-4.60	0.19
CF4587	109 / 129	19.56 +/- 0.64	17	-3.17	0.29
WBM1126	112 / 127	19.31 +/- 0.61	17	-4.41	0.28
CE4610	108 / 128	20.19 +/- 0.69	17	0.10	0.92

Strain	Events / n initial	Mean lifespan +/- SEM (Days)	Median lifespan	% mean lifespan change vs. WBM1126	P-value (log-rank) vs. N2
WBM1126	112 / 127	19.31 +/- 0.61	17		
CF4610	108 / 128	20.18 +/- 0.69	17	4.51	0.33

## References

- [Kamiyama 2016] Kamiyama, D. et al." Versatile protein tagging in cells with split fluorescent protein. " *Nature Communications* 7, 11046–9 <https://doi.org/10.1038/ncomms11046>
- [Leonetti 2016] Leonetti, M. D., Sekine, S., Kamiyama, D., Weissman, J. S. & Huang, B." A scalable strategy for high-throughput GFP tagging of endogenous human proteins." *Natl. Acad. Sci. U.S.A.* 113, E3501–8 (2016). <https://doi.org/10.1073/pnas.1606731113>



Hosted on

[GitHub Pages](#)