# Split-wrmScarlet and split-sfGFP: tools for faster, easier fluorescent labeling of endogenous proteins in *Caenorhabditis elegans*

Jérôme Goudeau (ID) ,[1],* Catherine S. Sharp,[2] Jonathan Paw,[1] Laura Savy,[3] Manuel D. Leonetti,[3] Andrew G. York,[1] Dustin L. Updike (ID) ,[2] Cynthia Kenyon,[1],* and Maria Ingaramo[1],*

[1]Calico Life Sciences LLC, South San Francisco, CA 94080, USA
[2]Mount Desert Island Biological Laboratory, Bar Harbor, ME 04672, USA
[3]Chan Zuckerberg Biohub, San Francisco, CA 94158, USA

*Corresponding authors: Calico Life Sciences LLC, 1170 Veterans Boulevard, South San Francisco, CA 94080, USA. jerome@calicolabs.com (J.G.); cynthia@calicolabs.com (C.K.); ingaramo@calicolabs.com (M.I.)

## Abstract

We create and share a new red fluorophore, along with a set of strains, reagents and protocols, to make it faster and easier to label endogenous *Caenorhabditis elegans* proteins with fluorescent tags. CRISPR-mediated fluorescent labeling of *C. elegans* proteins is an invaluable tool, but it is much more difficult to insert fluorophore-size DNA segments than it is to make small gene edits. In principle, high-affinity asymmetrically split fluorescent proteins solve this problem in *C. elegans*: the small fragment can quickly and easily be fused to almost any protein of interest, and can be detected wherever the large fragment is expressed and complemented. However, there is currently only one available strain stably expressing the large fragment of a split fluorescent protein, restricting this solution to a single tissue (the germline) in the highly autofluorescent green channel. No available *C. elegans* lines express unbound large fragments of split red fluorescent proteins, and even state-of-the-art split red fluorescent proteins are dim compared to the canonical split-sfGFP protein. In this study, we engineer a bright, high-affinity new split red fluorophore, split-wrmScarlet. We generate transgenic *C. elegans* lines to allow easy single-color labeling in muscle or germline cells and dual-color labeling in somatic cells. We also describe a novel expression strategy for the germline, where traditional expression strategies struggle. We validate these strains by targeting split-wrmScarlet to several genes whose products label distinct organelles, and we provide a protocol for easy, cloning-free CRISPR/Cas9 editing. As the collection of split-FP strains for labeling in different tissues or organelles expands, we will post updates at doi.org/10.5281/zenodo.3993663

Keywords: *C. elegans*; Cas9; CRISPR; genome engineering; GFP; protein localization; mScarlet; germline; wrmScarlet

## Introduction

Genetically expressed fluorophores are essential tools for visualizing and quantifying cellular proteins. In *Caenorhabditis elegans*, fluorescent proteins have traditionally been introduced on extra-chromosomal arrays (Kimble *et al.* 1982; Mello et al. 1991) or via MosSCI-based integration (Frøkjær-Jensen *et al.* 2008, 2012). These methods have enabled important discoveries but can also lead to artifacts due to supraphysiological gene-expression levels and lack of endogenous regulatory control. In recent years, the repertoire of *C. elegans* transgenic tools has expanded (see Nance and Frøkjær-Jensen 2019, for review), particularly due to advances in CRISPR/Cas9 genome-editing technologies (Paix *et al.* 2014; Dickinson and Goldstein 2016). CRISPR/Cas9 allows precise transgene insertion by homology-directed repair (HDR) and can be used to label an endogenous gene at its native locus with a fluorescent protein (Friedland *et al.* 2013; Dokshin *et al.* 2018; Farboud *et al.* 2019; Vicencio *et al.* 2019).

However, relative to CRISPR/Cas9-mediated integration of smaller transgenes, genomic insertion of large DNA fragments like those encoding fluorescent proteins remains a challenge, both because repair with double-stranded templates is less efficient than repair with single-stranded oligodeoxynucleotide donors (ssODN) (Farboud *et al.* 2019), and because of the requirement for cloning to prepare the HDR donor template. Recent methods such as "hybrid" (Dokshin *et al.* 2018) and "nested" (Vicencio *et al.* 2019) CRISPR remove the need for cloning but still require preparation of the DNA template or several rounds of injections and selection of transgenic progeny. As a result, using CRISPR with small ssODN templates is currently faster, easier, cheaper, and more efficient than with large templates. In our lab, we routinely make *C. elegans* genome edits with short ssODNs with almost guaranteed success. In contrast, in our experience, large edits using double-stranded DNA templates have higher failure rates and are more time-consuming.
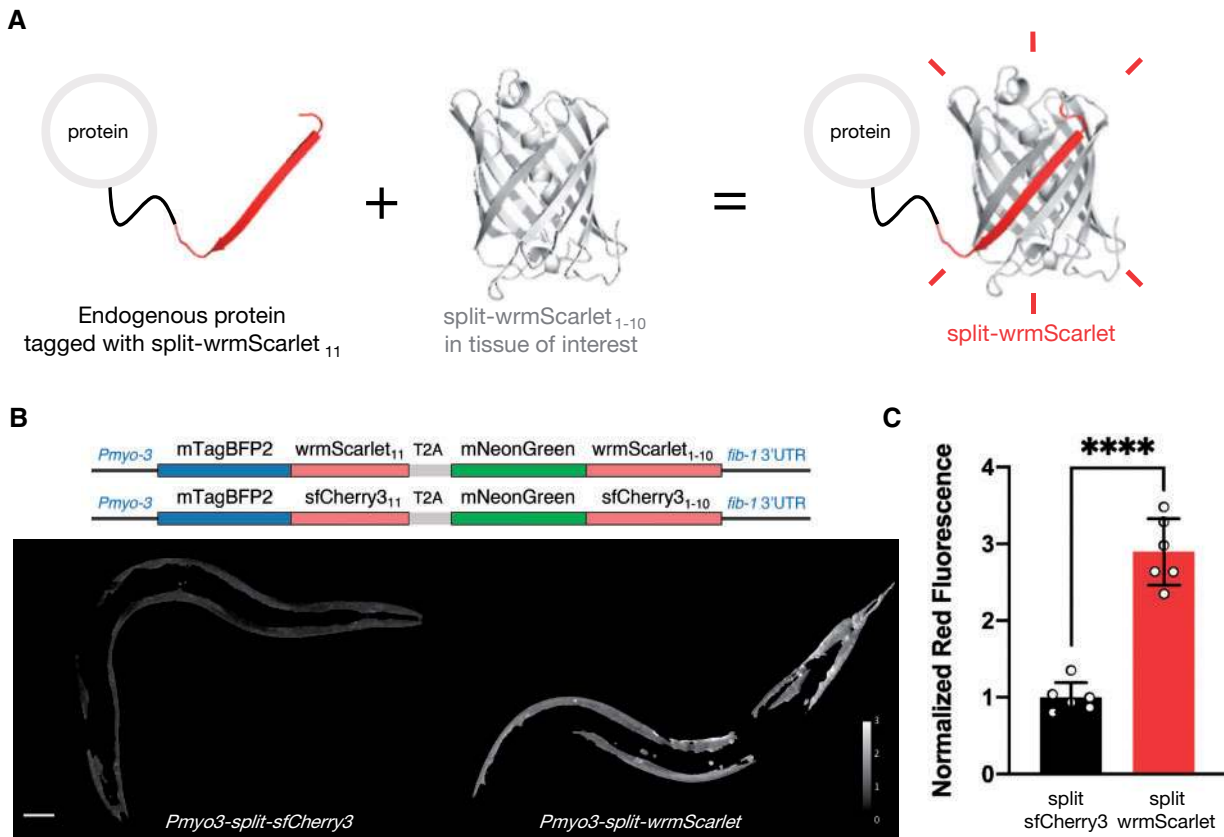
**Figure 1** Engineering and evaluating split-wrmScarlet. (A) Principle of endogenous protein labeling with split-wrmScarlet. The protein structure from split-wrmScarlet was generated using Phyre2 and PyMOL. (B) Schematic of the plasmids encoding split-wrmScarlet and split-sfCherry3. Each plasmid consists of the large $FP_{1-10}$ sequence fused to mNeonGreen, and the corresponding small $FP_{11}$ sequence fused to mTagBFP2. The T2A sequence ensures that mTagBFP2::$FP_{11}$ and the corresponding mNeonGreen::$FP_{1-10}$ are separated. The images are representative displays of the ratio of red to green fluorescence intensity from images acquired under identical conditions after background subtraction and masking with the same threshold. Scale bar, 50 μm. (C) Emission intensities from split-sfCherry3 and split-wrmScarlet normalized to mNeonGreen. Mean ± SD. Circles are individuals ($n = 6$ for each split fluorescent protein). ****$P < 0.0001$.

Our preferred approach is to combine the utility of full-length fluorescent proteins with the convenience of short genomic edits, by using high-affinity asymmetrically split fluorescent proteins (Cabantous *et al.* 2005). These fluorophores typically separate a GFP-like protein between the 10th and 11th strands of the beta barrel, splitting it asymmetrically into a large ($FP_{1-10}$) and a small ($FP_{11}$) fragment. The fragments are not individually fluorescent, but upon binding one another, recapitulate the fluorescent properties of an intact fluorophore (Figure 1A). Unlike the low-affinity split fluorescent proteins used in BiFC assays (Hu *et al.* 2002), high-affinity binding between the fragments is critical here. Our preferred approach for tagging a new cellular protein begins with a *C. elegans* strain expressing the large $FP_{1-10}$ fragment in cells of interest, unattached to any cellular protein. This way, only the small $FP_{11}$ fragment (<72 nt) needs to be inserted to tag the target protein, which will only fluoresce in compartments where it can bind the large fragment. These short insertions tend to be faster, easier, and more reliable than inserting a >600 nt full-length fluorescent protein (Paix *et al.* 2015; Prior *et al.* 2017, Dokshin *et al.* 2018; Richardson *et al.* 2018). Therefore, collections of *C. elegans* lines stably expressing the large $FP_{1-10}$ in different tissues are an invaluable resource allowing rapid fluorescent tagging in a cell type of choice. Stable lines with red $FP_{1-10}$ fragments would be especially useful, given *C. elegans*' substantial autofluorescence in the GFP channel.

Green and red asymmetrically split fluorescent proteins have been used to combine cell and protein specificity in *C. elegans*

neurons and synapses (Noma *et al.* 2017; Feng *et al.* 2019; He *et al.* 2019); however, these strains used extrachromosomal arrays, not stable lines, which are more time-consuming to maintain and can have variable expression levels. To the best of our knowledge, there is only one available unbound $FP_{1-10}$ stable *C. elegans* line, which expresses sfGFP$_{1-10}$ in the germline (Hefel and Smolikove 2019), and there are no available lines with red $FP_{1-10}$ fragments. Existing red split fluorophores are also much dimmer in *C. elegans* than green ones, despite recent improvements like split-sfCherry3 (Feng *et al.* 2019). In addition, we often struggle to express genome-integrated full-length fluorescent protein fusions in the germline, potentially due to generational silencing.

Here, we describe tools that reduce these obstacles for convenient fluorescent labeling of endogenous *C. elegans* proteins. We engineer split-wrmScarlet, a new split red fluorescent protein based on mScarlet (Bindels *et al.* 2017; El Mouridi *et al.* 2017), which is three times brighter in worms than split-sfCherry3 (https://www.addgene.org/138966). We generate and share *C. elegans* lines carrying single-copy insertions of split-wrmScarlet$_{1-10}$ expressed broadly in somatic cells (https://cgc.umn.edu/strain/CF4582), and specifically in muscle (https://cgc.umn.edu/strain/CF4610). We also describe a novel approach to make *C. elegans* lines with robust germline expression of exogenous proteins that appears to be resistant to generational silencing. We use this approach to make a germline-specific split-wrmScarlet$_{1-10}$ strain (https://cgc.umn.edu/strain/DUP237).

We provide a protocol for an easy, cloning-free method to label endogenous genes with $FP_{11}$s using CRISPR/Cas9, commercially available synthetic ssODN, and microinjection (dx.doi.org/10.17504/protocols.io.bamkic4w). We validate this protocol by targeting split-wrmScarlet$_{11}$ to six different genes whose products have distinct cellular locations. We show that labeling with tandem split-wrmScarlet$_{11}$-repeats increases fluorescence *in vivo*, and we provide the plasmid necessary to generate the dsDNA template (https://www.addgene.org/158533). We also generate a strain expressing an integrated copy of sfGFP$_{1-10}$ (Pédelacq *et al.* 2006) broadly in somatic cells (https://cgc.umn.edu/strain/CF4587), and a strain expressing sGFP2$_{1-10}$ (Köker *et al.* 2018) specifically in the germline (https://cgc.umn.edu/strain/DUP223). Finally, we generate a dual-color strain expressing both sfGFP$_{1-10}$ and split-wrmScarlet$_{1-10}$ in somatic cells (https://cgc.umn.edu/strain/CF4588), for two-color applications such as colocalization studies or organelle interaction. We hope that these resources will facilitate the study of *C. elegans* biology. As the collection of split-FP strains and related resources for labeling different tissues, organelles and proteins expands, we will post updates at doi.org/10.5281/zenodo.3993663.

## Materials and methods
### Mutagenesis and screening

For the initial screenings in *Escherichia coli*, we introduced a 32 amino-acid spacer between the 10th and 11th β-strands of full-length mScarlet in a pRSET vector (Feng *et al.* 2017). This starting construct was nonfluorescent, but we restored low fluorescence levels by introducing the superfolder mutation G220A. Semi-random mutagenesis was carried out using rolling-circle amplification with NNK primers at positions I8, K10, F15, G32, Q43, A45, K46, L47, G52, G53, D60, S63, P64, Q65, F66, S70, R71, T74, K75, D79, Y84, W94, R96, T107, V108, Q110, E115, L125, R126, T128, K139, K140, W144, E145, S147, T148, E149, R150, I162, K163, M164, L175, F178, K179, K183, K185, K186, N195, R198, I202, T203, S204, D208, Y209, T210, V211, V212, E213, Q214, Y215, E216, R217, S218, E219, A220, H222, S223, T224, G225, G226, M227, D228, and E229 with Phusion polymerase (NEB) in GC buffer, followed by pooling of the PCR products, *DpnI* digestion and transformation into BL21(DE3) *E. coli*. These positions covered areas deemed important for brightness or stability, and the interface between $FP_{11}$ and $FP_{1-10}$. Primers were resynthesized if a mutation interfered with neighboring mutagenic primer binding. The brightest three to five colonies were identified using a Leica M165 FC fluorescent stereomicroscope, and their plasmid DNA subjected to a new mutagenesis round. After five rounds, we separated the two fragments of a version of split-wrmScarlet (which had fluorescence comparable to mScarlet) into two *S. cerevisiae* plasmids to test for complementation. Because we did not detect fluorescence, we continued selection using two plasmids in yeast. For screening on two plasmids, a pRSET vector expressing split-wrmScarlet$_{1-10}$ and a pD881-MR vector (ATUM) expressing mTagBFP-split-wrmScarlet$_{11}$ (without the MDELYK tail from the C-terminus) were used to perform the semi-random mutagenesis. The libraries were co-electroporated into *E. coli* and expression was induced with 1% rhamnose and 1 mM IPTG. The library was enriched for fluorescent clones using FACS, and then subcloned to make pRS-GPD-split-wrmScarlet$_{1-10}$ and p416-TEF-membrane-mTagBFP-split-wrmScarlet$_{11}$. The yeast plasmids were co-transformed into a URA$^-$, HIS$^-$, LEU$^-$, MET$^-$ *S. cerevisiae* strain and selected for in SC media without uracil and histidine, and FACS was used again for enrichment of clones with the highest red to

blue ratio. After three rounds of semi-random mutagenesis with the two-plasmid strategy, a final round of random mutagenesis was performed using the GeneMorph II kit (Agilent). Yeast plasmids are available through Addgene (https://www.addgene.org/158585/, https://www.addgene.org/158584/), and *E. coli* plasmid sequences are present in Supplementary Table S5.

### *C. elegans* strains and maintenance

Animals were cultured under standard growth conditions with *E. coli* OP50 at 20°C (Brenner 1974). Strains generated in this work are listed in the Supplementary Table S3.

### Nucleic acid reagents

Synthetic nucleic acids were purchased from Integrated DNA Technologies (IDT), GenScript or Genewiz. For knock-in of a single split-wrmScarlet$_{11}$ or sfGFP$_{11}$ sequence, 200-mer HDR templates were ordered in ssODN form (synthetic ssODN) from IDT. For knock-in of split-wrmScarlet$_{11}$ repeats, HDR templates were ordered in dsDNA form (plasmids) from GenScript or Genewiz. For plasmids injected as extrachromosomal arrays, sequences were synthesized and cloned into the pUC57 vector (Genewiz). The complete set of crRNAs and DNA sequences and plasmids used for the experiments described here can be found in Supplementary Tables S1, S4, and S5.

### Strain generation: CRISPR/Cas9-triggered homologous recombination

CRISPR insertions were performed using published protocols (Paix *et al.* 2015; Paix *et al.* 2016). Ribonucleoprotein complexes (protein Cas9, tracrRNA, and crRNA) and DNA templates were microinjected into the gonad of young adults using standard methods (Evans 2006). Injected worms were singled and placed at 25°C overnight. All crRNA and DNA template sequences used to generate the strains described in this work are listed in the Supplementary Table S4. Split-wrmScarlet$_{11}$ and sfGFP$_{11}$ integrants were identified by screening for fluorescence in the $F_1$ or $F_2$ progeny of injected worms. The co-CRISPR *dpy-10(cn64)* mutation was used as a marker when generating nonfluorescent strains. The CF4582 strain *muIs252[Peft-3::split-wrmScarlet$_{1-10}$::unc-54 3'UTR Cbr-unc-119(+)] II; unc-119(ed3) III* was generated by replacing the *tir-1::mRuby* sequence from the strain CA1200 *ieSi57 II; unc-119(ed3) III* (Zhang *et al.* 2015) with the split-wrmScarlet$_{1-10}$ sequence. The CF4587 strain *muIs253[Peft-3::sfGFP$_{1-10}$::unc-54 3'UTR Cbr-unc-119(+)] II; unc-119(ed3) III* was generated by replacing the *let-858* promoter from the strain COP1795 *knuSi785 [pNU1687(Plet-858::sfGFP$_{1-10}$::unc-54 3'UTR unc-119(+))] II; unc-119(ed3) III* with the *eft-3* (also known as *eef-1A.1*) promoter. Both CF4582 and CF4587 strains were generated using long, partially single-stranded DNA donors (Dokshin *et al.* 2018). The CF4610 strain *muIs257[Pmyo-3::split-wrmScarlet$_{1-10}$::unc-54 3'UTR] I* was generated by inserting the split-wrmScarlet$_{1-10}$ sequence in the WBM1126 strain following the SKI LODGE protocol (Silva-García *et al.* 2019). The strains PHX731 *vha-13(syb731[wrmScarlet::vha-13]) V* and PHX1049 *vha-13(syb1049[gfp::vha-13]) V* were generated by SunyBiotech's CRISPR services. Strains were genotyped by Sanger sequencing of purified PCR products (Genewiz).

### Strain generation: Mos1-mediated single-copy insertion

The COP1795 strain was generated by NemaMetrix's MosSCI services. The PHX1797 strain was generated by SunyBiotech's MosSCI services, using a codon-optimized sequence of split-wrmScarlet$_{1-10}$ with three introns, and engineered to avoid

piRNA recognition transgene silencing (Wu *et al.* 2018; Zhang *et al.* 2018; Supplementary Table S1).

## Strain generation: genetic crosses

The following *C. elegans* strains were created by standard genetic crosses: CF4588 *muIs253[Peft-3::sfGFP$_{1-10}$::unc-54 3'UTR Cbr-unc-119(+)] muIs252[Peft-3::split-wrmScarlet$_{1-10}$::unc-54 3'UTR Cbr-unc-119(+)] II; unc-119(ed3) III* and CF4602 *muIs253[Peft-3::sfGFP$_{1-10}$::unc-54 3'UTR Cbr-unc-119(+)] muIs252[Peft-3::split-wrmScarlet$_{1-10}$::unc-54 3'UTR Cbr-unc-119(+)] II; unc-119(ed3) III; fib-1(muIs254[split-wrmScarlet$_{11}$::fib-1]) his-3(muIs255[his-3::sfGFP$_{11}$]) V*. Nonfluorescent parental lines CF4582, CF4587 and CF4610 generated using *dpy-10(cn64)* co-CRISPR were backcrossed at least once.

## Strain generation: plasmid microinjection

*Peft-3::3NLS::mTagBFP2::split-wrmScarlet$_{11}$::T2A::mNeonGreen::split-wrmScarlet$_{1-10}$::fib-1 3'UTR*, *Peft-3::3NLS::mTagBFP2::sfCherry3$_{11}$::T2A::mNeonGreen::sfCherry3$_{1-10}$::fib-1 3'UTR*, *Pmyo-3::mTagBFP2::split-wrmScarlet$_{11}$::T2A::mNeonGreen::split-wrmScarlet$_{1-10}$::fib-1 3'UTR*, or *Pmyo-3::mTagBFP2::sfCherry3$_{11}$::T2A::mNeonGreen::sfCherry3$_{1-10}$::fib-1 3'UTR* constructs were microinjected at (20 ng/µL) using a standard microinjection procedure (Mello et al. 1991). Germline gene expression was achieved using a microinjection-based protocol with diluted transgenic DNA (Kelly *et al.* 1997), *Psun-1::mNeonGreen::linker::split-wrmScarlet$_{11}$::tbb-2 3'UTR* construct (5 ng/µL) was co-injected with PvuII-digested genomic DNA fragments from *E. coli* (100 ng/µL). Plasmid sequences are listed in Supplementary Table S5.

## Germline strain generation: *glh-1::T2A::split-wrmScarlet$_{1-10}$* and *glh-1::T2A::sGFP2$_{1-10}$*

Using CRISPR/Cas9, the C-terminus of *glh-1* was tagged with either T2A::split-wrmScarlet$_{1-11}$ or T2A::sGFP2$_{1-11}$, a split superfolder GFP variant optimized for brightness and photostability (Köker *et al.* 2018). Fluorescence originating from these full-length fusions was present throughout the cytoplasm and nuclei of adult germ cells and gametes, with the maternally deposited signal persisting through the early stages of embryogenesis and larval development (Supplementary Figure S8, A and B, top panels). After verifying fluorescence, we used a precise CRISPR/Cas9 deletion of either split-wrmScarlet$_{11}$ or sGFP2$_{11}$ to convert these FP$_{1-11}$ strains into FP$_{1-10}$ strains, DUP237 *glh-1(sam140[glh-1::T2A::split-wrmScarlet$_{1-10}$]) I* and DUP223 *glh-1(sam129[glh-1::T2A::sGFP2$_{1-10}$]) I* and corroborated the absence of fluorescence (Supplementary Figure S8, A and B, middle panels). The crRNAs, ssDNAs and dsDNA template sequences are described in Supplementary Tables S1–S4.

## Microscopy

Confocal fluorescence imaging was performed using NIS Elements imaging software on a Nikon confocal spinning disk system equipped with an Andor EMCCD camera, a CSU-X1 confocal scanner (Yokogawa), 405, 488, and 561 nm solid-state lasers, and 455/50, 525/26 and 605/70 nm emission filters. Transgenic animals expressing sfGFP$_{11}$ or split-wrmScarlet$_{11}$ were screened using a Leica M165 FC fluorescent stereomicroscope equipped with a Sola SE-V with GFP and mCherry filters.

## Image analysis

Images were analyzed using Fiji. Image manipulations consisted of maximum intensity projections along the axial dimension, rolling ball radius background subtraction, smoothing, and LUT minimum and maximum adjustments. Masks were created by thresholding and setting the pixels under the threshold cutoff to NaN. Plotting of values per pixel was carried out in python 3, using numpy and matplotlib. When performing normalizations for split-sfCherry3 versus split-wrmScarlet, the red channel was divided by the green channel (mNeonGreen::FP$_{1-10}$) because the localization of both fragments is expected to be the same (cytosolic). For normalization of signals where mTagBFP::FP$_{11}$ was targeted to the membrane, the blue channel was used instead of the green channel.

## Mounting worms for microscopy

Pads made of 3% agarose (GeneMate) were dried briefly on Kimwipes (Kimtech) and transferred to microscope slides. Approximately 10 µL of 2 mM levamisole (Sigma) was pipetted onto the center of the agarose pad. Animals were transferred to the levamisole drop, and a coverslip was placed on top before imaging.

## Brood size analysis

Eight single synchronized adults grown at 20°C were transferred to fresh plates every 24 h until cessation of reproduction, and the number of viable progeny produced by each worm was scored.

## Developmental toxicity assay

Ten N2E wild-type animals were microinjected with either *Peft-3::3NLS::mTagBFP2::split-wrmScarlet$_{11}$::T2A::mNeonGreen::split-wrmScarlet$_{1-10}$::fib-1 3'UTR* or *Peft-3::3NLS::mTagBFP2::sfCherry3$_{11}$::T2A::mNeonGreen::sfCherry3$_{1-10}$::fib-1 3'UTR* construct at (20 ng/µL) and were singled. mNeonGreen-positive F$_1$ animals were scored and their development was monitored for up to 5 days from egg-laying. The number of fluorescent dead eggs, arrested larvae (*i.e.*, animals never reaching adulthood) or adults were scored for each group.

## Comparison of split-sfCherry3 to split-wrmScarlet in muscle

Ten N2E wild-type animals were microinjected with either *Pmyo-3::mTagBFP2::split-wrmScarlet$_{11}$::T2A::mNeonGreen::split-wrmScarlet$_{1-10}$::fib-1 3'UTR*, or *Pmyo-3::mTagBFP2::sfCherry3$_{11}$::T2A::mNeonGreen::sfCherry3$_{1-10}$::fib-1 3'UTR* constructs were microinjected at (20 ng/µL). F$_1$ animals expressing mNeonGreen in muscle were selected for comparison.

## Lifespan assays

NGM plates were supplemented with 5-Fluorouracil (5-FU, Sigma, 15 µM) (Goudeau *et al.* 2011) in order to prevent progeny from hatching and with kanamycin sulfate to prevent bacterial contamination (Sigma, 25 µg/mL). Animals fed with kanamycin-resistant OP50 were scored manually as dead or alive, from their L4 larval stage defined as day 0. A worm was considered alive if it moved spontaneously or, in cases where it wasn't moving, if it responded to a light touch stimulus with a platinum wire. Animals that crawled off the plates, had eggs that accumulated internally, burrowed or ruptured were censored and included in the analysis until the time of censorship.

## Structure prediction and rendering of split-wrmScarlet

Phyre2 was used to predict the three-dimensional modeling in intensive mode with default parameters (Kelley *et al.* 2015). The 3D model obtained was visualized using PyMOL (v2.2.0).

## Statistical analysis

Differences in fluorescence intensity between groups were compared using unpaired *t*-test with Welch's correction. Data are presented as means ± SD. Kaplan-Meier estimates of survival curves were calculated using *survival* (v2.38–3) and *rms* (v4.5–0) R packages and differences were tested using log-rank test. The number of animals used in each experiment is indicated in the figure legends.

## Data availability

Strains expressing a single-copy of split-wrmScarlet$_{1-10}$ and/or sfGFP$_{1-10}$ CF4582 (https://cgc.umn.edu/strain/CF4582), CF4587 (https://cgc.umn.edu/strain/CF4587), CF4588 (https://cgc.umn.edu/strain/CF4588), CF4610 (https://cgc.umn.edu/strain/CF4610), DUP223 (https://cgc.umn.edu/strain/DUP223), and DUP237 (https://cgc.umn.edu/strain/DUP237) are available via the *Caenorhabditis* Genetics Center (CGC). The vectors pJG100 carrying *Peft-3::split-wrmScarlet*$_{1-10}$*::unc-54* 3'UTR (www.addgene.org/138966), pJG103 carrying split-wrmScarlet$_{11}$ x3 tandem repeats (www.addgene.org/158533), yeast plasmids p416-TEF-membrane localization signal-mTagBFP-split-wrmScarlet$_{11}$-TEF terminator (www.addgene.org/158585) and pRS423-GPD-split-wrmScarlet$_{1-10}$-CYC1 terminator (www.addgene.org/158584) are deposited, along with sequences and maps, at Addgene. Other strains and plasmids are available upon request. The authors state that all data necessary for confirming the conclusions presented here are represented fully within the article. A detailed protocol to generate *C. elegans* with sfGFP$_{11}$ and/or split-wrmScarlet$_{11}$ integrants is available at dx.doi.org/10.17504/protocols.io.bamkic4w. Supplementary material is available at figshare: https://doi.org/10.25386/genetics.13635614.

## Results

### Split-wrmScarlet

To engineer split-wrmScarlet, first we introduced a 32 amino acid spacer between the 10th and 11th β-strands of full-length yeast-codon optimized mScarlet, following a strategy described previously (Feng *et al.* 2017). We subjected the spacer-inserted mScarlet sequence to several rounds of semi-random mutagenesis in *E. coli*, generating a version with fluorescence comparable to the full-length mScarlet when expressed in bacteria. However, upon separating the two fragments into two *S. cerevisiae* plasmids to test for complementation, we observed no detectable fluorescence in yeast. We decided to continue with several rounds of selection of new mutant libraries in yeast using FACS, by fusing the small fragment (without the MDELYK C-terminus residues) from our brightest *E. coli* clone to a plasma-membrane-targeted blue FP (mTagBFP), and expressing the large fragment from a high-copy number vector containing a strong promoter. The brightest resulting protein, which we named split-wrmScarlet, contained 10 amino acid substitutions relative to the C-terminal truncated mScarlet (Supplementary Figure S1, A and B). Fluorescence microscopy of yeast containing both plasmids corroborated that split-wrmScarlet showed the expected localization and can reach brightness comparable to that of intact mScarlet in yeast (Supplementary Figure S2, A and B).

### Split-wrmScarlet is threefold brighter than split-sfCherry3 in *C. elegans* muscles

In order to compare split-wrmScarlet to split-sfCherry3, the brightest published red split-FP at the time of the experiment, we combined the FP$_{1-10}$ and FP$_{11}$ fragments into a single plasmid for each fluorophore. Specifically, we generated worm-codon-optimized plasmids encoding three nuclear localization signals (NLS), mTagBFP2, FP$_{11}$, a T2A peptide-bond-skipping sequence, mNeonGreen and the corresponding FP$_{1-10}$, driven by the *eft-3* promoter (Supplementary Figure S3A). Each FP$_{11}$ was linked to mTagBFP2 in order to reduce the risk of proteolysis of the short peptide, and mNeonGreen was linked to FP$_{1-10}$ to monitor its expression, and for normalization purposes. Each construct was injected into wild-type animals and fluorescent progeny were analyzed. Unexpectedly, split-sfCherry3 turned out to be toxic when expressed ubiquitously, whereas 99% of split-wrmScarlet-overexpressing worms became viable adults (Supplementary Figure S3B).

In an attempt to reduce split-sfCherry3-associated toxicity, we modified our construct by using the muscle-specific *myo-3* promoter and removing the NLS sequence (Figure 1B). We did not detect toxicity associated with the expression of these constructs and were able to compare the fluorescence of split-sfCherry3 and split-wrmScarlet in young adults. Red fluorescence emitted from split-wrmScarlet was 2.9-fold higher than that of split-sfCherry3 when normalized to the mNeonGreen signal (Figure 1, B and C and Supplementary Figure S4). We also observed a 60% higher expression level of mTagBFP2 in the split-wrmScarlet-expressing animals (Supplementary Figure S4). It is worth noting that differences in expression levels could influence both brightness and toxicity comparisons. A more controlled way to compare the split FPs at similar expression levels would be to make single-copy genomic insertions of these constructs at a neutral site in the genome.

### Split-wrmScarlet$_{11}$-mediated tagging in all somatic tissues or specifically in muscles

Our protein-tagging approach was analogous to existing split-FP methods developed for human cells (Kamiyama *et al.* 2016, Leonetti *et al.* 2016) and *C. elegans* (Hefel and Smolikove 2019). It requires split-wrmScarlet$_{1-10}$ (*i.e.*, just the large fragment of split-wrmScarlet without the 11th β-strand) to be expressed in the cell or tissue of interest, and the small split-wrmScarlet$_{11}$ fragment to be inserted at an endogenous locus to tag a protein of interest (Figure 1A).
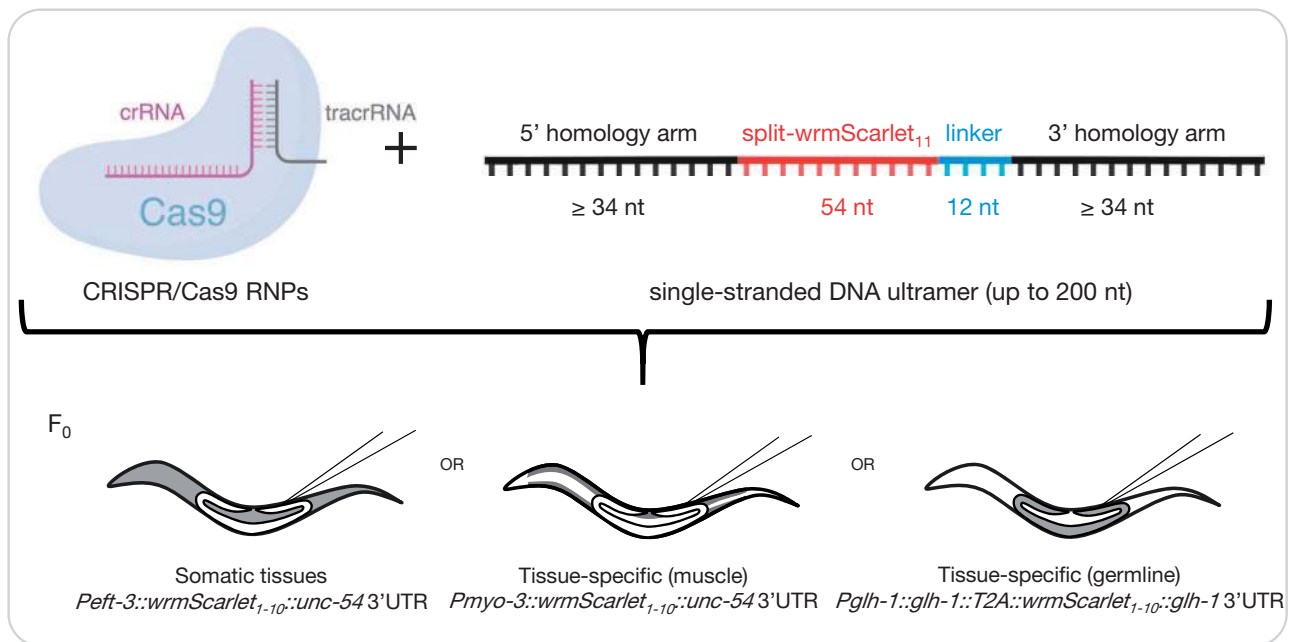
To build strains expressing single-copy insertions of split-wrmScarlet$_{1-10}$, we first optimized its sequence for *C. elegans* codon usage (Redemann *et al.* 2011) and included three introns (Supplementary Table S1). The strain expressing split-wrmScarlet$_{1-10}$ throughout the soma (driven by the *eft-3* promoter and *unc-54* 3'UTR) was generated by editing the genome of the existing MosSCI line CA1200 (Zhang *et al.* 2015) and replacing the sequence encoding *tir-1::mRuby* with *split-wrmScarlet*$_{1-10}$ using CRISPR/Cas9 and hybrid DNA templates (Paix *et al.* 2015; Dokshin *et al.* 2018; Supplementary Table S4). In order to perform tissue-specific labeling, we generated a strain expressing muscle-specific split-wrmScarlet$_{1-10}$ using the SKI-LODGE system in the strain WBM1126 (Silva-García *et al.* 2019; Supplementary Table S4). The expression of split-wrmScarlet$_{1-10}$ in these two lines did not affect the number of viable progeny (Supplementary Figure S5A) nor lifespan (Supplementary Figure S5B and Table S6), suggesting that the expression of split-wrmScarlet$_{1-10}$ had no deleterious effect. To tag a gene of interest with the split-wrmScarlet$_{11}$

fragment, we used microinjection of preassembled Cas9 ribonucleoproteins, because this method enables high-efficiency genome editing in worms (Paix *et al.* 2015). The most efficient insertion of short sequences in *C. elegans* was previously shown to be achieved using ssODNs (Paix *et al.* 2015; Prior *et al.* 2017; Dokshin *et al.* 2018). A great advantage of this strategy is that all of the components required for editing are commercially available or can be synthesized rapidly in the lab (Leonetti *et al.* 2016). Synthetic ssODNs have a typical size limit of 200 nt. The small size of split-wrmScarlet$_{11}$ (18–24 a.a.) is key: 200 nt can encompass split-wrmScarlet$_{11}$ (66–84 nt, including a 4 a.a. linker) flanked by two homology arms >34 nt (up to 67–58 nt) for HDR. In principle, a few days after the somatic and/or muscle-specific split-wrmScarlet$_{1-10}$ strain(s) are microinjected, progeny can be screened for red fluorescence, genotyped and sequenced to check the accuracy of editing (Figure 2; a detailed protocol is available
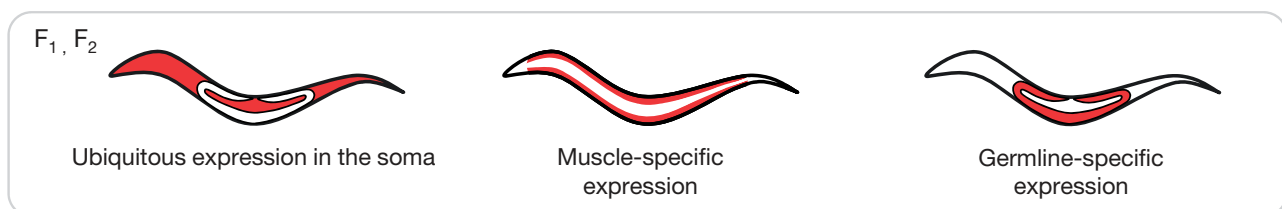
at dx.doi.org/10.17504/protocols.io.bamkic4w). If desired, co-CRISPR strategies such as *dpy-10(cn64)* (Paix *et al.* 2015) or co-injection with pRF4 (Dokshin *et al.* 2018) can be used to screen for correct candidates and to control for microinjection efficacy and payload toxicity.

To test our approach, we used it to tag six proteins with distinct subcellular localizations. Starting with the somatic split-wrmScarlet$_{1-10}$ parental strain CF4582, we introduced split-wrmScarlet$_{11}$ at the N-terminus of TBB-2, FIB-1 or VHA-13 or at the C-terminus of EAT-6, HIS-3, and TOMM-20 (Supplementary Table S4). These proteins mark the cytoskeleton, nucleoli, lysosomes, plasma membrane, nuclei, and mitochondria, respectively. Importantly, for tagging transmembrane proteins, the split-wrmScarlet$_{11}$ tag was introduced at the terminus exposed to the cytosol. Split-wrmScarlet fluorescence from all six proteins matched their expected subcellular localization in somatic cells (Figure 3,



**1. Microinjection into split-wrmScarlet$_{1-10}$ strains of choice**

crRNA tracrRNA
Cas9
CRISPR/Cas9 RNPs

5' homology arm — split-wrmScarlet$_{11}$ linker — 3' homology arm
≥ 34 nt — 54 nt — 12 nt — ≥ 34 nt
single-stranded DNA ultramer (up to 200 nt)

F$_0$

Somatic tissues
*Peft-3::wrmScarlet$_{1-10}$::unc-54* 3'UTR

OR

Tissue-specific (muscle)
*Pmyo-3::wrmScarlet$_{1-10}$::unc-54* 3'UTR

OR

Tissue-specific (germline)
*Pglh-1::glh-1::T2A::wrmScarlet$_{1-10}$::glh-1* 3'UTR

**2. Screen for fluorescent split-wrmScarlet$_{11}$-positive progeny**

F$_1$, F$_2$

Ubiquitous expression in the soma

Muscle-specific expression

Germline-specific expression

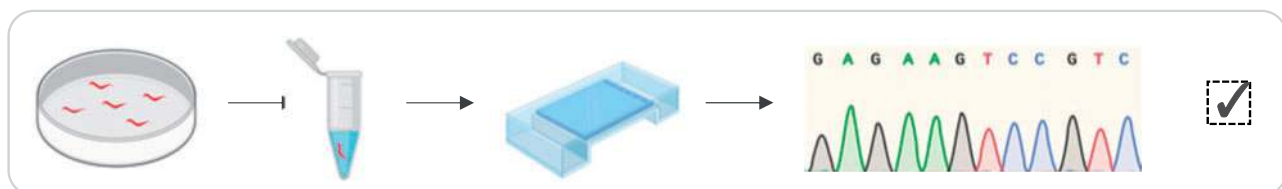**3. Lysis, PCR genotyping and sequencing**

G A G A A G T C C G T C

**Figure 2** Split-wrmScarlet$_{11}$-mediated tagging. Schematic representation of the split-wrmScarlet tagging workflow to visualize endogenous proteins specifically in muscles, germline, or throughout the soma. Some illustrations were created with BioRender.com.
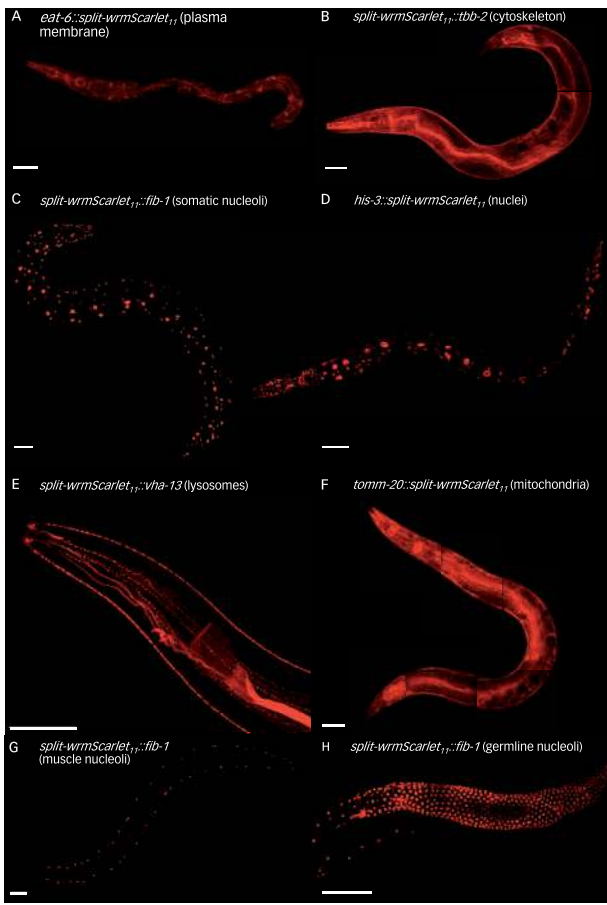
**Figure 3** Tissue-specific split-wrmScarlet labeling of proteins with distinct subcellular locations. Endogenous proteins tagged with split-wrmScarlet$_{11}$ in animals expressing split-wrmScarlet$_{1–10}$ in somatic tissues, in muscles or in the germline. (A–F) Confocal images of worms expressing somatic split-wrmScarlet$_{1–10}$ and (A) EAT-6::split-wrmScarlet$_{11}$ (plasma membrane), (B) split-wrmScarlet$_{11}$::TBB-2 (cytoskeleton), (C) split-wrmScarlet$_{11}$::FIB-1 (nucleoli), (D) HIS-3::split-wrmScarlet$_{11}$ (nuclei), (E) split-wrmScarlet$_{11}$::VHA-13 (lysosomes), or (F) TOMM-20::split-wrmScarlet$_{11}$ (mitochondria). (G) Transgenic worm expressing split-wrmScarlet$_{1–10}$ in muscle and split-wrmScarlet$_{11}$::FIB-1. (H) Transgenic worm expressing split-wrmScarlet$_{1–10}$ in the germline and split-wrmScarlet$_{11}$::FIB-1. (A–H) Maximum intensity projections of 3D stacks shown. Scale bars, 50 μm.

A–F). To test the muscle-specific split-wrmScarlet$_{1–10}$ line CF4610, we tagged the N-terminus of endogenous FIB-1 with split-wrmScarlet$_{11}$ and confirmed the fluorescence from nucleoli in muscle cells (Figure 3G). Together, our results show that split-wrmScarlet enables rapid fluorescent tagging of proteins with disparate cytoplasmic or nuclear locations expressed from their endogenous loci.

The 18 a.a. split-wrmScarlet$_{11}$ sequence used for these experiments ends with two glycines. In mammalian cells, C-terminal gly–gly sequences have been reported to function as degradation signals (Koren *et al.* 2018). Our TOMM-20::split-wrmScarlet$_{11}$ had a spontaneous mutation of the last glycine to a stop codon (Supplementary Table S4), which could be problematic if the protein degradation mechanism, DesCEND (destruction via C-end degrons) operates in *C. elegans*. However, we do not detect differences in protein abundance of HIS-3 versus HIS-3::split-wrmScarlet$_{11}$ by western blots in *C. elegans* (Supplementary Figure S11). We also do not detect differences in protein abundance of mScarlet truncated to end in gly–gly compared to

mScarlet ending in MDELYK via fluorescence in yeast. Nonetheless, we recommend using a 24 a.a. split-wrmScarlet$_{11}$ sequence YTVVEQYEKSVARHCTGGMDELYK when labeling proteins at their C-terminus to avoid the possibility that split-wrmScarlet$_{11}$ ending in gly–gly could function as a degron. This modified sequence still fits within the 200 nt ssODN synthesis limit and works at least as well as the 18 a.a. split-wrmScarlet$_{11}$ sequence (Supplementary Figure S6).

## Split-wrmScarlet$_{11}$-mediated tagging in the germline

Our initial attempt to use split-wrmScarlet in the germline failed. We made a single-copy integrated Psun-1::split-wrmScarlet$_{1–10}$::sun-1 3'UTR strain via MosSCI, but when we injected a plasmid encoding mNeonGreen::split-wrmScarlet$_{11}$, we observed green fluorescence, but no red fluorescence (Supplementary Figure S7, A and B), suggesting the absence of split-wrmScarlet$_{1–10}$ expression. We suspected germline silencing of the germline-expressed split-wrmScarlet$_{1–10}$, so we attempted an alternative expression approach by taking advantage of the germline-helicase protein GLH-1, which is highly expressed and germline-specific (Marnik *et al.* 2019). We fused a T2A::split-wrmScarlet$_{1–10}$ sequence to the C-terminus of the endogenous *glh-1* gene using CRISPR/Cas9. The high expression of GLH-1 yielded high expression of split-wrmScarlet$_{1–10}$, and the T2A separated split-wrmScarlet$_{1–10}$ from GLH-1 (Liu *et al.* 2017). The *glh-1*::T2A::split-wrmScarlet$_{1–10}$ strain (https://cgc.umn.edu/strain/DUP237) can be used like our other tissue-specific strains for germline-specific tagging. To demonstrate this, we tagged the N-terminus of endogenous FIB-1 with split-wrmScarlet$_{11}$, and we observed red fluorescence localized to the nucleoli specifically in the germline and embryos, as we hoped (Figure 3H; Supplementary Figure S8, A and C). Finally, we note that the strategy used to express wrmScarlet$_{1-10}$ or sGFP2$_{1-10}$ in the germline, by tagging the 3' end of the endogenous *glh-1* with T2A::FP$_{1-10}$ with CRISPR/Cas9, could be used to express any other protein of choice.

## Split-wrmScarlet$_{11}$ tandem repeats increase fluorescence

To benchmark the fluorescence intensity of split-wrmScarlet against its full-length counterpart, we first generated endogenously-expressed wrmScarlet::vha-13 (El Mouridi *et al.* 2017) transgenic animals and compared their fluorescence to split-wrmScarlet$_{11}$::vha-13 in worms expressing split-wrmScarlet$_{1–10}$ somatically (Figure 4, A and B). At the vha-13 locus, split-wrmScarlet was about half as bright as a full-length fluorophore (48%), a ratio comparable to that of split-mNeonGreen2 and its full-length counterpart in human cells (Feng *et al.* 2017).

Since visualizing endogenous proteins of low abundance can be challenging, it is key to address this limitation. Increasing the number of FP$_{11}$ domains tagged to an endogenous protein multiplies the number of the corresponding FP$_{1–10}$s recruited, increasing the overall fluorescent signal in human cells (Leonetti *et al.* 2016) and in *C. elegans* (He *et al.* 2019; Hefel and Smolikove 2019). To demonstrate that split-wrmScarlet fluorescence is enhanced by split-wrmScarlet$_{11}$ tandem repeats, we introduced two split-wrmScarlet$_{11}$ domains at the N-terminus of VHA-13 and three split-wrmScarlet$_{11}$ domains at the C-terminus of HIS-3 in animals expressing somatic split-wrmScarlet$_{1–10}$. Compared to animals carrying a single split-wrmScarlet$_{11}$ at the identical locus, carrying two split-wrmScarlet$_{11}$s increased overall fluorescence by 1.5-fold, while carrying three increased it by 2.3-fold (Figure 4, C and D). Note that our three-split-wrmScarlet$_{11}$ tandem
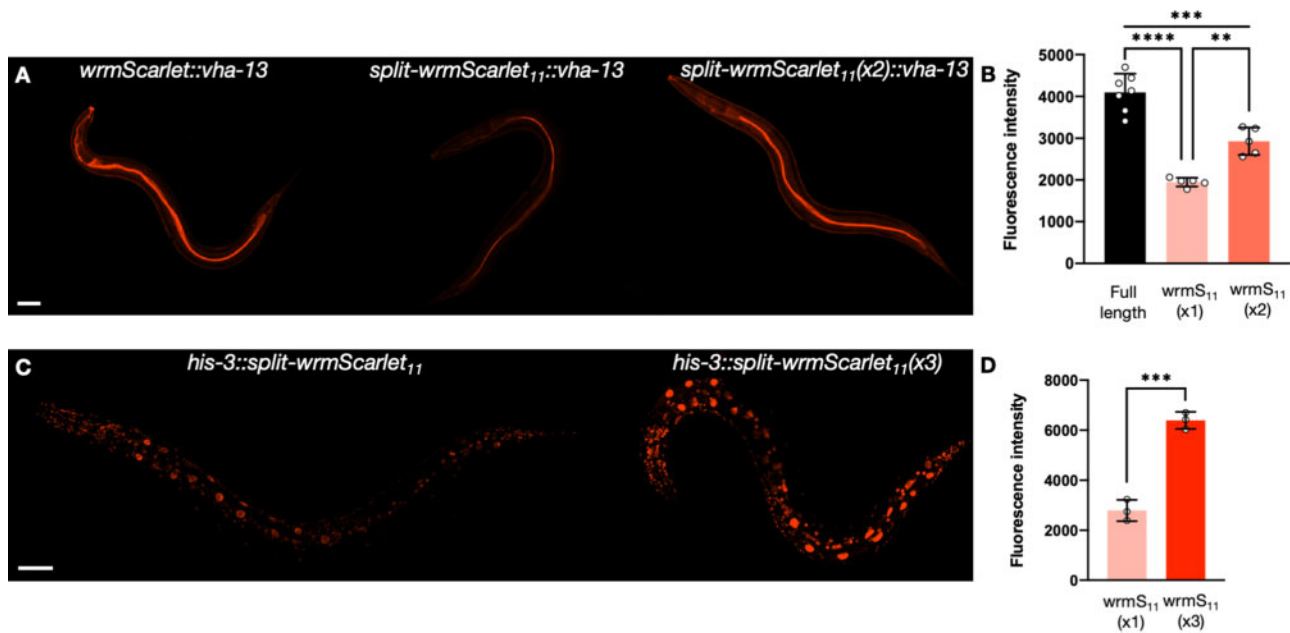
**Figure 4** Split-wrmScarlet$_{11}$ tandem repeats increase fluorescence. (A) Images of animals carrying either wrmScarlet, split-wrmScarlet$_{11}$ or two tandem repeats of split-wrmScarlet$_{11}$ inserted at the endogenous VHA-13 N-terminus. (B) Emission intensities of animals carrying wrmScarlet, split-wrmScarlet$_{11}$ or dual split-wrmScarlet$_{11}$ inserted at the VHA-13 N-terminus. Mean ± SD. Circles are individuals. ****$P < 0.0001$, ***$P < 0.001$, **$P < 0.005$. (C) Images of animals carrying either a single split-wrmScarlet$_{11}$ or three tandem repeats of split-wrmScarlet$_{11}$ inserted at the HIS-3 C-terminus. (D) split-wrmScarlet emission intensities from animals carrying a single split-wrmScarlet$_{11}$ or three tandem repeats of split-wrmScarlet$_{11}$ knock-in at the HIS-3 C-terminus. Mean ± SD. Circles are individuals. ***$P < 0.001$. Images from each comparison were taken under identical instrument conditions using confocal microscopy and are shown using identical brightness and contrast settings. Images shown are from a single confocal plane. Scale bars, 50 μm.

sequence exceeds the 200 nt ssODN synthesis limit, so we used dsDNA donor templates for these constructions (Supplementary Table S4).

## sfGFP$_{11}$-mediated tagging in somatic cells

Split-sfGFP has been used successfully in worms before (Noma *et al.* 2017; He *et al.* 2019; Hefel and Smolikove 2019). However, there is still a need for a strain that ubiquitously expresses sfGFP$_{1–10}$ in the soma from an integrated single-copy insertion in order to avoid heterogeneous expression and time-consuming manual maintenance. To build this strain, we codon-optimized the original *sfGFP$_{1–10}$* sequence for *C. elegans* and included one intron (Cabantous *et al.* 2005; Redemann *et al.* 2011; Supplementary Table S1). We initially generated a strain expressing sfGFP$_{1–10}$ driven by the *let-858* promoter and *unc-54* 3'UTR using MosSCI, but later replaced the *let-858* promoter with the *eft-3* promoter using CRISPR/Cas9 and hybrid DNA donor template because we observed that P*eft-3* resulted in significantly higher levels of gene expression (Paix *et al.* 2015; Dokshin *et al.* 2018; Supplementary Table S4). To validate this strain, we inserted sfGFP$_{11}$ at the N-terminus of lysosomal VHA-13 or at the C-terminus of nuclear-localized HIS-3 (Figure 5, A and B). Both strains yielded relatively bright signals in accordance with their predicted subcellular localization. We generated eGFP::VHA-13 transgenic animals and compared their fluorescence to sfGFP$_{11}$::VHA-13 in worms expressing sfGFP$_{1–10}$ somatically (Supplementary Figure S9, A and B). At the *vha-13* locus, split-sfGFP was about a third as bright as a full-length eGFP. It is worth noting that this comparison is not perfect, in part due to the presence of the six superfolder mutations S30R, Y39N, N105T, Y145F, I171V, and A206V in sfGFP but not in eGFP.

## sGFP2$_{11}$-mediated tagging in the germline

We also generated a germline-specific sGFP2$_{1–10}$ strain using a similar strategy (Supplementary Figure S8B, Supplementary Table S1). Split-sGFP2 is a split-superfolder GFP variant optimized for brightness and photostability (Köker *et al.* 2018). To test this germline-specific line DUP223 (https://cgc.umn.edu/strain/DUP223), we tagged the C-terminus of endogenous PGL-1 with sGFP2$_{11}$ and confirmed the green fluorescence from P-granules (Supplementary Figure S8B, lower panel).

## Dual color protein labeling with split-wrmScarlet and split-sfGFP

Finally, to allow two-color imaging in the soma, we crossed the strains CF4592 P*eft-3::sfGFP$_{1–10}$*; *his-3::sfGFP$_{11}$* and CF4601 P*eft-3::split-wrmScarlet$_{1–10}$*; *split-wrmScarlet$_{11}$::fib-1*. This cross resulted in the line CF4588 P*eft-3::sfGFP$_{1–10}$* P*eft-3::split-wrmScarlet$_{1–10}$* (https://cgc.umn.edu/strain/CF4588) as well as the dually labeled strain CF4602 P*eft-3::sfGFP$_{1–10}$* P*eft-3::split-wrmScarlet$_{1–10}$*; *split-wrmScarlet$_{11}$::fib-1 his-3::sfGFP$_{11}$* (Figure 5C). The fluorescent signals from both split-FPs appeared in their respective subcellular compartments, suggesting the two systems are compatible. We note an additional advantage of the strain CF4588: the loci of *split-wrmScarlet$_{1–10}$* and *sfGFP$_{1–10}$* are genetically linked (only 0.96 cM apart), which facilitates outcrossing when needed. In addition, all our parental *C. elegans* lines expressing *split-wrmScarlet$_{1–10}$* and *sfGFP$_{1–10}$* are viable homozygotes, so the strains do not require special maintenance.

## The current split-wrmScarlet is not detectable in mammalian cells

We failed to detect split-wrmScarlet in mammalian cells, despite our efforts to rescue its fluorescence by screening a mammalian-
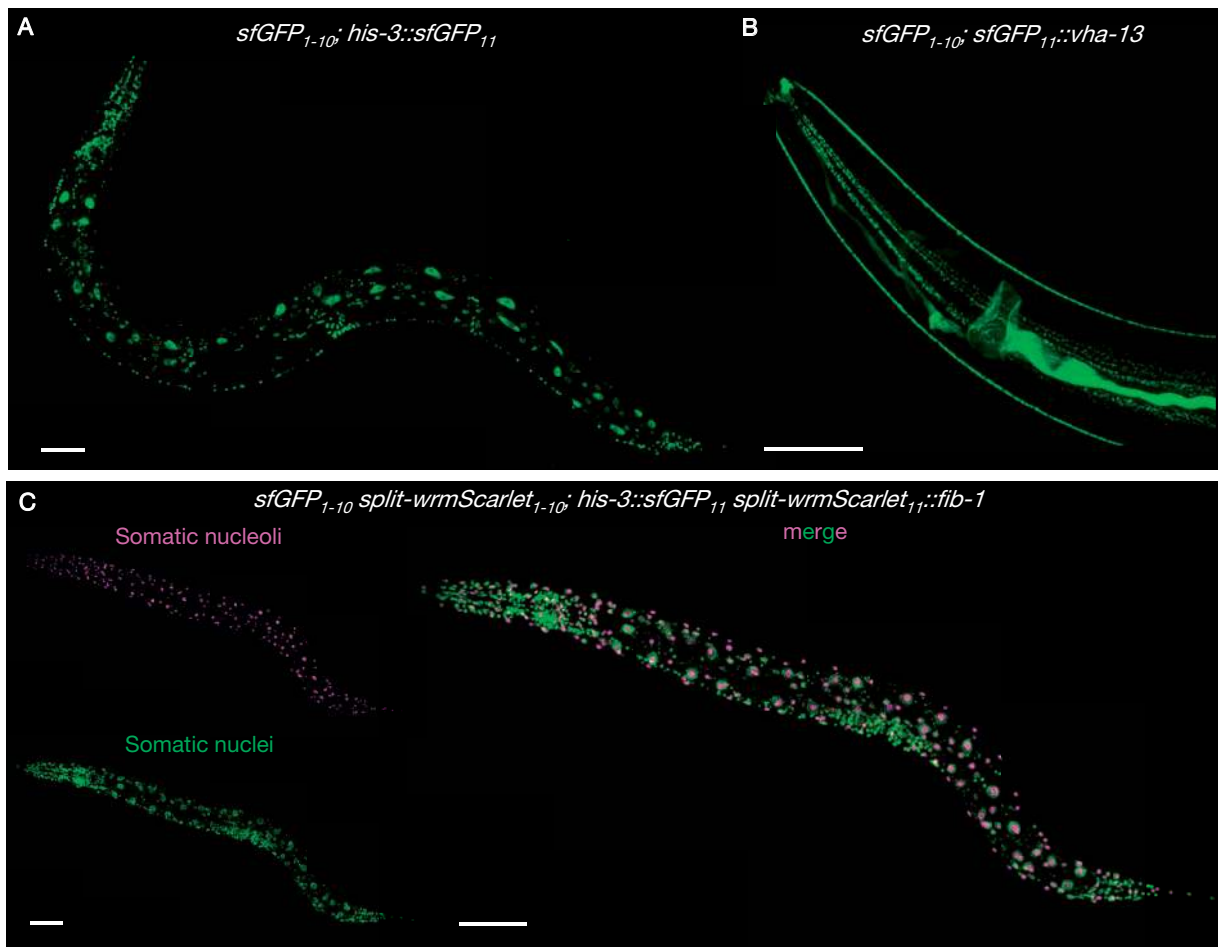
**Figure 5** Split-sfGFP and split-wrmScarlet dual-color protein labeling. Images of animals stably expressing sfGFP$_{1-10}$ in somatic tissues (A) CF4592 *muIs253[Peft-3::sfGFP$_{1-10}$::unc-54 3'UTR Cbr-unc-119(+)] II; unc-119(ed3) III; his-3(muIs255[his-3::sfGFP$_{11}$])* V or (B) CF4589 *muIs253[Peft-3::sfGFP$_{1-10}$::unc-54 3'UTR Cbr-unc-119(+)] II; unc-119(ed3) III; vha-13(muIs268[sfGFP$_{11}$::vha-13])* V. (C) Dual color protein labeling with split-wrmScarlet and split-sfGFP in somatic cells. Composite display of red and green channels of animals expressing split-wrmScarlet$_{1-10}$ and sfGFP$_{1-10}$ in somatic tissues, HIS-3::sfGFP$_{11}$ and split-wrmScarlet$_{11}$::FIB-1; CF4602 *muIs253[Peft-3::sfGFP$_{1-10}$::unc-54 3'UTR Cbr-unc-119(+)] muIs252[Peft-3::split-wrmScarlet$_{1-10}$::unc-54 3'UTR Cbr-unc-119(+)] II; unc-119(ed3) III; fib-1(muIs254[split-wrmScarlet$_{11}$::fib-1]) his-3(muIs255[his-3::sfGFP$_{11}$])* V. Maximum intensity projections of 3D stacks shown. Scale bars, 50 μm.

codon-optimized split-wrmScarlet$_{11}$ single/double mutant library in HEK293T cells (Supplementary Figure S10, A and B, and Supplementary text).

## Discussion

In this study, we describe several new tools for rapid CRISPR-mediated labeling of endogenously expressed proteins using split fluorophores. While these tools are powerful and relatively easy to implement, several considerations should be taken into account when using this method.

First, detection of a given protein labeled with an FP$_{11}$ can only occur in a cellular compartment where the corresponding FP$_{1-10}$ is present. Proteins tagged with split-wrmScarlet$_{11}$ or sfGFP$_{11}$ generated in this work were either exposed to the cytosol or nucleoplasm (nuclei or nucleoli), where split-wrmScarlet$_{1-10}$ and/or sfGFP$_{1-10}$ were present. For proteins or epitopes located within the lumen of organelles, such as mitochondria or the endoplasmic reticulum, one might need to generate and validate *C. elegans* lines expressing split-wrmScarlet$_{1-10}$ or sfGFP$_{1-10}$ containing a mitochondrial localization sequence or ER signal peptide and retention signals, respectively. These approaches have been used

successfully in mammalian cells with split-sfGFP when tagging ER-resident polypeptides (Kamiyama *et al.* 2016) and with split-sfCherry2 to detect proteins present in the mitochondrial matrix (Ramadani-Muja *et al.* 2019).

Second, when labeling proteins with split-wrmScarlet at the C-terminus, we recommend using the 24 a.a. split-wrmScarlet$_{11}$ sequence YTVVEQYEKSVARHCTGGMDELYK. As described in the Results section, our 18 a.a. split-wrmScarlet$_{11}$ fragment ends in gly–gly, which has been shown to be a degradation signal in mammalian cells. We cannot exclude the possibility that ending in gly–gly can be detrimental in *C. elegans*. The 24 a.a. split-wrmScarlet$_{11}$ still fits within a 200 nt ssODN donor template with a 12 nt linker and up to 58 nt homology arms and is at least as bright as the 18 a.a. split-wrmScarlet$_{11}$ (Supplementary Figure S6).

Third, as for any other protein tag, it is important to select, when possible, a site that is unlikely to interfere with protein folding, function or localization (Snapp 2005; Nance and Frøkjær-Jensen 2019). For example, N-termini of membrane- and organelle-resident proteins often contain signal peptides or localization signals, and C-termini may contain degron sequences that regulate protein turnover. Interestingly, there are examples

of proteins that become toxic when tagged with a full-length GFP, but tolerate labeling with a split fluorescent protein. For example, SYP-4 was reported to be mostly functional when endogenously tagged with sfGFP$_{11}$ in a strain expressing sfGFP$_{1-10}$ specifically in the germline, but not functional when labeled with full-length GFP (Hefel and Smolikove 2019).

Fourth, for proteins of interest present at low levels, we provided an alternative protocol to insert an additional two or three split-wrmScarlet$_{11}$ fragments, which increases the overall fluorescence substantially. However, the number of split-wrmScarlet$_{11}$ fragments could likely be increased further, to at least seven tandem repeats, based on approaches used successfully with split-sfGFP in human cells (Feng *et al.* 2017) and *C. elegans* (Noma *et al.* 2017; He *et al.* 2019; Hefel and Smolikove 2019).

Fifth, we would like to emphasize differences between our technique and the bimolecular fluorescence-complementation (BiFC) assay. When used together, high-affinity green and red split fluorescent proteins can provide information on co-localization, but unlike BiFC split proteins (Hu *et al.* 2002), they are not intended to assess protein-protein interactions directly. This is because BiFC split proteins require finely tuned affinities that do not disrupt the underlying interaction being studied. In our approach, only the split-wrmScarlet$_{11}$ fragment is attached to a protein of interest, the split-wrmScarlet$_{1-10}$ fragment is expressed in excess and unattached.

Finally, we would like to note that despite being three times brighter than the latest split-sfCherry3 in worms, our current split-wrmScarlet was not visible in the mammalian cell line we examined (Supplementary Figure S10). Its ability to fluoresce is not restricted to worms, because it can reach wild-type levels of brightness in yeast. We do not know the basis for this discrepancy. It is possible that the concentration of the split-wrmScarlet$_{1-10}$ fragment in mammalian cells is too low to drive complementation with split-wrmScarlet$_{11}$. This could potentially be overcome by further mutagenizing split-wrmScarlet and screening for fluorescence at low expression levels in mammalian cells.

In conclusion, we believe our system can substantially increase the speed, efficiency, and ease of *in vivo* microscopy studies in *C. elegans*. We expect it to facilitate two-color and co-localization experiments and to find wide use in the worm community. We believe that these strains could facilitate novel or large-scale experiments, such as efforts to tag the entire genome of *C. elegans*.

## Conflicts of interest

None declared.

## Literature cited

Bindels DS, Haarbosch L, van Weeren L, Postma M, Wiese KE, *et al.* 2017. mScarlet: a bright monomeric red fluorescent protein for cellular imaging. Nat Methods 14:53–56.

Brenner S. 1974. The genetics of *Caenorhabditis elegans*. Genetics. 77: 71–94.

Cabantous S, Terwilliger TC, Waldo GS. 2005. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. Nat Biotechnol. 23:102–107.

Dickinson DJ, Goldstein B. 2016. CRISPR-Based Methods for *Caenorhabditis elegans* Genome Engineering. Genetics. 202: 885–901.

Dokshin GA, Ghanta KS, Piscopo KM, Mello CC. 2018. Robust genome editing with short single-stranded and long, partially single-stranded DNA donors in *Caenorhabditis elegans*. Genetics. 210:781–787.

El Mouridi S, Lecroisey C, Tardy P, Mercier M, Leclercq-Blondel A, *et al.* 2017. Reliable CRISPR/CAS9 genome engineering in *Caenorhabditis elegans* using a single efficient sgRNA and an easily recognizable phenotype. G3 (Bethesda). 7:1429–1437.

Evans, TC, editor. 2006. Transformation and microinjection In: The C. elegans Research Community. WormBook. doi: 10.1895/wormbook.1.108.1, http://www.wormbook.org.

Farboud B, Severson AF, Meyer BJ. 2019. Strategies for efficient genome editing using CRISPR-Cas9. Genetics. 211:431–457.

Feng S, Sekine S, Pessino V, Li H, Leonetti MD, *et al.* 2017. Improved split fluorescent proteins for endogenous protein labeling. Nat Commun. 8:370.

Feng S, Varshney A, Coto Villa D, Modavi C, Kohler J, *et al.* 2019. Bright split red fluorescent proteins for the visualization of endogenous proteins and synapses. Commun Biol. 2:344–312.

Friedland AE, Tzur YB, Esvelt KM, Colaiácovo MP, Church GM, *et al.* 2013. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. Nat Methods 10:741–743.

Frøkjær-Jensen C, Davis MW, Ailion M, Jorgensen EM. 2012. Improved Mos1-mediated transgenesis in *C. elegans*. Nat Methods. 9:117–118.

Frøkjær-Jensen C, Wayne Davis M, Hopkins CE, Newman BJ, Thummel JM, *et al.* 2008. Single-copy insertion of transgenes in *Caenorhabditis elegans*. Nat Genet. 40:1375–1383.

Goudeau J, Bellemin S, Toselli-Mollereau E, Shamalnasab M, Chen Y, *et al.* 2011. Fatty acid desaturation links germ cell loss to longevity through NHR-80/HNF4 in *C. elegans*. PLoS Biol. 9:e1000599.

He S, Cuentas-Condori A, Miller DM. 2019. NATF (Native and Tissue-Specific Fluorescence): a strategy for bright,

tissue-specific GFP labeling of native proteins in *Caenorhabditis elegans*. Genetics. 212:387–395.

Hefel A, Smolikove S. 2019. Tissue-Specific Split sfGFP system for streamlined expression of GFP tagged proteins in the *Caenorhabditis elegans* Germline. G3(Bethesda). 9:1933–1943.

Hu CD, Chinenov Y, Kerppola TK. 2002. Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. Mol Cell. 9:789–798.

Kamiyama D, Sekine S, Barsi-Rhyne B, Hu J, Chen B, *et al.* 2016. Versatile protein tagging in cells with split fluorescent protein. Nat Commun. 7:11046–11049.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 10:845–858.

Kelly WG, Xu S, Montgomery MK, Fire A. 1997. Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* Gene. Genetics. 146:227–238.

Kimble J, Hodgkin J, Smith T, Smith J. 1982. Suppression of an amber mutation by microinjection of suppressor tRNA in *C. elegans*. Nature. 299:456–458.

Köker T, Fernandez A, Pinaud F. 2018. Characterization of split fluorescent protein variants and quantitative analyses of their self-assembly process. Sci Rep. 8:5344.

Koren I, Timms RT, Kula T, Xu Q, Li MZ, *et al.* 2018. The eukaryotic proteome is shaped by E3 ubiquitin ligases targeting C-terminal degrons. Cell. 173:1622–1635.e14.

Liu Z, Chen O, Wall J, Zheng M, Zhou Y, *et al.* 2017. Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. Sci Rep. 7:2193.

Leonetti MD, Sekine S, Kamiyama D, Weissman JS, Huang B. 2016. A scalable strategy for high-throughput GFP tagging of endogenous human proteins. Proc Natl Acad Sci USA. 113:E3501–E3508.

Marnik EA, Fuqua JH, Sharp CS, Rochester JD, Xu EL. 2019. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J. 213:923–939.

Mello CC, Kramer JM, Stinchcomb D, Ambros V. 1991. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J. 10:3959–3970.

Nance J, Frøkjær-Jensen C. 2019. The *Caenorhabditis elegans* transgenic toolbox. Genetics. 212:959–990.

Noma K, Goncharov A, Ellisman MH, Jin Y. 2017. Microtubule-dependent ribosome localization in *C. elegans* neurons. eLife. 6:218.

Paix A, Folkmann A, Rasoloson D, Seydoux G. 2015. High efficiency, homology-directed genome editing in *Caenorhabditis elegans* using CRISPR-Cas9 ribonucleoprotein complexes. Genetics. 201:47–54. doi:10.1534/genetics.115.179382/-/DC1.

Paix A, Schmidt H, Seydoux G. 2016. Cas9-assisted recombineering in *C. elegans*: genome editing using *in vivo* assembly of linear DNAs. Nucleic Acids Res. 44:e128. [10.1093/nar/gkw/502]

Paix A, Wang Y, Smith HE, Lee C-YS, Calidas D, *et al.* 2014. Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 Sites in *Caenorhabditis elegans*. Genetics. 198: 1347–1356.

Pédelacq J-D, Cabantous S, Tran T, Terwilliger TC, Waldo GS. 2006. Engineering and characterization of a superfolder green fluorescent protein. Nat Biotechnol. 24:79–88.

Prior H, Jawad AK, MacConnachie L, Beg AA. 2017. Highly efficient, rapid and co-CRISPR-independent genome editing in *Caenorhabditis elegans*. G3 (Bethesda). 7:3693–3698.

Ramadani-Muja J, Gottschalk B, Pfeil K, Burgstaller S, Rauter T, *et al.* 2019. Visualization of Sirtuin 4 distribution between mitochondria and the nucleus, based on bimolecular fluorescence self-complementation. Cells. 8:1583.

Redemann S, Schloissnig S, Ernst S, Pozniakowsky A, Ayloo S, *et al.* 2011. Codon adaptation-based control of protein expression in *C. elegans*. Nat Methods. 8:250–252.

Richardson CD, Kazane KR, Feng SJ, Zelin E, Bray NL, *et al.* 2018. CRISPR-Cas9 genome editing in human cells occurs via the Fanconi anemia pathway. Nat Genet. 50:1132–1139.

Silva-García CG, *et al.* 2019. Single-copy knock-in loci for defined gene expression in *Caenorhabditis elegans*. G3 (Bethesda). 9: 2195–2198.

Snapp E. 2005. Design and use of fluorescent fusion proteins in cell biology. Curr Protoc Cell Biol. 27:21.4.1–21.4.13.

Vicencio J, Martínez-Fernández C, Serrat X, Cerón J. 2019. Efficient generation of endogenous fluorescent reporters by nested CRISPR in *Caenorhabditis elegans*. Genetics. 211:1143–1154.

Wu W-S, Huang W-C, Brown JS, Zhang D, Song X, *et al.* 2018. pirScan: a webserver to predict piRNA targeting sites and to avoid transgene silencing in *C. elegans*. Nucleic Acids Res. 46:W43–W48.

Zhang L, Ward JD, Cheng Z, Dernburg AF. 2015. The auxin-inducible degradation (AID) system enables versatile conditional protein depletion in *C. elegans*. Development. 142:4374–4384.

Zhang D, Tu S, Stubna M, Wu W-S, Huang W-C, *et al.* 2018. The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. Science. 359:587–592.

*Communicating editor: H. Buelow*