Bright split red fluorescent proteins with enhanced complementation efficiency for the tagging of endogenous proteins and visualization of synapses

Siyu Feng[1], Aruna Varshney[2], Doris Coto Villa[2], Cyrus Modavi[3], John Kohler[4], Fatima Farah[2], Nebat Ali[2], Joachim Dieter Mueller[4], Miri VanHoven[2], Bo Huang[5,6,7,*]

[1] The UC Berkeley-UCSF Graduate Program in Bioengineering, San Francisco, CA 94143 USA;
[2] Department of Biological Sciences, San Jose State University, San Jose, CA 95192 USA;
[3] Department of Bioengineering and Therapeutic Sciences, University of California in San Francisco, San Francisco, CA 94143, USA;
[4] School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA;
[5] Department of Pharmaceutical Chemistry, University of California in San Francisco, San Francisco, CA 94143, USA;
[6] Department Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94143, USA;
[7] Chan Zuckerberg Biohub, San Francisco, CA 94158, USA.

* Correspondence should be addressed to bo.huang@ucsf.edu

Abstract

Self-associating split fluorescent proteins (FPs) have been widely used for labeling proteins, scaffolding protein assembly and detecting cell-cell contacts. Newly developed self-associating split FPs, however, have suffered from suboptimal fluorescence signal. Here, by investigating the complementation process, we have demonstrated two approaches to improve split FPs: assistance through SpyTag/SpyCatcher interaction and directed evolution. The latter has yielded two split sfCherry3 variants with substantially enhanced overall brightness, facilitating the tagging of endogenous proteins by gene editing. Based on sfCherry3, we have further developed a new red-colored trans-synaptic marker called Neuroligin-1 sfCherry3 Linker Across Synaptic Partners (NLG-1 CLASP) for multiplexed visualization of neuronal synapses in living animals, demonstrating its broad applications.

Introduction

Self-associating split fluorescent proteins (FPs) are a powerful tool for protein labeling and live-cell imaging. In this system, the eleventh β-strand of FP (FP_{11}, 16 amino acids) is separated out from the remainder of FP ($FP_{1\text{-}10}$) and is genetically fused to the protein of interest (POI). Specific fluorescence signal is detected when $FP_{1\text{-}10}$ reconstitutes with the on-target FP_{11} to generate a functional fluorescent protein. Since the initial development of self-associating split $GFP_{1\text{-}10/11}$ (*1*), this technology has been modified and adapted for an extensive range of applications including protein labeling and visualization (*2-4*), scaffolding protein assembly (*3*), protein solubility and aggregation assays (*5, 6*), and monitoring the membrane fusion process (*7*). One prominent application is generating a library of human cells with fluorescently tagged endogenous proteins via CRISPR/Cas9-mediated homology-directed repair (*4*). The small size of the GFP_{11} tag markedly improves the knock-in efficiency and simplifies the donor DNA preparation. In another application, the split $GFP_{1\text{-}10/11}$ system has also been utilized to visualize synapses in living nervous systems by Neuroligin-1 GFP Reconstitution Across Synaptic Partners (NLG-1 GRASP) (*8*).

While all these applications have focused on a single, green-colored channel, expanding the color palette will greatly benefit the investigation of more complex biological systems by enabling multicolor imaging. For this purpose, we have recently developed a red-colored split $sfCherry_{1\text{-}10/11}$ (*3*) and then a brighter split $sfCherry2_{1\text{-}10/11}$ (*9*), enabling dual-color endogenous labeling in human cells using orthogonal FP_{11} tags (*9*) and visualization of *Listeria* protein secretion in infection (*10*). However, unlike split $GFP_{1\text{-}10/11}$, which is as bright as its full-length, split $sfCherry2_{1\text{–}10/11}$ produces substantially lower overall fluorescence signal than its full-length counterpart.

Here, we have characterized the complementation mechanism of split FP systems by examining their overall and single-molecule fluorescence brightness. The results suggest a 2-step complementation model in which the affinity between the $FP_{1\text{-}10}$ and FP_{11} fragment is the major limitation to the overall fluorescence signal. Based on this model, we have devised a SpyTag/SpyCatcher-assisted approach to improve the complementation efficiency of $sfCherry2_{1\text{-}10/11}$. Furthermore, we have engineered two split sfCherry3 variants with much-enhanced complementation efficiency through a combination of cycles of directed evolution and structure-based site-directed mutagenesis. For tagging endogenous proteins by gene editing, sfCherry3 improves the sorting efficiency for successfully knocked-in cells by 5-10 fold in six tested targets, as compared to sfCherry2. Moreover, we have also developed of a new red-colored trans-synaptic marker called Neuroligin-1 sfCherry3 Linker Across Synaptic Partners (NLG-1 CLASP). We established that like NLG-1 GRASP, NLG-1 CLASP labels connections between correct synaptic partners, and has a spatial pattern similar to that predicted by electron micrograph reconstruction (*11, 12*). As a validation, NLG-1 CLASP labeling is disrupted by loss of the *clr-1* gene, which is required for synaptic partner recognition (*13*).

Results

1. The complementation efficiency of split sfCherry2 and split mNeonGreen2 can be explained by a dynamic association/dissociation equilibrium model.

Previously, we have shown that split GFP$_{1\text{-}10/11}$ has nearly identical overall brightness as its full-length counterpart, whereas both split mNeonGreen2$_{1\text{-}10/11}$ (mNG2$_{1\text{-}10/11}$) and split sfCherry2$_{1\text{-}10/11}$ are substantially dimmer (*9*). This sub-optimal performance of mNG2$_{1\text{-}10/11}$ and sfCherry2$_{1\text{-}10/11}$ could be attributed to either (a) the lower molecular brightness of complemented split FPs or (b) incomplete complementation between the FP$_{1\text{-}10}$ and FP$_{11}$ fragments. To test the first possibility, we measured the single-molecule brightness of these three split FPs and their full-length counterparts in living cells using fluorescence fluctuation spectroscopy (*14*). We observed no significant difference in single-molecule brightness between the split and the full-length FPs in all three cases (Fig. 1A). Therefore, incomplete complementation should be the cause of the reduced overall fluorescence signal. The ratio of overall fluorescence between split and full-length FPs then reflects the complementation efficiency.

For mNG2$_{1\text{-}10/11}$ and sfCherry2$_{1\text{-}10/11}$, to determine their complementation efficiency and compare them to GFP$_{1\text{-}10/11}$, we took a similar approach as previously done (*9*). We transiently expressed in HEK 293T cells the full-length FP or the two fragments: FP$_{1\text{-}10}$ and FP$_{11}$ on a well-folded carrier protein. We quantified whole cell fluorescence by flow cytometry while using a co-expressed infrared fluorescent protein mIFP to measure the expression level. mIFP was linked to the full-length FP or the FP$_{11}$ fragment through a P2A self-cleavage site to ensure equimolar expression. Fluorescence intensities in infrared and green/red channels of each single cell events were displayed in log-log scale scatter plots (Figs. 1B-D).

In the cases of full-length FPs, as expected, single cell fluorescence intensities followed the trend line with a slope of 1 because mIFP was expressed equimolarly (Fig. 1E). In the split cases, split GFP$_{1\text{-}10/11}$ almost completely followed the same diagonal line as its full-length counterpart (except for a subtle deviation at the lower-expression end) (Fig. 1B), suggesting a complementation efficiency of almost 100% across a wide range of expression levels. In contrast, both split mNG2$_{1-10/11}$ and split sfCherry2$_{1-10/11}$ deviated from the trend lines of their full length counterparts (Figs. 1C & D). This observation prompted us to consider a 2-step complementation process for self-associating split FPs: the two fragments undergo a reversible association/dissociation equilibrium before entering an irreversible process (fig. S1) of folding and/or chromophore maturation:

$$\mathrm{FP}_{1-10} + \mathrm{FP}_{11} \rightleftharpoons \mathrm{FP}^{*}_{1-10/11} \xrightarrow{\text{folding/maturation}} \mathrm{FP}_{1-10/11} \rightarrow \text{degradation}$$

We have verified that in our co-transfection scheme, the expression levels (concentrations) of FP$_{1-10}$ and FP$_{11}$ fragments are proportional through an independent experiment in cells expressing mIFP-P2A-sfCherry2$_{11}$-Carrier and TagBFP-P2A-sfCherry2$_{1-10}$ (Fig. 1E). Then, at a steady state, this 2-step model predicts the following relationship between the two channels of flow cytometry (see Supplementary Notes):

$$[\mathrm{mIFP}] \propto \sqrt{K_D{}'[\mathrm{FP}_{1-10/11}]} + [\mathrm{FP}_{1-10/11}]$$

where $[\mathrm{mIFP}]$ represents the sum of concentrations of uncomplemented and complemented FP$_{11}$, $[\mathrm{FP}_{1-10/11}]$ is the concentration of matured FP$_{1-10/11}$, and $K_D{}'$ is the effective dissociation constant of the overall complementation/maturation process. When $K_D{}'$ is much lower than the expression level of FP$_{11}$ (the case of GFP$_{1-10/11}$), the second term dominates, leading to a near proportional relationship between the two flow cytometry channels (slope of 1 in log-log plot, Fig. 1B). In the opposite case of high $K_D{}'$ (the cases of mNG2$_{1-10/11}$ and sfCherry2$_{1-10/11}$), the first term dominates, resulting in log-log scatter plots following more closely to slope-2 trend lines, matching our observations (Figs. 1C&D).

2. SpyTag/SpyCatcher interaction improves complementation of split sfCherry2.

Our model indicates that the complementation efficiency of split mNG2$_{1-10/11}$ and sfCherry2$_{1-10/11}$ improves with raised local concentration of fragments, leading to enhanced overall fluorescence signal. Therefore, we sought to utilize a pair of high-affinity binding partners to bring the two fragments into spatial proximity. Because a major advantage of the split FP$_{1-10/11}$ is to label endogenous proteins through knocking-in the short FP$_{11}$ peptide, it is preferable to have a small binding partner for the FP$_{11}$ fragment. For this purpose, we chose the SpyTag/SpyCatcher (*15*) system, a peptide-protein pair that undergoes irreversible binding through formation of an isopeptide bond. The 13-amino-acid (aa) SpyTag is sufficiently short that even when concatenated with GFP$_{11}$, the resulting sequence remains small enough for knock-in using synthetic oligo donor DNAs (*16*).

We examined SpyTag/SpyCatcher-assisted complementation on sfCherry2$_{1\text{-}10/11}$ using the recently improved Spy002 pair (*17*) (Fig. 2A Upper panel). We fused SpyCatcher to the N-terminus of sfCherry2$_{1\text{-}10}$ (the C-terminus is the split site) through a flexible linker in either 6 aa or 15 aa length (Fig. 2A Lower panel). We generated concatenated tags with SpyTag on either the N- or C-terminus of sfCherry2$_{11}$ with a double-glycine linker. We performed similar flow cytometry experiments as described earlier except that mIFP was replaced by TagBFP. Among the four possible combinations of binders/tags (fig. S2), SpyCatcher-6aa-sfCherry2$_{1\text{-}10}$ with SpyTag-sfCherry2$_{11}$-TagBFP demonstrated the most pronounced shift towards a trend line with a slope of 1 in the scatter plot (Fig. 2B). We have further verified that a shorter SpyTag-sfCherry2$_{11}$ fusion without the double-gylcine spacer behaved as efficiently (data not shown).

We validated the improvement in overall brightness for cellular microscopy by labeling the N-terminus of histone 2B (H2B) with SpyTag-sfCherry2$_{11}$-TagBFP and co-expressing it with either sfCherry2$_{1\text{-}10}$ or SpyCatcher-6aa-sfCherry2$_{1\text{-}10}$ fusion in HEK 293T cells. We observed the Spy-assisted system (Fig. 2C) could mark the nuclei with much stronger fluorescence signal with the same expression vectors.

3. Engineering split sfCherry3 for better complementation efficiency

Previously, we engineered sfCherry2$_{1\text{-}10/11}$ using a spacer-insertion strategy (*9*). This strategy was based on inserting a 32-aa spacer between the sfCherry$_{1\text{-}10}$ and sfCherry$_{11}$ coding regions. Beyond allowing mutagenesis of both fragments in a single PCR amplicon, the spatial constraints imposed by the linker were hypothesized to assist in the detection of the original mutations by raising the local concentrations of complimentary fragments. To increase the screening stringency for complementation-enhancing mutations, we chose to express the fragments of sfCherry2 as separate proteins using a pETDuet vector (Fig. 3A). Considering short peptides are prone to degradation in *E. coli*, we fused the sfCherry2$_{11}$ sequence to the N-terminus of a well-folding carrier protein (SpyCatcher in our case).

We subjected the sfCherry2$_{1\text{-}10}$ fragment to four rounds of error-prone PCR mutagenesis and screening. In every round, a mixture of ~20 brightest variants were selected for the next round.

The final isolated mutants were then subjected to one round of DNA shuffling. We have not mutated sfCherry2$_{11}$ so that all variants still bind the identical sfCherry2$_{11}$ peptide. In the end, sfCherry3C with 5 substitutions, K45R, G52D, T106A, K182R, N194D (numbering starts from the 1st Glu after the starting codon Met) was identified as the best variant after the directed evolution (Fig. 3B). All mutations were mapped to either surface orientated residues (T106A, K182R) or locations potentially interacting with the sfCherry2$_{11}$ peptide (K45R, G52D, and N194D).

To further improve the complementation efficiency of sfCherry3C, we introduced rational mutations inspired by a mCherry mutant named cp193g7 that is tolerant of circular permutations near our split-site (*18*), because this mutant contains multiple similar mutations as in sfCherry (*19*) and sfCherry2 (*9*). We combinatorially introduced the remaining mutations of cp193g7 (I7F, F65L, and L83W) into sfCherry3C through site-directed mutagenesis. Only the variant containing a single L83W mutation gave brighter signal than sfCherry3C, which we designated as sfCherry3V (Fig. 3C). Complemented sfCherry3C$_{1\text{-}10/11}$ has an identical emission spectrum as that of sfCherry2 (both split and full-length), whereas the emission spectrum of complemented sfCherry3V$_{1\text{-}10/11}$ is blue-shifted by 5 nm (Fig. 3D). Fluorescence fluctuation spectroscopy indicates that sfCherry3C$_{1\text{-}10/11}$ has the same single-molecule brightness as sfCherry2 (fig. S3). On the other hand, sfCherry3V$_{1\text{-}10/11}$ is dimmer at the single-molecule level (fig. S3), which might be attributed to a difference in two-photon excitation of the blue-shifted chromophore at 1000 nm in the fluorescence fluctuation spectroscopy measurement.

Next, we performed the flow cytometry analysis (Fig. 3E) on HEK 293T cells co-transfected with mIFP-P2A-sfCherry2$_{11}$-Carrier and sfCherryX$_{1\text{-}10}$ (X being 2, 3C or 3V). By fitting the data points (excluding those below a threshold above the scattering background) to a line with a fixed slope of 2 (fig. S4, see Supplementary Notes), we found a substantial up-shift of the fitted lines from sfCherry2 to sfCherry3C and sfCherry3V. The shifts corresponded to increases in binding affinity by 2.5-fold and 8.2-fold, respectively, assuming the same single-molecule brightness under one-photon excitation for flow cytometry. This enhancement in binding affinity should lead to improved complementation efficiency (hence the overall brightness) at the same expression level of the FP$_{1\text{-}10}$ fragment.

4. Efficient endogenous protein labeling in human cells using sfCherry3 variants.

One unique application of FP_{11} tag is to generate library-scale fluorescently labelled endogenous proteins through genetic knock-in by homology-directed DNA repair. The small 16-aa size of FP_{11} allows us to fit its DNA sequence and short homology arms (~70 nt on either side) into commercially available 200 nt single-strand oligo-DNA (ssDNA). By electroporating Cas9/sgRNA ribonucleoprotein (RNP) and donor ssDNA into cells constitutively expressing the corresponding $FP_{1\text{-}10}$ fragment, robust generation of FP-labeled human cell lines becomes fast and cost-effective (*4*). Multicolor knock-in has also been demonstrated by using orthogonal split FP systems to visualize differential distribution and interaction of multiple endoplasmic reticulum proteins (*9*).

The overall increased brightness of complemented sfCherry3 variants make them superior to sfCherry2 in the application of labeling endogenous proteins through knock-in. Because sfCherry3 and sfCherry2 share the same FP_{11} fragment, we adopted a reversed strategy as our previously reported one: knock-in of sfCherry2_{11} into HEK 293T wild-type cells through electroporation, followed by lentivirus infection for the three $\text{sfCherry}_{1\text{-}10}$ variants (schematics in Fig. 4A). A total of six sfCherry2_{11} cell lines were created, with knock-ins at: lamin A/C (LMNA, inner nuclear membrane), clathrin light chain A (CLTA), RAB11A, heterochromatin protein 1 β (HP1b), endoplasmic reticulum proteins SEC61b (translocon complex) and ARL6IP1 (tubular ER). We compared Fluorescence-Activated Cell Sorting (FACS) enrichment efficiency for each cell line after infection with lentivirus (fig. 4B to 4G). In all examined targets, the $\text{sfCherry3C}_{1\text{-}10}$ and $\text{sfCherry3V}_{1\text{-}10}$ groups displayed remarkable population enhancement in the red-fluorescence-positive gate compared to $\text{sfCherry2}_{1\text{-}10}$, rendering the sorting process substantially faster. Practically, for targets like CLTA, SEC61b or ARL6IP1, we were able to gate the fluorescent population around the clear peak and have 5-10 fold higher yield of isolating cells with successful knock-in.

We further confirmed our knock-ins were on-target through confocal microscopy imaging (Fig. 4B to 4G). Noticeably, $\text{sfCherry3V}_{1\text{-}10}$ groups have a reduced tendency to show fluorescent puncta from lysosomes (Fig. 4G) which is reported in our previous work (*9*). This makes the $\text{sfCherry3V}_{1\text{-}10}$ the preferred protein fragment when labeling endogenous proteins involved in the endomembrane system.

5. NLG-1 CLASP visualizes specific subsets of synapses in live animals.

To visualize synapses between specific sets of pre- and postsynaptic neurons in live animals, the trans-synaptic marker NLG-1 GRASP was designed using split GFP fragments (*8*). Complementary split GFP$_{1\text{-}10/11}$ fragments were linked via a flexible linker to the transmembrane synaptic protein Neuroligin, which localizes to both pre- and postsynaptic sites in *C. elegans* (*8*). When the two neurons in which the complementary pre- and postsynaptic markers are expressed form synapses, the split GFP fragments come into contact, reconstitute and fluoresce (Fig. 5A). Using NLG-1 GRASP, we discovered that the recognition between two synaptic partners, the PHB sensory neurons and the AVA interneurons, is mediated the secreted ligand UNC-6/Netrin, its canonical receptor UNC-40/Deleted in Colorectal Cancer (*20*), and the receptor protein tyrosine phosphatase (RPTP) CLR-1 (*13*). NLG-1 GRASP has also been adapted to many other systems, indicating that this technology is transferable (*21*). The addition of a red fluorescent trans-synaptic marker would greatly expand this system.

Leveraging the brightened split-sfCherry3 tagging system, we have developed the split sfCherry-based NLG-1 CLASP. The left and right PHB sensory neurons, located in the posterior of *C. elegans*, form the majority of their synapses with the left and right AVA and PVC interneurons (*11, 12*). In this study, we focused on synapses formed between the left and right PHB neurons with the left and right AVA interneurons. To visualize PHB-AVA synapses with this newly developed split sfCherry3C, constructs were generated in which the sequence encoding the large fragment sfCherry3C$_{1\text{-}10}$ was linked to the N-terminus of the *neuroligin-1 (nlg-1)* cDNA after the *nlg-1* signal sequence via a flexible 12GS linker. This half of the marker was expressed in PHB neurons using a promoter that within the posterior half of the worm, is specific for these presynaptic neurons ($_{p}$*gpa-6*) (*22*). The complementary small fragment sfCherry3C$_{11}$ was similarly linked to *nlg-1* and expressed in AVA neurons using a promoter that within the posterior half of the worm, is specific for these postsynaptic neurons ($_{p}$*flp-18*) (*23*). Transgenic lines carrying both $_{p}$*gpa-6::nlg-1::sfCherry3C*$_{1\text{-}10}$ and $_{p}$*flp-18::nlg-1::sfCherry3C*$_{11}$ were generated. In wild-type animals, the distribution of red fluorescent puncta was similar to those observed in animals labeled with PHB-AVA NLG-1 GRASP (Fig. 5), and to that described by serial electron microscopy reconstruction (*11, 12*). To determine if these puncta were indeed synaptic, we introduced NLG-1 CLASP into *clr-1/RPTP* synaptogenesis mutants. In *clr-1/RPTP* mutants, PHB-AVA NLG-1 GRASP fluorescence intensity is dramatically reduced, and a PHB circuit-specific behavior is disrupted, indicating a reduction in synaptogenesis between the two neurons (*13*). We found that NLG-1 CLASP fluorescence intensity was also dramatically reduced in *clr-1/RPTP* mutants (Fig. 5), further indicating that NLG-1 CLASP puncta are synaptic.

Discussion

Taken together, we have proposed a dynamic association/dissociation model to explain the complementation process of self-associating split FP system. We have performed whole cell fluorescence intensity measurement and single-molecule brightness quantification to validate this model. Inspired by our model, we have adopted the SpyTag/SpyCatcher system to enhance the local concentrations of fluorescent protein fragments, thereby improving its complementation efficiency and overall brightness. We have also designed an engineering strategy by expressing the two fragments individually through a pETDuet vector. Using this platform, we have developed a bright sfCherry3C$_{1\text{-}10/11}$ (and its variant sfCherry3V$_{1\text{-}10/11}$) with enhanced complementation efficiency, enabling the high-efficient generation of human cell lines with endogenously sfCherry-labeled proteins. Moreover, we have transformed the split sfCherry3C into a trans-synaptic marker called NLG-1 CLASP and have demonstrated the visualization of specific synapses in living animals.

The proposed 2-step model, consisting of a dynamic association/dissociation equilibrium followed by an irreversible folding/maturation process, can be generalized to other split fluorescent proteins, including the non-self-associating ones used to monitor protein-protein interaction in bimolecular fluorescence complementation (BiFC) assays (*24*) (*25*). In fact, there is no definitive boundary between the non-self-associating and self-associating split FPs. Instead, their only difference is in the spontaneous binding affinity, which can be characterized by K_D' in our model. For BiFC analysis, this affinity is one of the major determinant of the sensitivity. If it is too low, the split FP will fail to produce sufficient complementation signal even when the probed molecular interaction does occur. On the other hand, the opposite extreme leads to high background complementation as observed in certain split constructs (*6*). While this K_D' is not straightforward to measure biochemically *in vitro* due to the overall irreversible nature of complementation, our steady state flow cytometry analysis provides a reliable way to characterize it in cells. Our model also shows that the overall complemented signal is determined not only thermodynamically by the initial binding affinity and local fragment concentrations, but also kinetically by the rates of folding, chromophore maturation and protein degradation.

The assisted complementation demonstrated by SpyTag/SpyCatcher presents a simple way to improve complementation efficiency between FP fragments. This strategy can be expanded to multicolor imaging using FP_{11} tagging using orthogonal binder/tag pairs such as SnoopTag/SnoopCatcher (engineered by splitting an adhesin from *Streptococcus pneumoniae* (*26*)) or SsrA/SspB (a degradation tag and its adaptor protein from bacterial ClpX ATPase (*27, 28*)).

The substantial improvement of complementation efficiency in $sfCherry3_{1-10/11}$ is beneficial for a wide variety of applications ranging from protein labeling to scaffolding protein complexes, as well as monitoring cell-cell connections. A highly efficient complementation process not only guarantees an enhanced overall brightness, but also enables us to tune down the expression level of FP_{1-10} fragments which might otherwise exhaust the cellular machinery that maintaining the protein homeostasis (biogenesis, folding, trafficking and degradation of proteins). This benefit is essential in scenarios that are sensitive to the expression of exogenous proteins, such as tagging endogenous proteins in embryos. Moreover, our engineering platform based on pETDuet vector can be utilized to optimize other self-associating split FPs with insufficient complementation efficiency, such as $mNeonGreen2_{1-10/11}$.

Utilizing the split $sfCherry3_{1-10/11}$ construct, we have established a completely orthogonal, red-colored trans-synaptic marker NLG-1-CLASP in addition to the original green fluorescent trans-synaptic marker NLG-1 GRASP. Similar cyan and yellow trans-synaptic markers (called dual-eGRASP) have recently been generated in vertebrates (*29*). The use of NLG-1 CLASP with these cyan and yellow trans-synaptic markers has the potential to allow simultaneous and differential labeling of synapses between a single neuron and three synaptic partners in live animals. Since many synaptic connections within the nematode have overlapping localizations within the nerve ring and other nerve bundles (*11, 12*), this tool will allow us to accurately visualize multiple subsets of a single neuron's connections. These tools may similarly be of use in densely innervated regions of the vertebrate nervous system, such as the hippocampus and cortex. Thus, we propose that NLG-1 CLASP will be a powerful tool with which to probe the development and plasticity of neural circuits within live animals.

Materials and Methods

Molecular cloning.

The DNA sequence of SpyCatcher 002 and SpyTag 002 (based on the reported sequence from (*17*)) were directly synthesized (Integrated DNA Technologies, IDT). The DNAs of histone H2B, TagBFP and mIFP were subcloned from mEmerald, TagBFP or mIFP fusion plasmids (cDNA source: the Michael Davidson Fluorescent Protein Collection at the UCSF Nikon Imaging Center) using Phusion High-Fidelity DNA Polymerase (Thermo Scientific). The P2A sequence are GCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTGGACCT. The lentiviral plasmids pSFFV-GFP$_{1\text{-}10}$, pSFFV-mNG2$_{1\text{-}10}$ and pSFFV-sfCherry2$_{1\text{-}10}$ were generated in our previous research (*9*). To build three pSFFV-mIFP-P2A-FP$_{11}$-SpyCatcher constructs (used in Figs. 1A-C), three DNA fragments encoding mIFP, P2A-FP$_{11}$ and SpyCatcher were ligated into linearized pSFFV vector (BamHI/NotI) using In-Fusion HD Cloning kit (Clontech) within one reaction. To construct pSFFV-TagBFP-P2A-sfCherry2$_{1\text{-}10}$ plasmid (used in Fig. 1D), two DNA fragments encoding TagBFP or P2A were ligated in to linearized pSFFV-sfCherry2$_{1\text{-}10}$ vector using In-Fusion. To generate pcDNA-SpyCatcher-6aa-sfCherry2$_{1\text{-}10}$ and pcDNA-SpyCatcher-15aa-sfCherry2$_{1\text{-}10}$ plasmids (used in Figs. 2A-B), PCR amplicons encoding SpyCatcher or sfCherry2$_{1\text{-}10}$ were cloned into digested pcDNA3.1 vectors (HindIII/BamHI), and the different linkers were achieved through designing overlapping primers with various linker lengths. To make pSFFV-SpyTag-sfCherry2$_{11}$-TagBFP and pSFFV-SpyTag-sfCherry2$_{11}$-TagBFP-H2B (used in Figs. 2B-C), DNA fragments encoding SpyTag-sfCherry2$_{11}$, TagBFP and H2B were ligated in to linearized pSFFV vector (BamHI/NotI) through In-Fusion.

The pETDuet-1 vector was kindly donated by Dr. Alexander Kintzer from Dr. Robert Stroud's laboratory at UCSF. We generated the initial plasmid for mutagenesis screening by two rounds of In-Fusion ligation reaction: inserting the PCR amplicon encoding sfCherry2$_{1\text{-}10}$ into the first multiple cloning site (MCS) of pETDuet-1 digested by NcoI, followed by inserting the DNA sequence encoding sfCherry2$_{11}$-SpyCatcher into the second MCS of pETDuet-1 digested by NdeI. The NcoI restriction site preserved in the final product, but the NdeI restriction site was destroyed. For the mammalian expression and lentiviral production, DNAs of sfCherry3C$_{1\text{-}10}$ were directly PCR amplified from identified pETDuet-1 construct (final mutant) and cloned into the lentiviral pSFFV vector (BamHI/NotI). To generate the sfCherry3V$_{1\text{-}10}$ variant, we introduced the point mutation L83W into pSFFV-sfCherry3C$_{1\text{-}10}$ plasmid using QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies). For the complete nucleotide sequence of sfCherry3C$_{1\text{-}10}$, sfCherry3V$_{1\text{-}10}$, SpyCatcher 002 and SpyTag 002, see Supplementary Table 1.

Constructs used in the NLG-1 CLASP application were generated using standard molecular techniques. To generate $_p$*gpa-6::nlg-1::sfCherry3C*$_{1\text{-}10}$ construct (MVC227), *sfCherry3C*$_{1\text{-}10}$ was amplified from *pSFFV-sfCherry3C*$_{1\text{-}10}$ using the following primers: MVP846 (AGCTGCTAGCATGGAACGCATTTATCTTCTTCTCCTTCTTTTTCTGCCCAGGATACGATCCATGGAGGAGGACAACATGG) and MVP847 (TCCGGAGCTCGTCCTCGTTGTGGCTGGT). The fragment was subcloned into $_p$*gpa-6::nlg-1::GFP*$_{1\text{-}10}$ (MVC6) (*20*), replacing GFP$_{1\text{-}10}$, using the NheI and SacI sites. To generate the $_p$*flp-18::nlg-1::sfCherry2*$_{11}$ (MVC228) construct, *sfCherry2*$_{11}$ was amplified from *H*$_2$*B-sfCherry2*$_{11}$ using the following primers: MVP848 (AGCTGCTAGCATGGAACGCATTTATCTTCTTCTCCTTCTTTTTCTGCCCAGGATACGATCCTACACCATCGTGGAGCAGT) and MVP849 (TCCGGAGCTCGGTGCTGTGTCTGGCCTC). The fragment was subcloned into $_p$*flp-18::nlg-1::GFP*$_{11}$ (MVC12)(*20*), replacing GFP$_{11}$, using the NheI and SacI sites**.**

Cell culture and lentiviral production.

Human HEK 293T cells (UCSF cell culture facility) were maintained in Dulbecco's modified Eagle medium with high glucose (Gibco), supplemented with 10% (vol/vol) FBS and 100 μg/ml penicillin/streptomycin (UCSF Cell Culture Facility). U2OS cells (American Type Culture Collection, Manassas, VA) were cultured in DMEM media, supplemented with 10% fetal bovine serum. All cells were grown at 37 °C and 5% CO_2 in a humidified incubator. For the lentiviral production, 1 X 10^6 HEK 293T cells were plated into T25 one day prior to transfection. 430 ng of pMD2.G plasmid, 3600 ng of pCMV-dR8.91 plasmid and 4100 ng of the lentiviral plasmid (pSFFV-sfCherry2$_{1\text{-}10}$, pSFFV-sfCherry3C$_{1\text{-}10}$ and pSFFV-sfCherry3V$_{1\text{-}10}$) were co-transfected into HEK 293T cells using FuGENE HD (Promega) following the manufacturer's recommended protocol. The virus containing supernatant is harvested 48 h after transfection and were centrifuged to pellet any packaging cells. Virus containing medium is used immediately or stored in -80 °C freezer for future use.

For single-molecule brightness measurement sample preparation, U2OS cells were grown in 24-well plates with #1.5 glass coverslip bottoms (In Vitro Scientific) and transfected ~24 hours before measurement using GenJet transfection reagent (SignaGen Laboratories) according to the manufacturer's instructions. Immediately before measurement, the growth media was exchanged with PBS buffer with calcium and magnesium (Gibco).

Sample preparation and data analysis in flow cytometry.

To characterize the relationship between complementation efficiency and expression level in split GFP, split mNG2 and split sfCherry2, we made pSFFV-mIFP_P2A_full-length-FP and pSFFV-mIFP_P2A_FP$_{11}$-SpyCatcher constructs. Corresponding to each scatter plots in Figures 1B-D, 3 X 10^4 HEK 293T cells grown on 48-well plate (Eppendorf) were co-transfected with (B) left: 100 ng pSFFV-mIFP_P2A_GFP$_{11}$-SpyCatcher with 200 ng pSFFV-GFP$_{1-10}$, right: 100 ng pSFFV-mIFP_P2A_GFP[full-length] with 200 ng pSFFV-GFP$_{1-10}$; (C) left: 100 ng pSFFV-mIFP_P2A_mNG2$_{11}$-SpyCatcher with 200 ng pSFFV-mNG2$_{1-10}$, right: 100ng pSFFV-mIFP_P2A_mNG2[full-length] with 200 ng pSFFV-mNG2$_{1-10}$; (D) left: 100 ng pSFFV-mIFP_P2A_sfCherry2$_{11}$-SpyCatcher with 200 ng pSFFV-sfCherry2$_{1-10}$, right: 100 ng pSFFV-mIFP_P2A_sfCherry2[full-length] with 200 ng pSFFV-sfCherry2$_{1-10}$. In Figure 1E, same cells were co-transfected with 100 ng pSFFV-mIFP_P2A_sfCherry2$_{11}$-SpyCatcher and 200 ng pSFFV-sfCherry2$_{1-10}$-TagBFP.

To test the Spy002 pair assisted complementation in sfCherry2, we built two tandem-binder constructs (pcDNA-SpyCatcher-6aa-sfCherry2$_{1-10}$ and pcDNA-SpyCatcher-15aa-sfCherry2$_{1-10}$) and two tandem-tag constructs (pSFFV-SpyTag-sfCherry2$_{11}$-TagBFP and pSFFV-sfCherry2$_{11}$-SpyTag-TagBFP). We then combinatorially co-transfected the same cells with 100ng tag construct plus 200 ng binder construct. In Figure 2B, to compare the complementation efficiency with or without Spy pair interaction, cells were co-transfected with either 100 ng pSFFV-SpyTag-sfCherry2$_{11}$-TagBFP plus 200 ng pcDNA-sfCherry2$_{1-10}$, or 100 ng pSFFV-SpyTag-sfCherry2$_{11}$-TagBFP plus 200 ng SpyCatcher-6aa-sfCherry2$_{1-10}$.

To validate the increased complementation of split sfCherry3 variants versus split sfCherry2, we made pSFFV-sfCherry3C$_{1-10}$ and pSFFV-sfCherry3V$_{1-10}$ constructs. Corresponding to each scatter plots in Figure 3E, same cells were co-transfected using with (left) 100ng pSFFV-mIFP_P2A_sfCherry2$_{11}$-SpyCatcher with 200ng pSFFV-sfCherry2$_{1-10}$; (middle) 100ng pSFFV-mIFP_P2A_sfCherry2$_{11}$-SpyCatcher with 200ng pSFFV-sfCherry3C$_{1-10}$; (right) 100ng pSFFV-mIFP_P2A_sfCherry2$_{11}$-SpyCatcher with 200ng pSFFV-sfCherry3V$_{1-10}$.

For flow cytometry analysis, 48 hours after transfection, transfected HEK 293T cells were digested with Trypsin-EDTA (0.25%) (Gibco) into single cells and re-suspended in 0.5 ml PBS solution. Analytical flow cytometry was carried out on a LSR II instrument (BD Biosciences) and cell sorting on a FACSAria II (BD Biosciences) in Laboratory for Cell Analysis at UCSF. Flow cytometry data analysis (gating and plotting) was conducted using the FlowJo software (FlowJo LLC).

Single-molecule brightness measurement and data analysis.

Fluorescence brightness measurements were carried out on a homebuilt two-photon microscope, which has been previously described (*30*). Pulsed laser light (100 fs pulses with a repetition frequency of 80 MHz) from a mode-locked Ti-Sapphire laser (Mai-Tai, Spectra Physics) was focused by a 63X C-Apochromat water immersion objective (NA = 1.2, Zeiss) to create two-photon excitation. The emitted fluorescence was collected by the same objective and separated from the excitation light by a dichroic mirror (675DCSXR, Chroma Technology). The fluorescence emission was separated into two detection channels by a 580nm dichroic mirror (585DCXR, Chroma Technology), and the green channel was further filtered by an 84nm-wide bandpass filter centered at 510 nm (FF01-510/84-25, Semrock). Avalanche photodiodes (SPCM-AQ-14, Perkin-Elmer) detected the fluorescence signal, and photon counts were recorded by a data acquisition card (FLEX02, Correlator.com) for ~60 seconds with a sampling frequency of 200 kHz. All measurements were carried out at an excitation wavelength of 1000 nm and a measured power after the objective of ~0.46 mW. The photon count record was analyzed to recover Mandel's Q parameter as previously described (*31*) using programs written for IDL 8.7 (Research Systems, Inc.). The Q value is converted into brightness λ, which represents the average fluorescence intensity per molecule, using the relation $Q = \gamma_2 \lambda T$, where T represents the sampling time and γ_2 is a shape-dependent factor whose value has been determined as described previously (*32*).

Mutagenesis and screening.

The amino-acid sequence of sfCherry2 and the split site were from our previous published literature (*9*). The sfCherry2$_{1\text{-}10}$ sequence was subjected to random mutagenic PCR (forward primer: AGGAGATATACCATGGAGGAGGACAAC, reverse primer: CTGCTGCCCATGTCAGTCCTCGTTGTG) using the GeneMorph II Random Mutagenesis Kit (Agilent Technologies). A high mutation rate protocol suggested in the instruction manual was adapted, with an initial target DNA amount of 0.2 µg and 30-cycle amplification. The cDNA library pool was gel-purified and ligated into a PCR-linearized pETDuet vector (forward primer: TGACATGGGCAGCAGCCA , reverse primer: CATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGG. The product only contains the sfCherry2$_{11}$-SpyCatcher coding sequence in the 2nd MCS but not the sfCherry2$_{1\text{-}10}$ in the 1st MCS.) using In-Fusion. The plasmid pool was then transformed into *E.coli* BL21 (DE3) electrocompetent cells (Lucigen) by electroporation using the Gene Pulser Xcell™

Electroporation Systems (BioRad). The expression library was evenly plated on nitrocellulose membrane (Whatman, 0.45 µm pore size), which was sitting on an LB-agar plate with 30 mg/ml kanamycin. After overnight growth at 37 °C, the nitrocellulose membrane was carefully transferred onto a new LB-agar plate containing 1 mM Isopropyl-β-D-thiogalactoside (IPTG) and 30 mg/ml kanamycin and cultured for another 3-6 hours at 37 °C to induce the protein production. We performed the clone screening by imaging the IPTG-containing LB-agar plate using a BioSpectrum Imaging System (UVP). The brightest candidates in each library were pooled (typically ~20 from approximately 10,000 colonies) and served as templates for the next round of directed evolution. The DNA sequences of selected constructs were confirmed by sequencing (Quintara Biosciences). For DNA shuffling, we adopted the protocol described in Yu et al (*33*). Specifically, we PCR amplified the brightest six sfCherry3$_{1-10}$ variants (forward primer: AGGAGATATACCATGGAGGAGGACAAC, reverse primer: CTGCTGCCCATGTCAGTCCTCGTTGTG) from the last round of random mutagenesis screening. PCR products of 651 bp were purified from 1% agarose gels using zymoclean gel DNA gel recovery kit (Zymo Research). The DNA concentrations were measured in Nanodrop and the fragments were mixed at equal amounts for a total of ~2 µg. The mixture was then digested with 0.5 unit DNase I (New England Biolabs) for 13min and terminated by heating at 95°C for 10 min. The DNase I digests were run on a 2% agarose gel, and the band with a size of 50-100 bp was selected and purified. 10 µl of purified fragments was added to 10 µl of Phusion High-Fidelity PCR Master Mix and reassembled with a PCR program of 30 cycles, with each cycle consisting of 95 °C for 60 s, 50 °C for 60 s and 72 °C for 30 s. After reassembly, 1 µl of this reaction was amplified by PCR. The shuffled library was then transformed, expressed and screened as described above. After the directed evolution was saturated (no apparent fluorescence increase in the induced colonies), the brightest clone was selected and the DNA sequences of the constructs were confirmed by sequencing (Quintara Biosciences). The emission spectra of split sfCherry variants (in *E.coli* solution culture expressing FPs) were measurement on a Synergy™ H4 Hybrid Multi-Mode Microplate Reader from UCSF Center for Advanced Technology.

sgRNA production, RNP nucleofection and lentiviral transduction

We purchased synthetic single strand DNA oligos from Integrated DNA Technologies (IDT). We prepared sgRNAs and Cas9/sgRNA RNP complexes following our published methods (*4*). Specifically, sgRNAs were obtained by *in vitro* transcribing DNA templates containing a T7 promoter (TAATACGACTCACTATAG), an sgRNA scaffold region and the gene-specific 20 nt sgRNA sequence. DNA templates were produced by overlapping PCR using a set of 4 primers: 3 common primers (forward primer T25: 5'- TAA TAC GAC TCA CTA TAG -3'; reverse primer BS7: 5'- AAA AAA AGC ACC GAC TCG GTG C -3' and reverse primer ML611: 5'- AAA AAA AGC ACC GAC TCG GTG CCA CTT TTT CAA GTT GAT AAC GGA CTA GCC TTA TTT AAA CTT GCT ATG CTG TTT CCA GCA TAG CTC TTA AAC -3') and one gene-specific primer (forward primer 5'- TAA TAC GAC TCA CTA TAG NNN NNN NNN NNN NNN NNN NNG TTT AAG AGC TAT GCT GGA A -3'). For each template, a 50 µL PCR was performed with Phusion® High-Fidelity PCR Master Mix (New England Biolabs) reagents with the addition of 1 µM T25, 1µM BS7, 20 nM ML611 and 20 nM gene-specific primer. The PCR product was purified and eluted in 12 µL of RNAse-free DNA buffer. Next, a 100-µL *in vitro* transcription reaction was performed with ~300 ng DNA template from PCR product and 1000 U of T7 RNA polymerase in buffer containing 40 mM Tris pH 7.9, 20 mM $MgCl_2$, 5 mM DTT, 2 mM spermidine and 2 mM of each NTP (New England BioLabs). Following a 4 h incubation at 37°C, the sgRNA product was purified and eluted in 15 µL of RNAse-free RNA buffer. The sgRNA was quality-checked by running 5 pg of the product on Mini-PROTEAN TBE Precast Gels (Bio-Rad Laboratories) at 200-voltage for 60~80 min.

For the knock-in of sfCherry2_{11} into HEK 293T wild-type (WT) cells, 200-nt homology-directed recombination (HDR) templates were ordered in single-stranded DNA (ssDNA) from IDT. For the complete set of DNA sequence used for sgRNA in vitro transcription or HDR templates, see Supplementary Tables 2 and 3. Cas9 protein (pMJ915 construct, containing two nuclear localization sequences) was expressed in *E.coli* and purified by the University of California Berkeley Macrolab. HEK 293T WT cells were treated with 200 ng/mL nocodazole (Sigma) for ~17 hours before electroporation to increase HDR efficiency. 100 pmol Cas9 protein and 130 pmol sgRNA were assembled into Cas9/sgRNA RNP complexes just before nucleofection and combined with 150 pmol HDR template in a final volume of 10 µL. Electroporation was performed in an Amaxa 96-well shuttle Nuleofector device (Lonza) using SF-cell line reagents (Lonza). Nocodazole-treated cells were resuspended to 10^4 cells/µL in SF solution immediately prior to electroporation. For each sample, 20 µL of cells was added to the 10 µL RNP/template mixture. Cells were immediately electroporated using the CM130 program and transferred to 48-well plate with pre-warmed medium. Electroporated cells were cultured and expanded for 7-10 days prior to lentiviral transduction.

One day before lentiviral transduction, knocked-in cells well split and seeded in 12-well plate at 6~8 X 10^4 per well for 4 wells. The confluency should reach 70~80% on the day of transduction.

The lentivirus titer of sfCherry2$_{1-10}$, sfCherry3C$_{1-10}$ or sfCherry3V$_{1-10}$ was quantified independently by Lenti-X GoStix Plus kit (Takara, Cat# 631280) immediately before infection. And the supernatant was diluted (around 1:5 dilution) with fresh medium to make sure the final virus concentrations are the same across groups. The experimental groups were treated with 1ml diluted sfCherry2$_{1-10}$, sfCherry3C$_{1-10}$ or sfCherry3V$_{1-10}$ viral supernatant supplemented with 10 ug polybrene (MilliporeSigma) respectively. The negative control was treated with fresh medium supplemented with the same concentration of polybrene. 24-hour after infection, the viral supernatant was swapped with fresh medium. After another 48 to 72 hours, the infected cells were harvested for flow cytometry analysis and cell sorting.

Fluorescence Microscopy

To validate the Spy-assisted complementation can improve the overall brightness in protein labeling (Fig. 2C), we transfected HEK 293T cells (1.5 X 10^4 per well) grown on an 8-well glass bottom chamber (Thermo Fisher Scientific) using FuGene HD (Promega). In order to achieve better cell attachment, 8-well chamber was coated with Fibronectin (Sigma-Aldrich) for 45 min before seeding cells. Total plasmid amount of 180 ng per well with the FP$_{11}$ to FP$_{1-10}$ ratio in 1:2 was used to achieve optimal labeling. 48 hours after transfection, cells were fixed with 4% paraformaldehyde and then imaged on a Nikon Ti-E inverted wide-field fluorescence microscope equipped with an LED light source (Excelitas X-Cite XLED1), a 40× 0.55 NA air objective (Nikon), a motorized stage (ASI) and an sCMOS camera (Tucsen).

Live-cell imaging (Figs. 4B-G) of sorted successful knock-in cells (with either sfCherry3C$_{1-10}$ or sfCherry3V$_{1-10}$ infection) was acquired on an inverted Nikon Ti-E microscope (UCSF Nikon Imaging Center), a Yokogawa CSU-W1 confocal scanner unit, a Plan Apo VC 100× 1.4 NA oil immersion objective, a stage incubator, an Andor Zyla 4.2 sCMOS camera and MicroManager2.0 software. Fluorescence recovery after photobleaching (FRAP) experiment (fig. S1) was performed on the same microscopy with a Vortran 473 nm laser for photobleaching and a Rapp Optoelectronic UGA-40 photobleaching system (COM5). The frame interval of the movie is 2.5 s and the total length is 10 min. The photobleaching was conducted by scanning the 473 nm (50 mW) laser over the region of interest at the 20 s of the movie. For optimal cell attachment, 8-well glass bottom chamber was coated with Fibronectin (Sigma-Aldrich) for 45 min before seeding sorted endogenously tagged HEK 293T cells. A background image (taken at the same imaging condition without putting on any sample) was subtracted from the live-cell microscopy images using the ImageJ software. Analysis of fluorescence microscopy images were performed in ImageJ.

Generation of NLG-1 CLASP in *C. elegans*

All *C. elegans* strains were maintained using standard protocols(*34*) and were raised on 60 mm Nematode Growth Media plates seeded with OP50 *Escherichia coli* at 20 °C. Wild-type strains were *C. elegans* variety Bristol, strain N2, and the mutant strain used for this study was *clr-1(e1745) II.* Transgenic strains include *wyEx1982 (20), wyEx1982; clr-1, iyEx368 and iyEx368; clr-1. wyEx1982* contains extra-chromosomal PHB-AVA NLG-1 GRASP marker (*$_p$gpa-6::nlg-1::GFP$_{1-10}$* (60 ng/µl), *$_p$flp-18::nlg-1::GFP$_{11}$* (30 ng/µl), *$_p$nlp-1::mCherry* (10 ng/µl), *$_p$flp-18::mCherry* (5 ng/µl) and *$_p$odr-1::DsRedII* (20 ng/µl)). *iyEx368* contains the extra-chromosomal PHB-AVA NLG-1 CLASP marker (*$_p$gpa-6::nlg-1::sfCherry3C$_{1-10}$* (68 ng/µl), *$_p$flp-18::nlg-1::sfCherry2$_{11}$* (38.6 ng/µl), *$_p$nlp-1::GFP* (4.5ng/µl), and *$_p$odr-1::DsRedII* (39 ng/µl)).

A Zeiss Axio Imager.A1 compound fluorescent microscope was used to capture images of live *C. elegans* under 63X magnification. Worms were paralyzed on 2% agarose pads using a 2:1 ratio of 0.3 M 2,3-butanedione monoxime (BDM) and 10 mM levamisole in M9 buffer. All micrographs taken of PHB-AVA NLG-1 GRASP and NLG-1 CLASP markers were of larval stage 4 animals. All data from micrographs were quantified using NIH ImageJ software. Intensity of PHB-AVA NLG-1 GRASP and PHB-AVA NLG-1 CLASP was measured as previously described (*13, 35*). Briefly, the intensity at each pixel within each synaptic puncta was measured using NIH ImageJ. To account for differences in background fluorescence, background intensity was estimated by calculating the minimum intensity value in a region immediately adjacent to the puncta. This minimum intensity value was then subtracted from the intensity for each pixel, and the sum of the adjusted values was calculated. For control, pictures of wild-type animals were also taken on the same day using the same settings.

Acknowledgement

We thank Dr. David Alexander Brown, Dr. Yina Wang for extensive discussion on data analysis, Dr. Bin Yang for help in fluorescence light microscopy, Alejandro Ramirez for preparing reagents and supplies, Dr. Kari Herrington for help in FRAP experiment, Dr. Noelle L'Etoile for advice on NLG-1 CLASP experiments, Dr. Xiaokun Shu for sharing the BioSpectrum Imaging System, and Dr. Joseph DeRisi for generously sharing the nucleofector device. This work is supported by the National Institutes of Health R21EB022798 and R01GM124334 (to B.H.), UCSF Program for Breakthrough Biomedical Research (Byers Award in Basic Science to B.H.), the National Institutes of Health R01GM064589 (to J.K. and J.D.M.), the National Institutes of Health R01NS087544 and SC3GM089595 (to M.V.), T34GM008253 MARC fellowship (to N.A.), and the National Science Foundation 1355202 (to M.V.). B.H. is a Chan Zuckerberg Biohub investigator.

Competing Financial Interest

The authors declare no competing financial interests.

Author Contributions

S.F. and B.H. conceived and designed the research; S.F. performed the flow cytometry, random mutagenesis and screening, CRISPR knock-in, FACS, confocal microscopy and FRAP experiments; J.K and J.D.M. performed single-molecule brightness measurement and analysis; S.F. and C.M. performed the molecular cloning and site-directed mutagenesis experiments; A.V., D.C.V., F.F., N.A. and M. V. perfored the NLG-1 CLASP experiments and analysis; S.F. and B.H. analyzed the data; S.F. and B.H. wrote the manuscript.

Figures

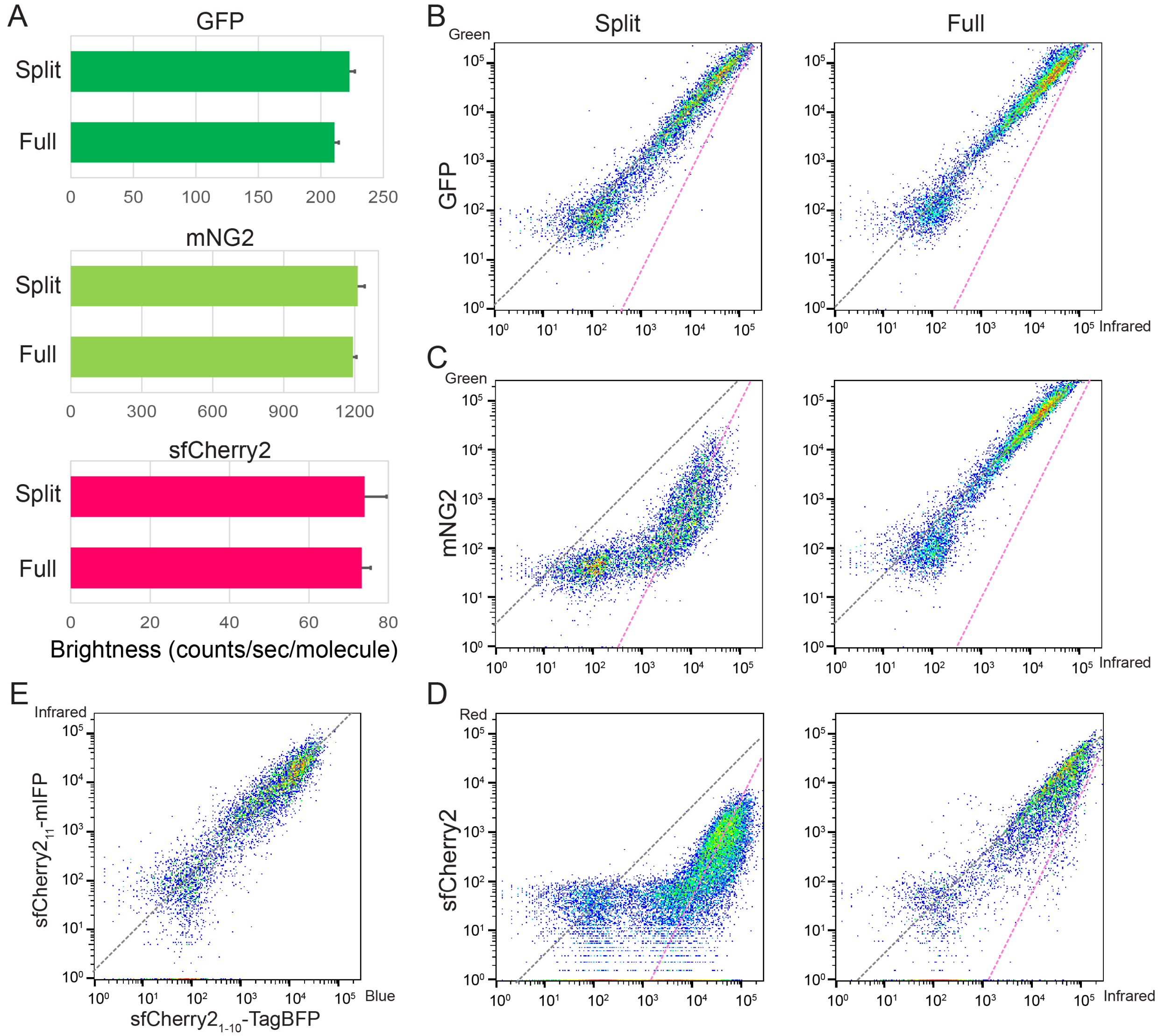


Fig. 1. Characterization of split fluorescent proteins. **(A)** Single molecule brightness measurement of split FPs and their full-length counterparts using fluorescence fluctuation spectroscopy. *N* = 10. Error bars are standard deviations. **(B to D)** Flow cytometry analysis of whole cell fluorescence in HEK 293T expressing either **(B)** GFP$_{1\text{-}10/11}$, **(C)** mNG2$_{1\text{-}10/11}$, and **(D)** sfCherry2$_{1\text{-}10/11}$ or their full-length counterparts. The x-axis is the log-scale infrared fluorescence intensity indicating the expression level, and the y-axis is the log-scale green (or red) fluorescence intensity. The grey dashed trend lines have a slope of 1 and intercepts are set to best follow the points in the right (full-length) panels. The pink dashed trend lines have a slope of 2 and intercepts set to best follow the points in the left (split) panels with the exception of GFP$_{1\text{-}10/11}$. **(E)** Expression levels of two fragments are proportional within a wide range of expression levels in a co-transfect experiment. The grey dashed trend line has a slope of 1.

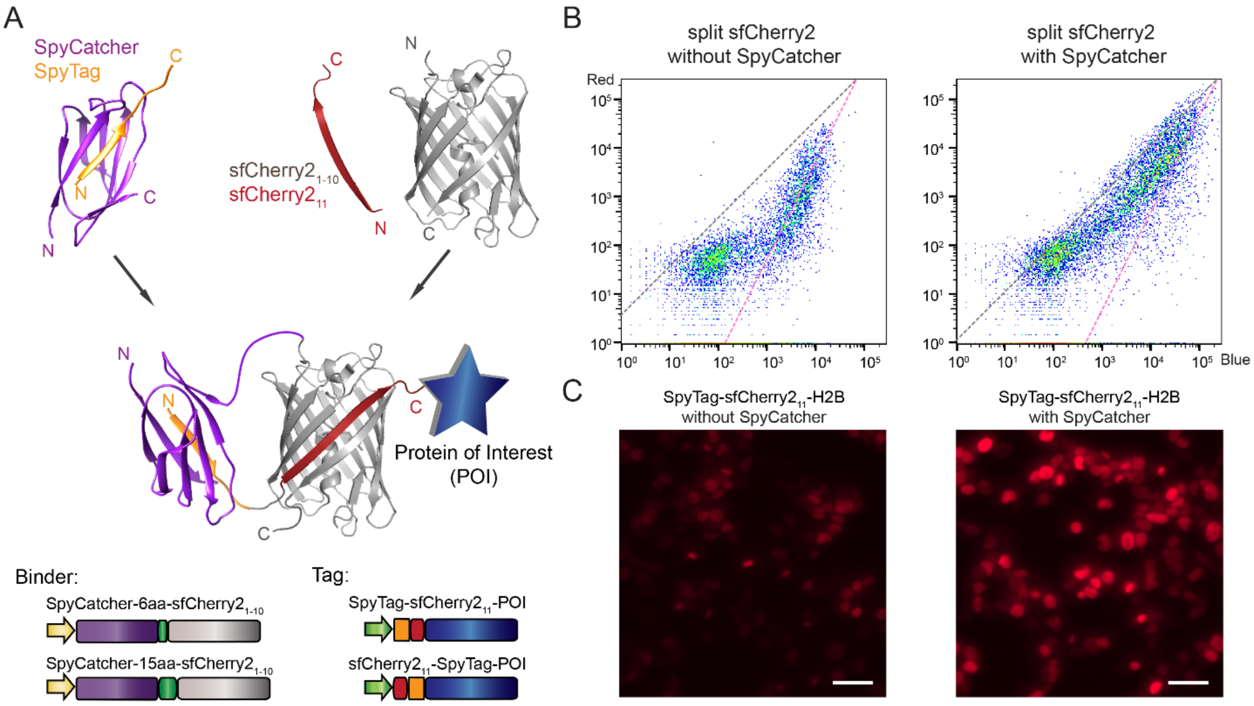

A
SpyCatcher
SpyTag
C
N
C
N
C
sfCherry2$_{1-10}$
sfCherry2$_{11}$
N
N
C
N
N
C
C
Protein of Interest
(POI)
Binder:
SpyCatcher-6aa-sfCherry2$_{1-10}$
SpyCatcher-15aa-sfCherry2$_{1-10}$
Tag:
SpyTag-sfCherry2$_{11}$-POI
sfCherry2$_{11}$-SpyTag-POI
B
split sfCherry2
without SpyCatcher
split sfCherry2
with SpyCatcher
Red
10^5
10^4
10^3
10^2
10^1
10^0
10^0
10^1
10^2
10^3
10^4
10^5
Blue
C
SpyTag-sfCherry2$_{11}$-H2B
without SpyCatcher
SpyTag-sfCherry2$_{11}$-H2B
with SpyCatcher

Fig. 2. SpyTag/SpyCatcher assisted complementation of split sfCherry2$_{1-10/11}$. (A) Schematic diagram of SpyTag/SpyCatcher-assisted complementation and construct design. **(B)** Flow cytometry analysis of whole cell fluorescence in HEK 293T cells expressing SpyTag-sfCherry2$_{11}$-TagBFP with either sfCherry2$_{1-10}$ alone or with SpyCatcher-6aa-sfCherry2$_{1-10}$. The grey dashed trend line has a slope of 1 and the pink one has a slope of 2. **(C)** Fluorescence microscopy of sfCherry2$_{11}$ labeled H2B without (left) or with (right) the assistance of SpyTag/SpyCatcher interaction. The imaging condition and brightness/contrast range were set the same for better comparison. Scale bars: 50 μm.

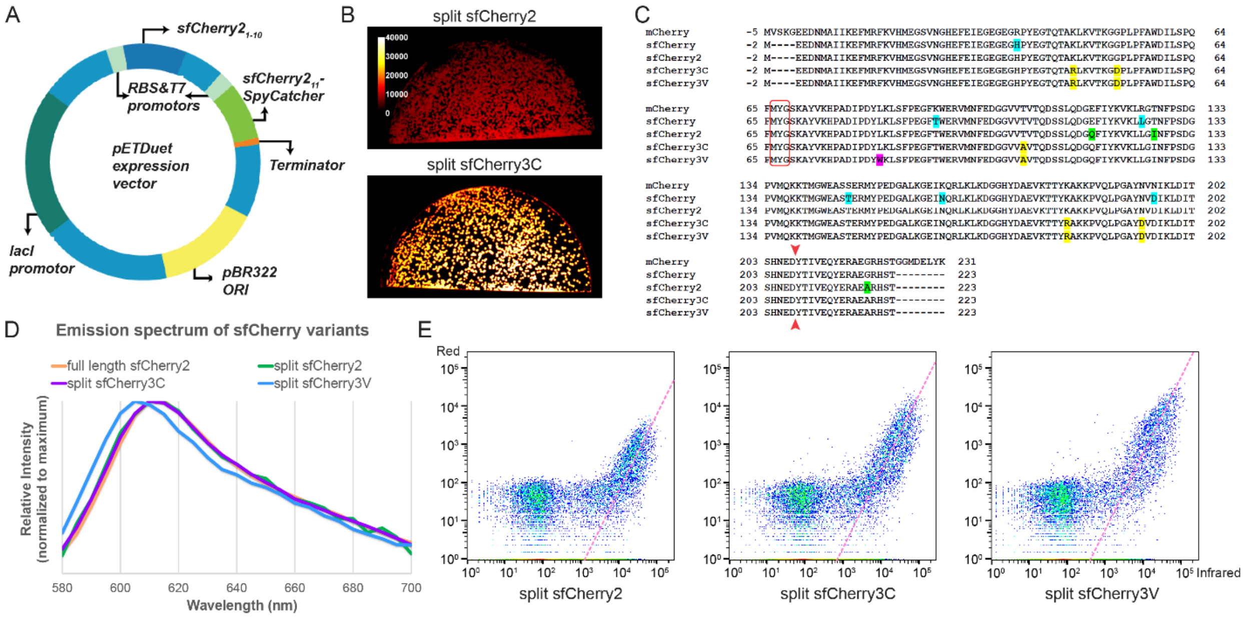


Fig. 3. Engineering and characterization of split sfCherry3. (A) Schematics of pETDuet-based engineering platform. **(B)** Fluorescence images of E.coli colonies expressing split sfCherry2 or sfCherry3C from the pETDuet constructs. **(C)** Protein sequence alignment of mCherry, sfCherry, sfCherry2, sfCherry3C and sfCherry3V. The amino acids forming the chromophore are indicated by a red box. The split site is indicated by the red arrow. Starting from mCherry, mutations introduced in sfCherry, sfCherry2, sfChery3C and sfCherry3V are highlighted in cyan, green, yellow and magenta, respectively. The overall alignment numbering follows that of sfCherry. **(D)** Emission spectra of sfCherry variants. **(E)** Flow cytometry analysis of whole cell fluorescence in HEK 293T cells expressing mIFP-P2A- sfCherry2$_{11}$-SpyCatcher and sfCherry2$_{1-10}$ (left)/ sfCherry3C$_{1-10}$ (middle)/ sfCherry3V$_{1-10}$ (right). The pink dashed lines are results from linear least-square fitting with a fixed slope of 2 (see fig. S4).

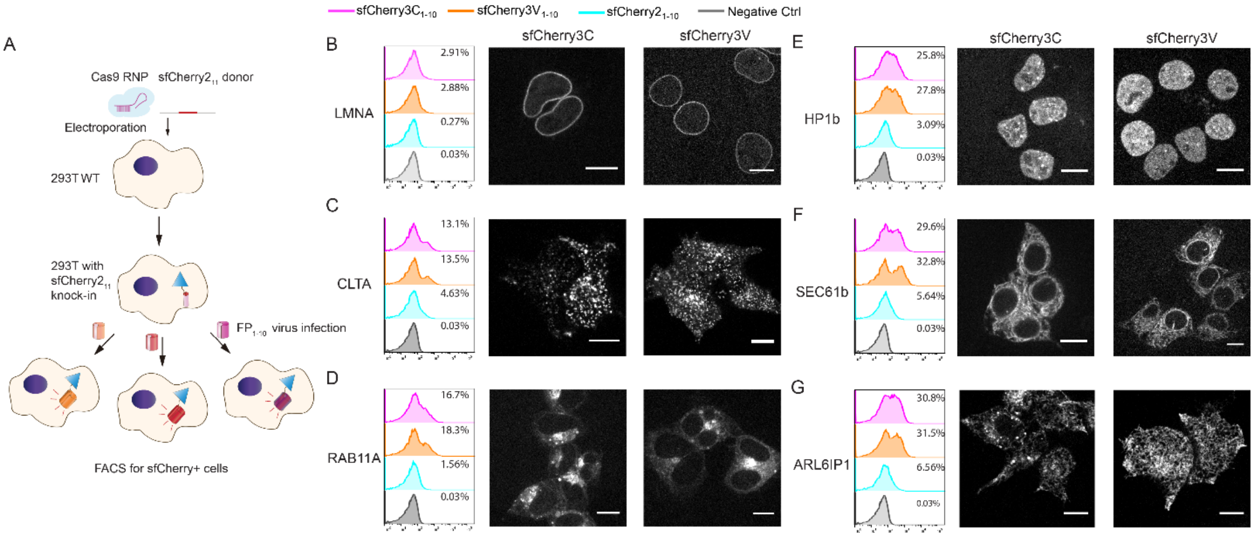

sfCherry3C1-10
sfCherry3V1-10
sfCherry2 1-10
Negative Ctrl
A
Cas9 RNP
sfCherry2 11 donor
Electroporation
293T WT
293T with sfCherry2 11 knock-in
FP 1-10 virus infection
FACS for sfCherry+ cells
B
LMNA
2.91%
2.88%
0.27%
0.03%
sfCherry3C
sfCherry3V
C
CLTA
13.1%
13.5%
4.63%
0.03%
D
RAB11A
16.7%
18.3%
1.56%
0.03%
E
HP1b
25.8%
27.8%
3.09%
0.03%
sfCherry3C
sfCherry3V
F
SEC61b
29.6%
32.8%
5.64%
0.03%
G
ARL6IP1
30.8%
31.5%
6.56%
0.03%

Fig. 4. Endogenous protein labeling in HEK 293T cells using sfCherry3 variants. (A) Schematic diagram of knock-in followed by virus infection and FACS enrichment. **(B to G)** Analysis of FACS sorting efficiency in six targets, **(B)** lamin A/C, **(C)** clathrin light chain A, **(D)** RAB11A, **(E)** heterochromatin protein 1 β, **(F)** ER translocon complex SEC61b and **(G)** ER tubule protein ARL6IP1, and visualization of sorted knock-in cells through confocal fluorescence microscopy. Scale bar: 10 μm.

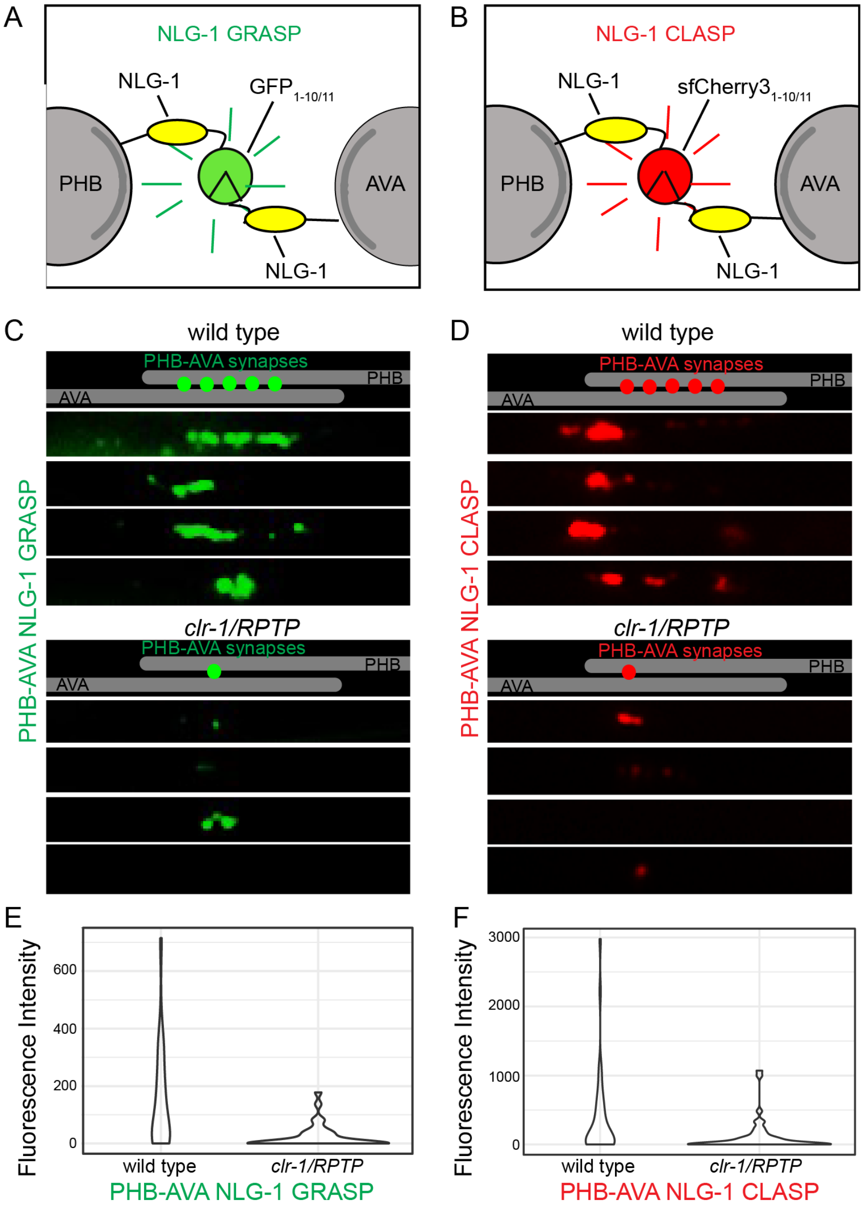

A
NLG-1 GRASP
NLG-1
GFP 1-10/11
PHB
AVA
NLG-1
B
NLG-1 CLASP
NLG-1
sfCherry3 1-10/11
PHB
AVA
NLG-1
C
wild type
PHB-AVA synapses
PHB
AVA
clr-1/RPTP
PHB-AVA synapses
PHB
AVA
PHB-AVA NLG-1 GRASP
D
wild type
PHB-AVA synapses
PHB
AVA
clr-1/RPTP
PHB-AVA synapses
PHB
AVA
PHB-AVA NLG-1 CLASP
E
Fluorescence Intensity
600
400
200
0
wild type
clr-1/RPTP
PHB-AVA NLG-1 GRASP
F
Fluorescence Intensity
3000
2000
1000
0
wild type
clr-1/RPTP
PHB-AVA NLG-1 CLASP

Fig. 5. NLG-1 CLASP visualizes specific subsets of synapses in live *C. elegans*. (A & B) Schematic diagram of GFP$_{1\text{-}10/11}$-based NLG-1 GRASP and sfCherry3C$_{1\text{-}10/11}$-based NLG-1 CLASP in PHB and AVA neurites. **(C & D)** Schematics and micrographs of NLG-1 GRASP and NLG-1 CLASP specifically labelling synaptic contacts between PHB and AVA neurons in *C. elegans*. PHB-AVA NLG-1 GRASP and NLG-1 CLASP fluorescence intensity are dramatically reduced in the synaptic partner recognition mutant *clr-1/RPTP*. **(E & F)** Quantification of the reduction in relative fluorescence intensities of NLG-1 GRASP and NLG-1 CLASP in *clr-1/RPTP* mutants. $N \geq 40$.

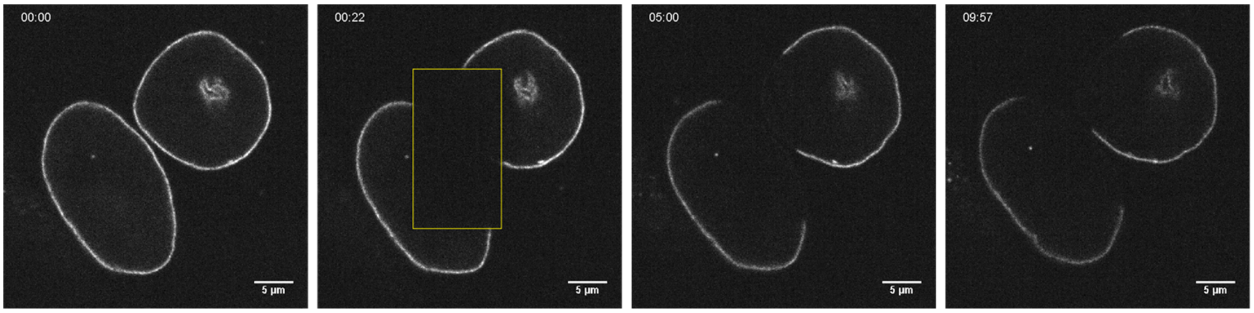
References

1. S. Cabantous, T. C. Terwilliger, G. S. Waldo, Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat Biotechnol* **23**, 102-107 (2005).
2. S. B. Van Engelenburg, A. E. Palmer, Imaging type-III secretion reveals dynamics and spatial segregation of Salmonella effectors. *Nat Methods* **7**, 325-330 (2010).
3. D. Kamiyama *et al.*, Versatile protein tagging in cells with split fluorescent protein. *Nat Commun* **7**, 11046 (2016).
4. M. D. Leonetti, S. Sekine, D. Kamiyama, J. S. Weissman, B. Huang, A scalable strategy for high-throughput GFP tagging of endogenous human proteins. *Proc Natl Acad Sci U S A* **113**, E3501-3508 (2016).
5. S. Cabantous *et al.*, Recent advances in GFP folding reporter and split-GFP solubility reporter technologies. Application to improving the folding and solubility of recalcitrant proteins from Mycobacterium tuberculosis. *J Struct Funct Genomics* **6**, 113-119 (2005).
6. W. Chun, G. S. Waldo, G. V. Johnson, Split GFP complementation assay: a novel approach to quantitatively measure aggregation of tau in situ: effects of GSK3beta activation and caspase 3 cleavage. *J Neurochem* **103**, 2529-2539 (2007).
7. H. Ishikawa, F. Meng, N. Kondo, A. Iwamoto, Z. Matsuda, Generation of a dual-functional split-reporter protein for monitoring membrane fusion using self-associating split GFP. *Protein Eng Des Sel* **25**, 813-820 (2012).
8. E. H. Feinberg *et al.*, GFP Reconstitution Across Synaptic Partners (GRASP) defines cell contacts and synapses in living nervous systems. *Neuron* **57**, 353-363 (2008).
9. S. Feng *et al.*, Improved split fluorescent proteins for endogenous protein labeling. *Nat Commun* **8**, 370 (2017).
10. D. Batan *et al.*, A Multicolor Split-Fluorescent Protein Approach to Visualize Listeria Protein Secretion in Infection. *Biophys J* **115**, 251-262 (2018).
11. J. G. White, E. Southgate, J. N. Thomson, S. Brenner, The structure of the nervous system of the nematode Caenorhabditis elegans. *Philos Trans R Soc Lond B Biol Sci* **314**, 1-340 (1986).

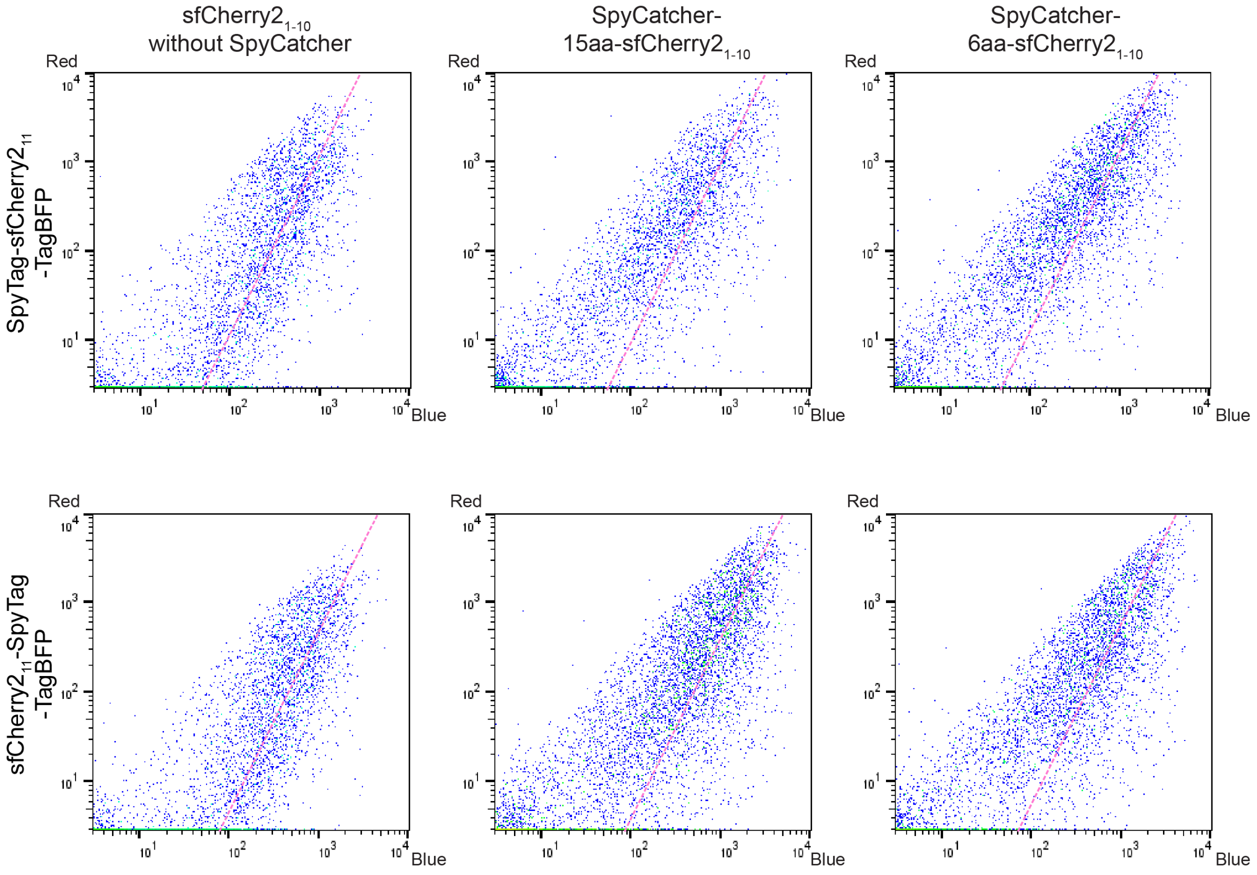
12. D. H. Hall, R. L. Russell, The posterior nervous system of the nematode Caenorhabditis elegans: serial reconstruction of identified neurons and complete pattern of synaptic interactions. *J Neurosci* **11**, 1-22 (1991).
13. A. Varshney *et al.*, The receptor protein tyrosine phosphatase CLR-1 is required for synaptic partner recognition. *PLoS Genet* **14**, e1007312 (2018).
14. B. D. Slaughter, R. Li, Toward quantitative "in vivo biochemistry" with fluorescence fluctuation spectroscopy. *Mol Biol Cell* **21**, 4306-4311 (2010).
15. B. Zakeri *et al.*, Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc Natl Acad Sci U S A* **109**, E690-697 (2012).
16. V. Pessino, Y. R. Citron, S. Feng, B. Huang, Covalent Protein Labeling by SpyTag-SpyCatcher in Fixed Cells for Super-Resolution Microscopy. *Chembiochem* **18**, 1492-1495 (2017).
17. A. H. Keeble *et al.*, Evolving Accelerated Amidation by SpyTag/SpyCatcher to Analyze Membrane Dynamics. *Angew Chem Int Ed Engl* **56**, 16521-16525 (2017).
18. H. J. Carlson, D. W. Cotton, R. E. Campbell, Circularly permuted monomeric red fluorescent proteins with new termini in the beta-sheet. *Protein Sci* **19**, 1490-1499 (2010).
19. H. B. Nguyen, L. W. Hung, T. O. Yeates, T. C. Terwilliger, G. S. Waldo, Split green fluorescent protein as a modular binding partner for protein crystallization. *Acta Crystallogr D Biol Crystallogr* **69**, 2513-2523 (2013).
20. J. Park *et al.*, A conserved juxtacrine signal regulates synaptic partner recognition in Caenorhabditis elegans. *Neural Dev* **6**, 28 (2011).
21. H. Lee, W. C. Oh, J. Seong, J. Kim, Advanced Fluorescence Protein-Based Synapse-Detectors. *Front Synaptic Neurosci* **8**, 16 (2016).
22. G. Jansen *et al.*, The complete family of genes encoding G proteins of Caenorhabditis elegans. *Nat Genet* **21**, 414-419 (1999).
23. C. Rogers *et al.*, Inhibition of Caenorhabditis elegans social feeding by FMRFamide-related peptide activation of NPR-1. *Nat Neurosci* **6**, 1178-1185 (2003).
24. T. K. Kerppola, Bimolecular fluorescence complementation: visualization of molecular interactions in living cells. *Methods Cell Biol* **85**, 431-470 (2008).

25. C. D. Hu, Y. Chinenov, T. K. Kerppola, Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol Cell* **9**, 789-798 (2002).
26. G. Veggiani *et al.*, Programmable polyproteams built using twin peptide superglues. *Proc Natl Acad Sci U S A* **113**, 1202-1207 (2016).
27. I. Levchenko, R. A. Grant, D. A. Wah, R. T. Sauer, T. A. Baker, Structure of a delivery protein for an AAA+ protease in complex with a peptide degradation tag. *Mol Cell* **12**, 365-372 (2003).
28. G. Guntas *et al.*, Engineering an improved light-induced dimer (iLID) for controlling the localization and activity of signaling proteins. *Proc Natl Acad Sci U S A* **112**, 112-117 (2015).
29. J. H. Choi *et al.*, Interregional synaptic maps among engram cells underlie memory formation. *Science* **360**, 430-435 (2018).
30. B. Wu, Y. Chen, J. D. Muller, Fluorescence fluctuation spectroscopy of mCherry in living cells. *Biophys J* **96**, 2391-2404 (2009).
31. A. Sanchez-Andres, Y. Chen, J. D. Muller, Molecular brightness determined from a generalized form of Mandel's Q-parameter. *Biophys J* **89**, 3531-3547 (2005).
32. P. J. Macdonald, Y. Chen, X. Wang, Y. Chen, J. D. Mueller, Brightness analysis by Z-scan fluorescence fluctuation spectroscopy for the study of protein interactions within living cells. *Biophys J* **99**, 979-988 (2010).
33. D. Yu *et al.*, An improved monomeric infrared fluorescent protein for neuronal and tumour brain imaging. *Nat Commun* **5**, 3626 (2014).
34. S. Brenner, The genetics of Caenorhabditis elegans. *Genetics* **77**, 71-94 (1974).
35. J. Park *et al.*, A conserved juxtacrine signal regulates synaptic partner recognition in Caenorhabditis elegans. *Neural Dev* **6**, 28 (2011).

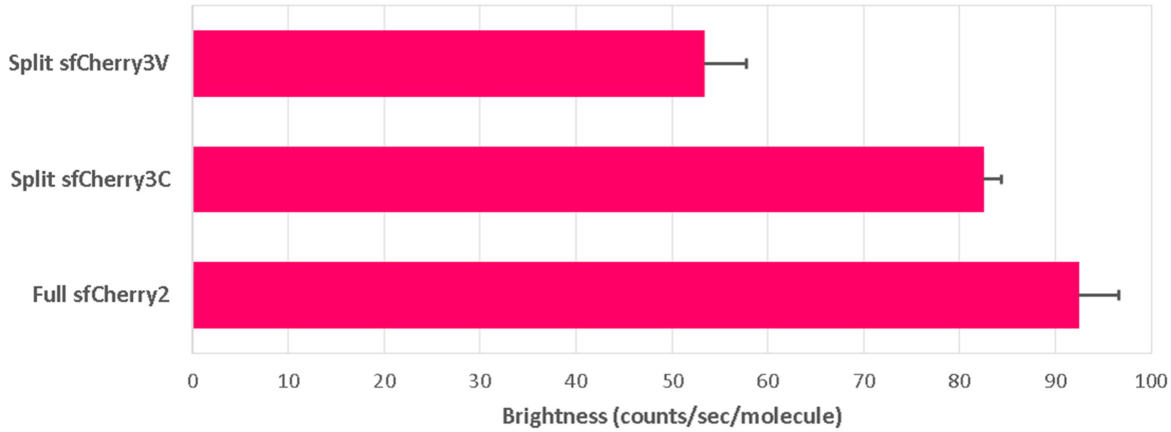
Supplementary Figures



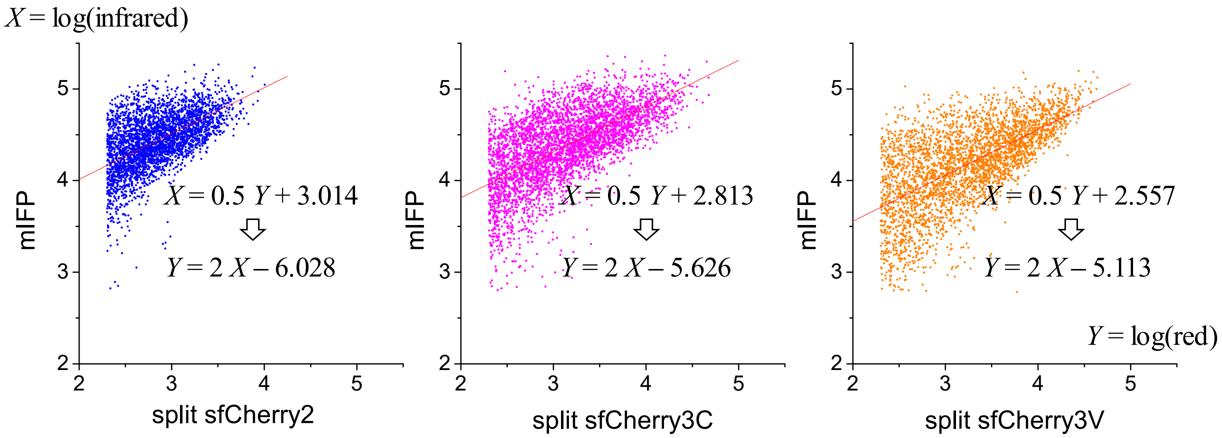
Supplementary Figure 1. Snapshots of fluorescence recovery after photobleaching (FRAP) experiment on endogenously mNeonGreen2$_{1-10/11}$ labeled Lamin A/C (nuclear envelope) in HEK 293T cells at different time points (before bleaching, 1st frame, 5 min and 10 min after bleaching). The yellow square marks the photobleaching area.



Supplementary Figure 2. Flow cytometry analysis of whole cell fluorescence in HEK 293T cells expressing different tandem tags (upper: SpyTag-sfCherry2$_{11}$-TagBFP, lower: sfCherry2$_{11}$- SpyTag-TagBFP) with different binders (left: sfCherry2$_{1-10}$ alone, middle: SpyCatcher-15aa-sfCherry2$_{1-10}$, right: SpyCatcher-6aa-sfCherry2$_{1-10}$). The pink dashed trend line has a slope of 2.



Supplementary Figure 3. Single-molecule brightness measurement of split sfCherry3V, split sfCherry3Cs and full-length sfCherry2 using fluorescence fluctuation spectroscopy by two-photon excitation.



Supplementary Figure 4. Least square linear fitting of flow cytometry points in Figure 3E with a fixed slope of 0.5. The *x*- and *y*-axes were flipped compared to main figures in order to properly handle the asymmetric errors in the two channels. Points with intensities within 3σ of the scattering background in either channel were excluded from the fitting. The intercept change in the fitted line (from 3.014 ± 0.004 to 2.813 ± 0.005 to 2.557 ± 0.006, ± standard error for least square fitting) was used to calculate the relative change in dissociation constant K_D' (see Supplementary Notes).

Supplementary Notes

A two-step model for $FP_{1\text{-}10/11}$ complementation

The overall complementation process of commonly used split fluorescent proteins is known to be irreversible [1, 2]. Here, we verified that in the case of the self-complementing split $mNG2_{1\text{-}10/11}$, the overall process is also irreversible by fluorescence recovery after photobleaching (FRAP) experiment. In HEK 293T cells with lamin A/C endogenously labeled by $mNG2_{11}$ and constitutively overexpressed $mNG2_{1\text{-}10}$ [3], bleaching a part of the nuclear lamina led to no observed fluorescence recovery in 10 min (Supplementary Figure 1, Supplementary Movie). Considering the fast maturation of mNeonGreen (< 10 min) [4], this lack of fluorescence recovery indicated that there was no exchange between the complemented $mNG2_{1\text{-}10}$ with free cytosolic $mNG2_{1\text{-}10}$.

On the other hand, a one-step, irreversible complementation would result in a direct proportional relationship between the complemented fluorescence signal and the fragment expression level. In this case, the FACS data points in Figure 1C (left) should follow a trend line with a slope of 1, whereas in reality, they fall closer to a trend line with a slope of 2. To explain this observation, we consider a two-step complementation model, in which the first step is a relatively fast reversible binding between the $FP_{1\text{-}10}$ and FP_{11} fragments, forming an intermediate $FP^{*}_{1\text{-}10/11}$ complex, followed by a second, slower step of irreversible folding and/or fluorophore maturation.

$$\mathrm{FP}_{1-10} + \mathrm{FP}_{11} \underset{\text{binding}}{\overset{K_D}{\rightleftharpoons}} \mathrm{FP}^{*}_{1-10/11} \xrightarrow[\text{folding/maturation}]{k_f} \mathrm{FP}_{1-10/11} \overset{k_{deg}}{\longrightarrow} \text{degradation}$$

At a steady state, the concentration of the complemented, functionally fluorescent species, $FP_{1\text{-}10/11}$, is related to the concentration of the intermediate complex, $FP^{*}_{1\text{-}10/11}$, by:

$$[\mathrm{FP}_{1-10/11}] = \frac{k_f}{k_{deg}}[\mathrm{FP}^{*}_{1-10/11}]$$

where k_f and k_{deg} are the folding/maturation rate and degradation rate, respectively. Generally, we would expect that the folding rate is much faster than the degradation rate, i.e. $k_f >> k_{deg}$ and hence $[\mathrm{FP}_{1-10/11}] >> [\mathrm{FP}^{*}_{1-10/11}]$. The concentration of $FP^{*}_{1\text{-}10/11}$ is further connected to the concentration of the free fragments as:

$$[\mathrm{FP}^{*}_{1-10/11}] = \frac{1}{K_D}[\mathrm{FP}_{1-10}][\mathrm{FP}_{11}]$$

where K_D is the dissociation constant for the initial binding step. In our experiment, FP_{11} and $FP_{1\text{-}10}$ are expressed from co-transfected plasmids. Figure 1E indicates that their expression levels in individual cells follow a proportional relationship:

$$[\mathrm{FP}_{1-10}] = a[\mathrm{FP}_{11}]$$

where a should be close to 1 considering that the same expression vectors are used for FP_{11} and $FP_{1\text{-}10}$. Taken together, we have:

$$[\mathrm{FP}_{1-10/11}] = \frac{1}{K_D{}'}[\mathrm{FP}_{11}]^2$$

where the effective dissociation constant

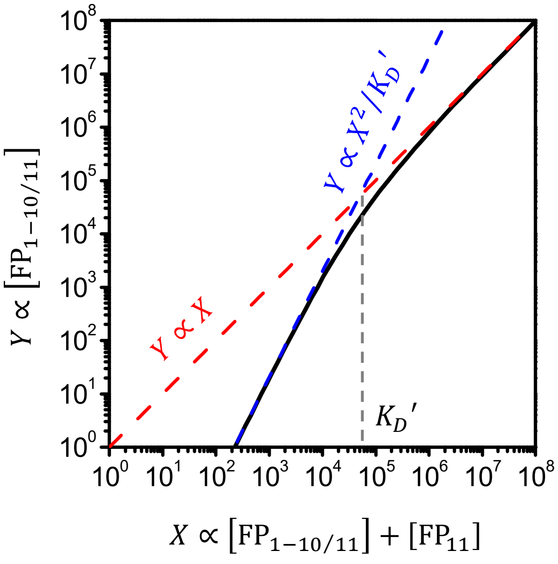
$$K_D{}' = \frac{k_{deg}}{k_f a} K_D$$

In our FACS experiment, the complemented fluorescence signal, Y, is proportional to $[\mathrm{FP}_{1-10/11}]$, whereas the expression marker (mIFP) signal, X, is actually proportional to the total concentration of complemented and uncomplemented FP_{11} fragment

$$X \propto [\mathrm{FP}_{1-10/11}] + [\mathrm{FP}^*_{1-10/11}] + [\mathrm{FP}_{11}] \approx [\mathrm{FP}_{1-10/11}] + [\mathrm{FP}_{11}] = [\mathrm{FP}_{1-10/11}] + \sqrt{K_D{}'[\mathrm{FP}_{1-10/11}]}$$

considering that $[\mathrm{FP}_{1-10/11}] >> [\mathrm{FP}^*_{1-10/11}]$.

The relationship between Y and X, on a log-log plot, follows the black curve on the simulated plot below:



At a high fragment expression level or with a small $K_D{}'$ (the case of $\mathrm{GFP}_{1\text{-}10/11}$), X is dominated by the first term of $[\mathrm{FP}_{1-10/11}]$, making the right end of the curve approaching a slope of 1. On the other hand, at a low fragment expression level or with a large $K_D{}'$ (the cases of $\mathrm{mNG2}_{1\text{-}10/11}$ and $\mathrm{sfCherry2}_{1\text{-}10/11}$), X is dominated by the second term of $[\mathrm{FP}_{11}] = \sqrt{K_D{}'[\mathrm{FP}_{1-10/11}]}$, making the left end of the curve approaching a slope of 2.

In this latter regime, only a small fraction of the FP_{11} fragment is complemented with the $\mathrm{FP}_{1\text{-}10}$ fragment to reconstitute a functional fluorescent protein. Consequently, the apparent brightness of the FP_{11}-labeled protein is lower than protein labeled by the full length FP. The complementation efficiency, defined as the fraction of FP_{11} complemented with $\mathrm{FP}_{1\text{-}10}$, is directly proportional to the concentration of $\mathrm{FP}_{1\text{-}10}$:

$$\text{complementation efficiency} = \frac{[\mathrm{FP}_{1-10/11}]}{[\mathrm{FP}_{1-10/11}] + [\mathrm{FP}^*_{1-10/11}] + [\mathrm{FP}_{11}]} \approx \frac{[\mathrm{FP}_{1-10/11}]}{[\mathrm{FP}_{11}]} = \frac{[\mathrm{FP}_{1-10}]}{K_D{}''}$$

where $K_D" = aK_D'$ (i.e. effective dissociation constant with respect to $FP_{1\text{-}10}$ instead of FP_{11}). This relationship holds true both for our FACS experiment and for any general applications of FP_{11} labeling. In other words, the complementation efficiency improves with higher expression level of the $FP_{1\text{-}10}$ fragment. Therefore, overexpression of $mNG2_{1\text{-}10}$ and $sfCherry2_{1\text{-}10}$ improves the apparent brightness of the FP_{11} labeled proteins. This improvement saturates when the concentration of $FP_{1\text{-}10}$ is approaching K_D''. Therefore, a trade-off must be considered between the diminishing brightness improvement and the potential cytotoxicity of high $FP_{1\text{-}10}$ expression.

Obviously, a better approach to improve the apparent brightness / complementation efficiency is to reduce K_D' (equivalent to reducing K_D''). This is exactly what we tried to achieve by engineering sfCherry3C and sfCherry3V. In these cases, the improvement of K_D' can be measured by fitting the FACS data with a line at a slope of 2 and then measure the change in the intercept (fig. S4).

From sfCherry2 to sfCherry3C, the K_D' decreased by:

$$\frac{K_D'(\text{sfCherry2})}{K_D'(\text{sfCherry3C})} = 10^{6.028-5.628} = 2.5$$

From sfCherry2 to sfCherry3V, the K_D' decreased by:

$$\frac{K_D'(\text{sfCherry2})}{K_D'(\text{sfCherry3V})} = 10^{6.028-5.113} = 8.2$$

References

1. Cabantous, S. and G.S. Waldo, *In vivo and in vitro protein solubility assays using split GFP.* Nat Methods, 2006. **3**(10): p. 845-54.
2. Kerppola, T.K., *Visualization of molecular interactions by fluorescence complementation.* Nat Rev Mol Cell Biol, 2006. **7**(6): p. 449-56.
3. Feng, S., et al., *Improved split fluorescent proteins for endogenous protein labeling.* Nat Commun, 2017. **8**(1): p. 370.
4. Shaner, N.C., et al., *A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum.* Nat Methods, 2013. **10**(5): p. 407-9.

Supplementary Tables

Supplementary Table 1: Sequence of sfCherry3C$_{1-10}$, sfCherry3V$_{1-10}$, SpyCatcher 002 and SpyTag 002.

	DNA Sequence
sfCherry3C$_{1-10}$	GAGGAGGACAACATGGCCATCATCAAGGAGTTCATGAGATTCAAGGTGCACATGGAGGGCAGCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCACCCCTACGAGGGCACCCAGACCGCCAGGCTGAAGGTGACCAAGGGCGACCCCCTGCCCTTCGCCTGGGACATCCTGAGCCCCCAGTTCATGTACGGCAGCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACCTGAAGCTGAGCTTCCCCGAGGGCTTCACCTGGGAGAGAGTGATGAACTTCGAGGACGGCGGCGTGGTGGCCGTGACCCAGGACAGCAGCCTGCAGGACGGCCAGTTCATCTACAAGGTGAAGCTGCTGGGCATCAACTTCCCCAGCGACGGCCCCGTGATGCAGAAGAAGACCATGGGCTGGGAGGCCAGCACCGAGAGAATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAACCAGAGACTGAAGCTGAAGGACGGCGGTCACTACGACGCCGAGGTGAAGACCACCTACAGGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACGACGTGGACATCAAGCTGGACATCACCAGCCACAACGAGGAC
sfCherry3V$_{1-10}$	GAGGAGGACAACATGGCCATCATCAAGGAGTTCATGAGATTCAAGGTGCACATGGAGGGCAGCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCACCCCTACGAGGGCACCCAGACCGCCAGGCTGAAGGTGACCAAGGGCGACCCCCTGCCCTTCGCCTGGGACATCCTGAGCCCCCAGTTCATGTACGGCAGCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTGGAAGCTGAGCTTCCCCGAGGGCTTCACCTGGGAGAGAGTGATGAACTTCGAGGACGGCGGCGTGGTGGCCGTGACCCAGGACAGCAGCCTGCAGGACGGCCAGTTCATCTACAAGGTGAAGCTGCTGGGCATCAACTTCCCCAGCGACGGCCCCGTGATGCAGAAGAAGACCATGGGCTGGGAGGCCAGCACCGAGAGAATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAACCAGAGACTGAAGCTGAAGGACGGCGGTCACTACGACGCCGAGGTGAAGACCACCTACAGGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACGACGTGGACATCAAGCTGGACATCACCAGCCACAACGAGGAC
SpyCatcher 002	GGCGCCATGGTAACCACCTTATCAGGTTTATCAGGTGAGCAAGGTCCGTCCGGTGATATGACAACTGAAGAAGATAGTGCTACCCATATTAAATTCTCAAAACGTGATGAGGACGGCCGTGAGTTAGCTGGTGCAACTATGGAGTTGCGTGATTCATCTGGTAAAACTATTAGTACATGGATTTCAGATGGACATGTGAAGGATTTCTACCTGTATCCAGGAAAATATACATTTGTCGAAACCGCAGCACCAGACGGTTATGAGGTAGCAACTGCTATTACCTTTACAGTTAATGAGCAAGGTCAGGTTACTGTAAATGGCGAAGCAACTAAAGGTGACGCTCATACT
SpyTag 002	GTGCCTACTATCGTGATGGTGGACGCCTACAAGCGTTACAAG

Supplementary Table 2: DNA sequence used for sgRNA in vitro transcription

Target Gene	Target Term	Sequence of DNA oligo for sgRNA synthesis
LMNA	N	TAATACGACTCACTATAGGCCATGGAGACCCCGTCCCAGGTTTAAGAGCTATGCTGGAA
CLTA	N	TAATACGACTCACTATAGGGCCATGGCGGGCAACTGAAGTTTAAGAGCTATGCTGGAA
RAB11A	N	TAATACGACTCACTATAGGGTAGTCGTACTCGTCGTCGGTTTAAGAGCTATGCTGGAA
HP1b	N	TAATACGACTCACTATAGGAAAGCTGGCGGGCACTATGGTTTAAGAGCTATGCTGGAA
SEC61b	N	TAATACGACTCACTATAGGCTTGTCTCCCTCTACAGCCGTTTAAGAGCTATGCTGGAA
ARL6IP1	N	TAATACGACTCACTATAGGATCCCCGAGACGATGGCGGGTTTAAGAGCTATGCTGGAA

Supplementary Table 3: Oligo-nucleotide donor DNA sequence

Target Gene-FP$_{11}$ Tag	DNA sequence
LMNA-sfCherry2$_{11}$	TCCTTCGACCCGAGCCCCGCGCCCTTTCCGGGACCCCTGCCCCGCGGGCAGCGCTGCCAACCTGCCGGCCATGTACACCATCGTGGAGCAGTACGAGAGAGCCGAGGCCAGACACAGCACCGGTGGCGGCGAGACCCCGTCCCAGCGGCGCGCCACCCGCAGCGGGGCGCAGGCCAGCTCCACTCCGCTGTCGCCCACCC
CLTA-sfCherry2$_{11}$	CGGGCGTGGTGTCGGTGGGTCGGTTGGTTTTTGTCTCACCGTTGGTGTCCGTGCCGTTCAGTTGCCCGCCATGTACACCATCGTGGAGCAGTACGAGAGAGCCGAGGCCAGACACAGCACCGGAGGTGGCATGGCTGAGCTGGATCCGTTCGGCGCCCCTGCCGGCGCCCCTGGCGGTCCCGCGCTGGGGAACGGAGTGG
RAB11A-sfCherry2$_{11}$	CCCTGCAGCGACGCCCCCTGGTCCCACAGATACCACTGCTGCTCCCGCCCTTTCGCTCCTCGGCCGCGCAATGTACACCATCGTGGAGCAGTACGAGAGAGCCGAGGCCAGACACAGCACCGGAGGTGGCGGCACCCGCGACGACGAGTACGACTACCTCTTTAAAGGTGAGGCCATGGGCTCTCGCACTCTACACAGTC
HP1b-sfCherry2$_{11}$	CTAATGCCCTTTTTATTTCATTTATCATTTTAGCAGCGTCACCCTTTACACCAGAAAGCTGGCGGGCACTATGTACACCATCGTGGAGCAGTACGAGAGAGCCGAGGCCAGACACAGCACCGGTGGCGGCGGGAAAAAACAAAACAAGAAGAAAGTGGAGGAGGTGCTAGAAGAGGAGGAAGAGGAATATGTGGTGGAAA
SEC61b-sfCherry2$_{11}$	GTGTCTAGGCCGGGGTTCTGGGGCAGGCCTGCCGCGCTCACCCGTCTGTCTGCTTGTCTCCCTCTACAGTACACCATCGTGGAGCAGTACGAGAGAGCCGAGGCCAGACACAGCACCGGTGGCGGCCCTGGTCCGACCCCCAGTGGCACTAACGTGGGATCCTCAGGGCGCTCTCCCAGCAAAGCAGTGGCCGCCCGGGC
ARL6IP1-sfCherry2$_{11}$	GCGGGTTTCGGTTGGAGGACTCGTTGGGGAGGTGGCCTGCGCTTGTAGAGACTGCATCCCCGAGACGATGTACACCATCGTGGAGCAGTACGAGAGAGCCGAGGCCAGACACAGCACCGGTGGCGGCGCGGAGGGAGATAATCGCAGCACCAACCTGCTGGTGAGTCCTGGCTGCCTGTCCCCCGGGAGCCGAGCGA

sfCherry2$_{11}$ sequence
Linker sequence
Coding region sequence