

## Supplementary Information for

### Improved Split Fluorescent Proteins for Endogenous Protein Labeling

Siyu Feng<sup>1</sup>, Sayaka Sekine<sup>2</sup>, Veronica Pessino<sup>3</sup>, Han Li<sup>4</sup>, Manuel D. Leonetti<sup>4</sup>, Bo Huang<sup>2, 5, 6\*</sup>

<sup>1</sup>The UC Berkeley-UCSF Graduate Program in Bioengineering, San Francisco, CA 94143 USA;

<sup>2</sup>Department of Pharmaceutical Chemistry, University of California in San Francisco, San Francisco, CA 94143, USA;

<sup>3</sup>Graduate Program of Biophysics, University of California, San Francisco, San Francisco, CA 94143, USA;

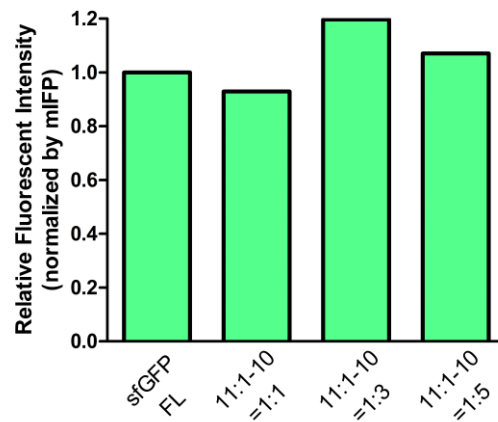
<sup>4</sup>Department of Cellular and Molecular Pharmacology, and Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, CA 94143, USA;

<sup>5</sup>Department Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94143, USA.

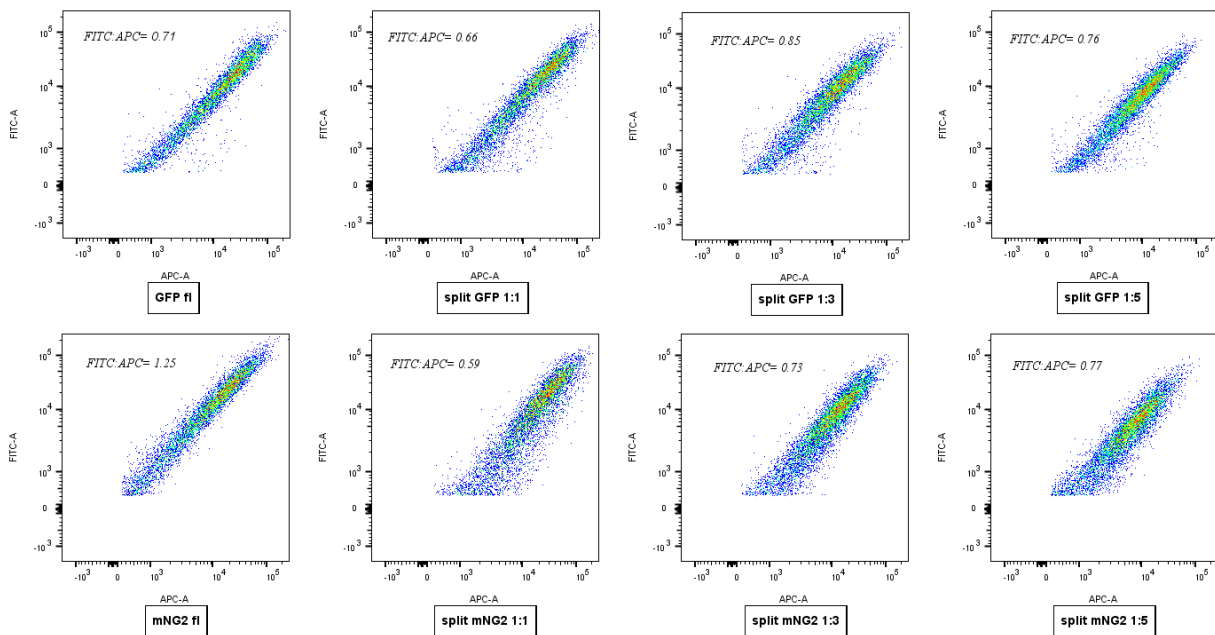
<sup>6</sup>Chan Zuckerberg Biohub, San Francisco, CA 94158, USA

\* Correspondence should be addressed to [bo.huang@ucsf.edu](mailto:bo.huang@ucsf.edu).

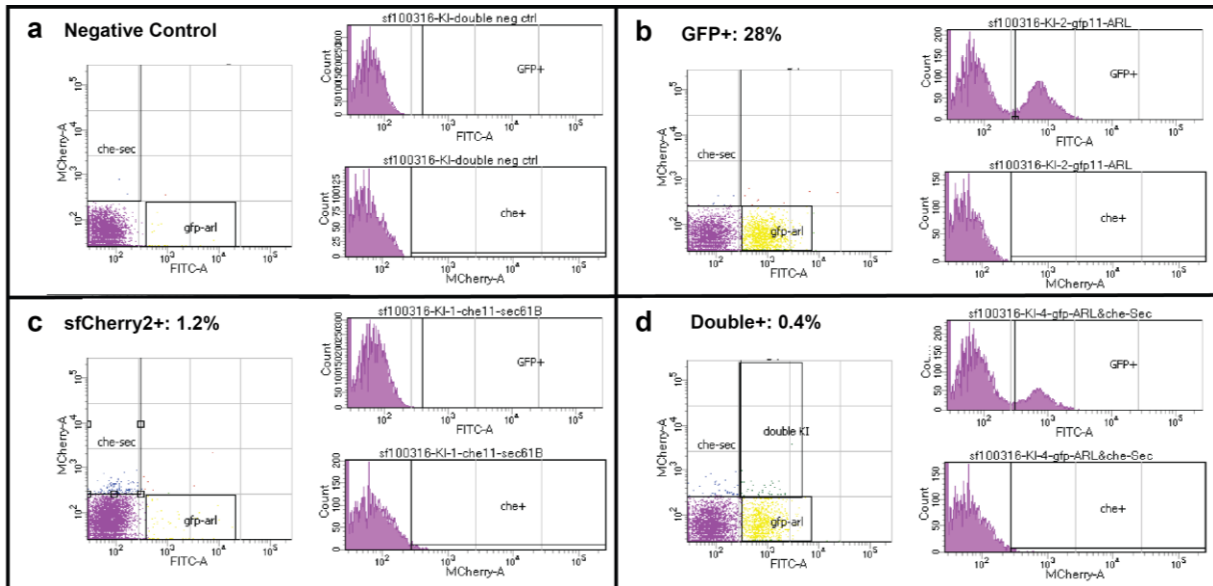
a



b



**Supplementary Figure 1 (a)** Whole cell fluorescence intensity of full length GFP and GFP<sub>11</sub>-CLTA/GFP<sub>1-10</sub>, measured by flow cytometry and normalized for expression level by mIFP fluorescence signal. Number of cells > 6000. **(b)** FACS raw data of bar graphs in Fig. 2e and Supplementary Fig. 1a. The X-axis is mIFP fluorescence intensity (ex=633 nm, em=710/50 nm) and the Y-axis is green fluorescence intensity (ex=488 nm, em=530/30 nm). The “FITC:APC” value is the population mean of green fluorescence normalized by mIFP fluorescence.



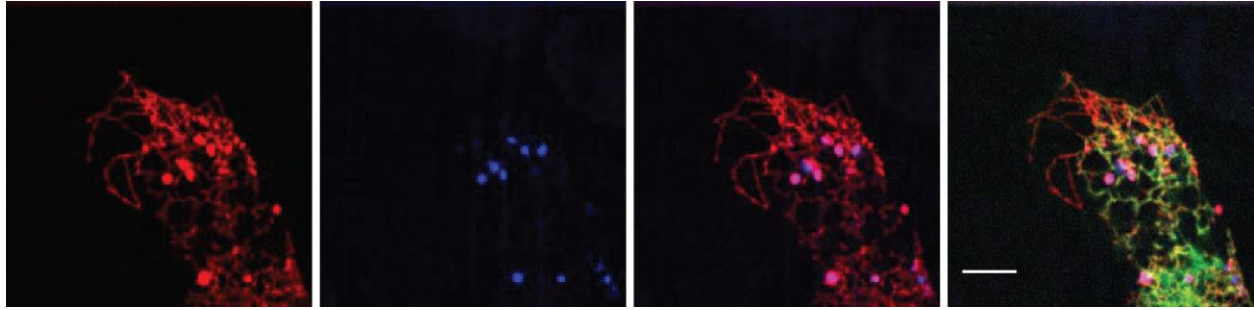
**Supplementary Figure 2.** FACS results of single knock-in or simultaneous double knock-in using GFP<sub>11</sub> or/and sfCherry2<sub>11</sub>. **(a)** Background fluorescence of HEK 293T cells stably expressing both GFP<sub>1-10</sub> and sfCherry2<sub>1-10</sub>. **(b)** Single knock-in of GFP<sub>11</sub> into ARL6IP1 with an efficiency of ~28% (GFP positive). **(c)** Single knock-in of sfCherry2<sub>11</sub> into Sec61B with an efficiency of ~1.2% (sfCherry2 positive). **(d)** Simultaneous double knock-in of GFP<sub>11</sub> and sfCherry2<sub>11</sub> into ARL6IP1 and Sec61B respectively, with an efficiency of ~0.4% (double positive).

ARL6IP1-sfCherry2<sub>11</sub>

Lysotracker staining

2-color merge

3-color merge (Sec61B)



**Supplementary Figure 3.** Lysotracker live-cell staining visualize the colocalization of lysosomes and sfCherry2 puncta in the same ARL6IP1-sfCherry2<sub>11</sub>/Sec61B-GFP<sub>11</sub> knock-in HEK 293T cell shown in Figures 6a and 6b. Scale bar: 5  $\mu$ m.

**Supplementary Table 1: Sequence of mNG2<sub>1-10</sub>/mNG2<sub>11</sub>, sfCherry2<sub>1-10</sub>/sfCherry2<sub>11</sub>, PAsfCherry2<sub>1-10</sub> and 32-residue spacer in the screening construct.**

	DNA Sequence
<b>mNG2<sub>1-10</sub></b>	ATGGTGAGCAAGGGTGAGGAGGATAACATGGCCTCTCTCCCAGCGACTCATGAGTTACA CATCTTTGGCTCCATCAACGGTGTGGACTTTGACATGGTGGGTCAGGGTACCGGCAATC CAAATGATGGTTATGAGGAGTTAAACCTGAAGTCCACCAAGGGTGACCTCCAGTTCTCCC CCTGGATTCTGGTCCCTCATATCGGGTATGGCTTCCATCAGTACCTGCCCTACCCTGACG GGATGTCGCCTTTCCAGGCCGCCATGGTAGATGGCTCCGGATACCAAGTCCATCGCACA ATGCAGTTTGAAGATGGTGCCTCCCTTACTGTTAACTACCGCTACACCTACGAGGGAAGC CACATCAAAGGAGAGGCCAGGTGATGGGGACTGGTTTCCCTGCTGACGGTCTGTGA TGACCAACACGCTGACCGCTGCGGACTGGTGCATGTCGAAGAAGACTTACCCCAACGA CAAAACCATCATCAGTACCTTTAAGTGGAGTTACACCACTGTAAATGGCAAACGCTACCG GAGCACTGCGCGGACCACCTACACCTTGGCAAGCCAATGGCGGCTAACTATCTGAAGA ACCAGCCGATGTACGTGTTCCGTAAGACGGAGCTCAAGCACTCCATG
<b>mNG2<sub>11</sub></b>	ACCGAGCTCAACTTCAAGGAGTGGCAAAGGCCTTACCGATATGATG
<b>sfCherry2<sub>1-10</sub></b>	ATGGAGGAGGACAACATGGCCATCATCAAGGAGTTCATGAGATTCAAGGTGCACATGGA GGGCAGCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCACCCCTA CGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGCGGCCCTGCCCTTCGC CTGGGACATCCTGAGCCCCAGTTCATGTACGGCAGCAAGGCCTACGTGAAGCACCCC GCCGACATCCCCGACTACCTGAAGCTGAGCTTCCCCGAGGGCTTACCTGGGAGAGAG TGATGAACTTCGAGGACGGCGCGTGGTGACCGTGACCCAGGACAGCAGCCTGCAGG ACGGCCAGTTCATCTACAAGGTGAAGCTGCTGGGCATCAACTTCCCCAGCGACGGCCC CGTGATGCAGAAGAAGACCATGGGCTGGGAGGCCAGCACCGAGAGAATGTACCCCGAG GACGGCGCCCTGAAGGGCGAGATCAACCAGAGACTGAAGCTGAAGGACGGCGGCCAC TACGACGCCGAGGTGAAGACCACCTACAAGGCCAAGAAGCCCCTGCAGCTGCCCGGC GCCTACAACGTGGACATCAAGCTGGACATCACCAAGCCACAACGAGGAC
<b>sfCherry2<sub>11</sub></b>	TACACCATCGTGGAGCAGTACGAGAGAGCCGAGGCCAGACACAGCACC
<b>PAsfCherry2<sub>1-10</sub></b>	ATGGAGGAGGATAACATGGCCATCATTAAAGGAGTTCATGCGTTCAAGGTGCACATGGAG GGGTCCGTGAACGGCCACGTGTTTCGAGATCGAGGGCGAGGGCGAGGGCCACCCCTAC GAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTACC TGGGACATCCTGTCCCCTCAATTCATGTACGGCTCCAATGCCTACGTGAAGCACCCCGC CGACATCCCCGACTACTTTAAGCTGTCCTTCCCCGAGGGCTTACCTGGGAGCGCGTGA TGAAATTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACG GTCAGTTCATCTACAAGGTGAAGCTGCTGGGCATCAACTTCCCCCTCCGACGGCCCCGTA ATGCAGAAGAAGACCATGGGCTGGGAGGCCCTCACCGAGCGGATGTACCCCGAGGACG GCGCCCTGAAGGGCGAGGTCAACCCGAGAGTGAAGCTGAAGGACGGCGGCCACTACG ACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCCTGCAGCTGCCCGGCGCCTA CAACGTCGACCGCAAGTTGGACATCACCTCACACAACGAGGAC
<b>32-residue spacer</b>	GACGTTGGTGGTGGCGGATCAGAAGGAGGCGGTAGCGGGGGCCCTGGTTCCGGGAGGG GAAGGTTCTGCTGGGGGAGGGAGCGCTGGCGGGGGTCT

**Supplementary Table 2: sgRNA sequence**

Target Gene	Target Term	Sequence of DNA oligo for sgRNA synthesis
SEC61B	N	TAATACGACTCACTATAGGCTTGTCTCCCTCTACAGCCGTTTAAGAGCTATGCTGGAA
LMNA	N	TAATACGACTCACTATAGGCCATGGAGACCCCGTCCCAGGTTTAAGAGCTATGCTGGAA
CLTA	N	TAATACGACTCACTATAGGGCCATGGCGGGCAACTGAAGTTTAAGAGCTATGCTGGAA
ARL6IP1	N	TAATACGACTCACTATAGGATCCCCGAGACGATGGCGGGTTTAAGAGCTATGCTGGAA
RAB11A	N	TAATACGACTCACTATAGGGTAGTCGTA CTGTCGTCGGTTTAAGAGCTATGCTGGAA
SPTLC1	C	TAATACGACTCACTATAGGGACTCTGCCTAGAGCAGGAGTTTAAGAGCTATGCTGGAA

**Supplementary Table 3: Oligo-nucleotide donor DNA sequence**

Target Gene-FP <sub>11</sub> Tag	DNA sequence
LMNA-mNG2 <sub>11</sub>	TCCTTCGACCCGAGCCCCGCGCCCTTTCCGGGACCCCTGCCCCGC GGGCAGCGCTGCCAACCTGCCGGCCATG <b>ACCGAGCTCAACTTCAAG</b> <b>GAGTGGCAAAGGCCTTTACCGATATGATGGAGGTGGC</b> ATGGAGAC CCCGTCCCAGCGGCGCGCCACCCGCAGCGGGGCGCAGGCCAGCT CCTACTCCGCTGTCGCCCA
CLTA-mNG2 <sub>11</sub>	CGGGCGTGGTGTCTGGTGGGTCGGTTGGTTTTTGTCTCACCGTTGGT GTCCGTGCCGTTTCAGTTGCCCGCCATG <b>ACCGAGCTCAACTTCAAGG</b> <b>AGTGGCAAAGGCCTTTACCGATATGATGGAGGTGGC</b> ATGGCTGAG CTGGATCCGTTCCGGCGCCCTGCCGGCGCCCTGGCGGTCCCGCG CTGGGGAACGGAGTGG
RAB11A-mNG2 <sub>11</sub>	TGCAGCGACGCCCTGGTCCACAGATACCACTGCTGCTCCCGCCCTTTC GCTCCTCGGCCGCGCAATG <b>ACCGAGCTCAACTTCAAGGAGTGGCAAAGGC</b> <b>CTTTACCGATATGATGGTTCTGGC</b> GGCGGCACCCGCGACGACGAGTACGA CTACCTCTTTAAAGGTGAGGCCATGGGCTCTCGCACTCTACACAGTC
SPTLC1-mNG2 <sub>11</sub>	GAACAAACAGAGGAAGAACTGGAGAGAGCTGCGTCCACCATCAAGGAGGTA GCCCAGGCAGTTCTGCTCGGT <b>GGCTCTGGCACCGAGCTCAACTTCAAGGAG</b> <b>TGGCAAAGGCCTTTACCGATATGATG</b> TAGGCAGAGTCCCGGGACCATGGCC TCCTGCCACACAACACGCAGAGAGGACTCAAGACTCCCGCTGGCCA
LMNA-sfCherry2 <sub>11</sub>	TCCTTCGACCCGAGCCCCGCGCCCTTTCCGGGACCCCTGCCCCGC GGGCAGCGCTGCCAACCTGCCGGCCATG <b>TACACCATCGTGGAGCAG</b> <b>TACGAGAGAGCCGAGGCCAGACACAGCACCCGGTGGCGGC</b> GAGACC CCGTCCCAGCGGCGCGCCACCCGCAGCGGGGCGCAGGCCAGCTC CACTCCGCTGTCGCCACCC
ARL6IP1-sfCherry2 <sub>11</sub>	GCGGGTTTCGGTTGGAGGACTCGTTGGGGAGGTGGCCTGCGCTTG TAGAGACTGCATCCCCGAGACGATG <b>TACACCATCGTGGAGCAGTACC</b> <b>AGAGAGCCGAGGCCAGACACAGCACCCGGTGGCGGC</b> GCGGAGGGA GATAATCGCAGCACCAACCTGCTGGTGAGTCCTGGCTGCCTGTCCC CCGGGAGCCGAGCGA
SEC61B- sfCherry2 <sub>11</sub>	GTGTCTAGGCCGGGGTTCTGGGGCAGGCCTGCCGCGCTCACCCGT CTGTCTGCTTGTCTCCCTCTACAG <b>TACACCATCGTGGAGCAGTACGA</b> <b>GAGAGCCGAGGCCAGACACAGCACCCGGTGGCGGC</b> CCTGGTCCGAC CCCCAGTGGCACTAACGTGGGATCCTCAGGGCGCTCTCCAGCAA GCAGTGGCCGCCCGGGC
SEC61B-GFP <sub>11</sub>	GTGTCTAGGCCGGGGTTCTGGGGCAGGCCTGCCGCGCTCACCCGT CTGTCTGCTTGTCTCCCTCTACAG <b>CGTGACCACATGGTCTTTCATGA</b> <b>GTATGTAAATGCTGCTGGGATTACAGGTGGCGGC</b> CCTGGTCCGACCC CCAGTGGCACTAACGTGGGATCCTCAGGGCGCTCTCCAGCAAAGC AGTGGCCGCCCGGGC

sfCherry2 <sub>11</sub> -GFP <sub>11</sub> - CLTA	AGCTGATAATACGACTCACTATAGGGCTCTGCAACACCGCCTAGACC GACCGGATACACGGGTAGGGCTTCCGCTTTACCCGTCTCCCTCCTGG CGCTTGTCTCCTCTCCCAGTCGGCACCACAGCGGTGGCTGCCGGG CGTGGTGTTCGGTGGGTTCGGTTGGTTTTTGTCTCACCGTTGGTGTCC GTGCCGTTTCAGTTGCCCGCCATGTACACCATCGTGGAGCAGTACGAG AGAGCCGAGGCCAGACACAGCACCGGTGGCTCTGGAAGTTCAGGT GGAGGCTCGCGTGACCACATGGTCCTTCATGAGTATGTAAATGCTGC TGGGATTACAGGAGGCGGTATGGCTGAGCTGGATCCGTTCCGGCGCC CCTGCCGGCGCCCCTGGCGGTCCCGCGCTGGGGAACGGAGTGGC CGGCGCCGGCGAAGAAGACCCGGCTGCGGCCTTCTTGGCGCAGCA AGAGAGCGAGATTGCGGGCATCGAGAACGACGAGGCCT
--	---

mNG2<sub>11</sub> Sequence

sfCherry2<sub>11</sub> sequence

GFP<sub>11</sub> sequence

Linker sequence

Coding region sequence