



# Rapid directed molecular evolution of fluorescent proteins in mammalian cells

Siranush Babakhanova<sup>1,2,3,4</sup> | Erica E. Jung<sup>5</sup> | Kazuhiko Namikawa<sup>6</sup> | Hanbin Zhang<sup>7,8,9</sup> | Yangdong Wang<sup>7,8,9</sup> | Oksana M. Subach<sup>10</sup> | Dmitry A. Korzhenevskiy<sup>10</sup> | Tatiana V. Rakitina<sup>10,11</sup> | Xian Xiao<sup>7,8,9</sup> | Wenjing Wang<sup>7,8,9</sup> | Jing Shi<sup>7,8,9</sup> | Mikhail Drobizhev<sup>12</sup> | Demian Park<sup>1,4</sup> | Lea Eisenhard<sup>6</sup> | Hongyun Tang<sup>7,8,9</sup> | Reinhard W. Köster<sup>6</sup> | Fedor V. Subach<sup>10</sup>  | Edward S. Boyden<sup>1,4,13,14,15,16</sup> | Kiryl D. Piatkevich<sup>7,8,9</sup> 

<sup>1</sup>Media Arts and Sciences, Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, USA

<sup>2</sup>Department of Physics, MIT, Cambridge, Massachusetts, USA

<sup>3</sup>Department of Electrical Engineering and Computer Science, MIT, Cambridge, Massachusetts, USA

<sup>4</sup>MIT McGovern Institute for Brain Research, MIT, Cambridge, Massachusetts, USA

<sup>5</sup>Department of Mechanical and Industrial Engineering, The University of Illinois at Chicago, Chicago, Illinois, USA

<sup>6</sup>Division of Cellular and Molecular Neurobiology, Zoological Institute, Technische Universität Braunschweig, Braunschweig, Germany

<sup>7</sup>School of Life Sciences, Westlake University, Zhejiang, Hangzhou, China

<sup>8</sup>Westlake Laboratory of Life Sciences and Biomedicine, Zhejiang, Hangzhou, China

<sup>9</sup>Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Zhejiang, Hangzhou, China

<sup>10</sup>National Research Center “Kurchatov Institute”, Moscow, Russian Federation

<sup>11</sup>Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, RAS, Moscow, Russian Federation

<sup>12</sup>Department of Microbiology and Immunology, Montana State University, Bozeman, Montana, USA

<sup>13</sup>Department of Biological Engineering, MIT, Cambridge, Massachusetts, USA

<sup>14</sup>Department of Brain and Cognitive Sciences, MIT, Cambridge, Massachusetts, USA

<sup>15</sup>Howard Hughes Medical Institute, Cambridge, Massachusetts, USA

<sup>16</sup>Koch Institute, MIT, Cambridge, Massachusetts, USA

## Correspondence

Edward S. Boyden, McGovern Institute and HHMI, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA.  
Email: edboyden@mit.edu

Kiryl D. Piatkevich, School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China.  
Email: kiryl.piatkevich@westlake.edu.cn

## Abstract

In vivo imaging of model organisms is heavily reliant on fluorescent proteins with high intracellular brightness. Here we describe a practical method for rapid optimization of fluorescent proteins via directed molecular evolution in cultured mammalian cells. Using this method, we were able to perform screening of large gene libraries containing up to  $2 \times 10^7$  independent random genes of fluorescent proteins expressed in HEK cells, completing one iteration of directed evolution in a course of 8 days. We employed this approach to develop a set of green and near-infrared fluorescent proteins with enhanced

Siranush Babakhanova, Erica E. Jung, Kazuhiko Namikawa, and Hanbin Zhang contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Protein Science* published by Wiley Periodicals LLC on behalf of The Protein Society.

### Funding information

National Natural Science Foundation of China, Grant/Award Number: 32050410298; NIH BRAIN, Grant/Award Number: U24NS109107; National Science Foundation, Grant/Award Number: 1734870; National Institutes of Health, Grant/Award Numbers: R01GM104948, R43MH109332, R01MH114031, R01DA045549, R01MH12297101, R01DA029639; German Research Foundation, Grant/Award Number: 241961032; Russian Science Foundation, Grant/Award Number: 17-14-01256; RFBR, Grant/Award Number: 19-04-00395; National Research Center Kurchatov Institute, Grant/Award Number: 2752; MRIC Funding, Grant/Award Number: 103536022023; 2020 BBRF Young Investigator Grant, Grant/Award Number: 28961

intracellular brightness. The developed near-infrared fluorescent proteins demonstrated high performance for fluorescent labeling of neurons in culture and in vivo in model organisms such as *Caenorhabditis elegans*, *Drosophila*, zebrafish, and mice. Spectral properties of the optimized near-infrared fluorescent proteins enabled crosstalk-free multicolor imaging in combination with common green and red fluorescent proteins, as well as dual-color near-infrared fluorescence imaging. The described method has a great potential to be adopted by protein engineers due to its simplicity and practicality. We also believe that the new enhanced fluorescent proteins will find wide application for in vivo multicolor imaging of small model organisms.

### KEYWORDS

directed molecular evolution, fluorescent proteins, in vivo fluorescence imaging, neuroimaging, optogenetics

## 1 | INTRODUCTION

Fluorescent proteins (FPs) became indispensable tools for in vivo imaging of cellular and subcellular structures in model organisms.<sup>1</sup> Since the cloning of the first green FP from jellyfish *Aequorea victoria* in 1992,<sup>2</sup> a myriad of chromoproteins with various spectral and biochemical properties have been cloned from diverse natural sources such as corals,<sup>3</sup> fish,<sup>4</sup> plants,<sup>5</sup> soil bacteria,<sup>6</sup> and cyanobacteria.<sup>7</sup> However, all naturally occurring chromoproteins have to be modified, optimized, or even reengineered in order to be utilized for fluorescence imaging in vivo.<sup>8–12</sup> Among all biochemical characteristics, intracellular brightness is one of the most crucial properties for *in vivo* applications that protein engineers and developers choose to optimize before everything else.<sup>13,14</sup> FPs are usually optimized via directed molecular evolution by iteratively generating and screening large gene libraries in bacterial cells.<sup>13–15</sup> However, high molecular brightness, commonly screened for in bacterial cells, does not always correspond to high intracellular brightness when expressed in cultured mammalian cells<sup>8,9,16</sup> or in vivo.<sup>16–18</sup> This phenomenon is particularly well documented for bacteriophytochrome-based FPs,<sup>16,19,20</sup> whose fluorescence relies on the incorporation of the chromophore biliverdin (BV) from the bulk.<sup>21</sup> Ideally, the development of functional proteins for *in vivo* imaging should be performed in an environment physiologically relevant to the final hosts to ensure proper protein folding, localization, and posttranslational modification.<sup>22,23</sup> Yeasts, although eukaryotic cells, which provide convenience of large gene library expression like bacteria, may not serve as an ideal host system

for FPs as it was shown that brightness in yeast cells is not necessarily retained in mammalian cells.<sup>24</sup> In this regard, vertebrate cell lines represent a promising expression host for directed molecular evolution of FPs. Indeed, several studies demonstrated the possibility to evolve FPs in cultured chicken<sup>25</sup> or mammalian cells.<sup>26–28</sup> However, the proposed methods did not find wide adaptation among protein engineers due to several limitations and drawbacks. First, all previously developed methods for directed molecular evolution in cultured cells from vertebrates involve establishing cell lines stably maintaining target genes, in a way where any given cell expresses ideally no more than one copy of a target gene.<sup>26–28</sup> However, both commonly used single gene copy delivery methods, such as electroporation and retroviral transduction, and establishing stable cell lines, are time-consuming and laborious, and complicated by apoptosis, low efficiency of stable gene integration, and long cell doubling time. Second, in situ diversification of target genes using a cytidine deaminase<sup>25,26,29</sup> or viral replication<sup>30</sup> has a low mutation rate of only 1–3 mutations per kilobase pair in comparison to 9–16 mutations per kilobase pair for regular error-prone PCR typically used for FP development. Higher mutation rates can be achieved with CRISPR/Cas9 editing technology<sup>28</sup> or via in vitro mutagenesis,<sup>27</sup> but comes with a limited library size of 10<sup>5</sup>–10<sup>6</sup> independent clones. Third, recovery of target genes after screening and sorting is typically done from large pools (10<sup>3</sup>–10<sup>7</sup>) of collected cells with subsequent random picking of just a few (10–200) clones from the pool that significantly reduces the chance of finding the best variant.<sup>26–28</sup> As a result, the previously reported methods for evolution of FPs in mammalian cells were

not adopted by protein engineers and the developed FPs did not find wide utilization for in vivo imaging.

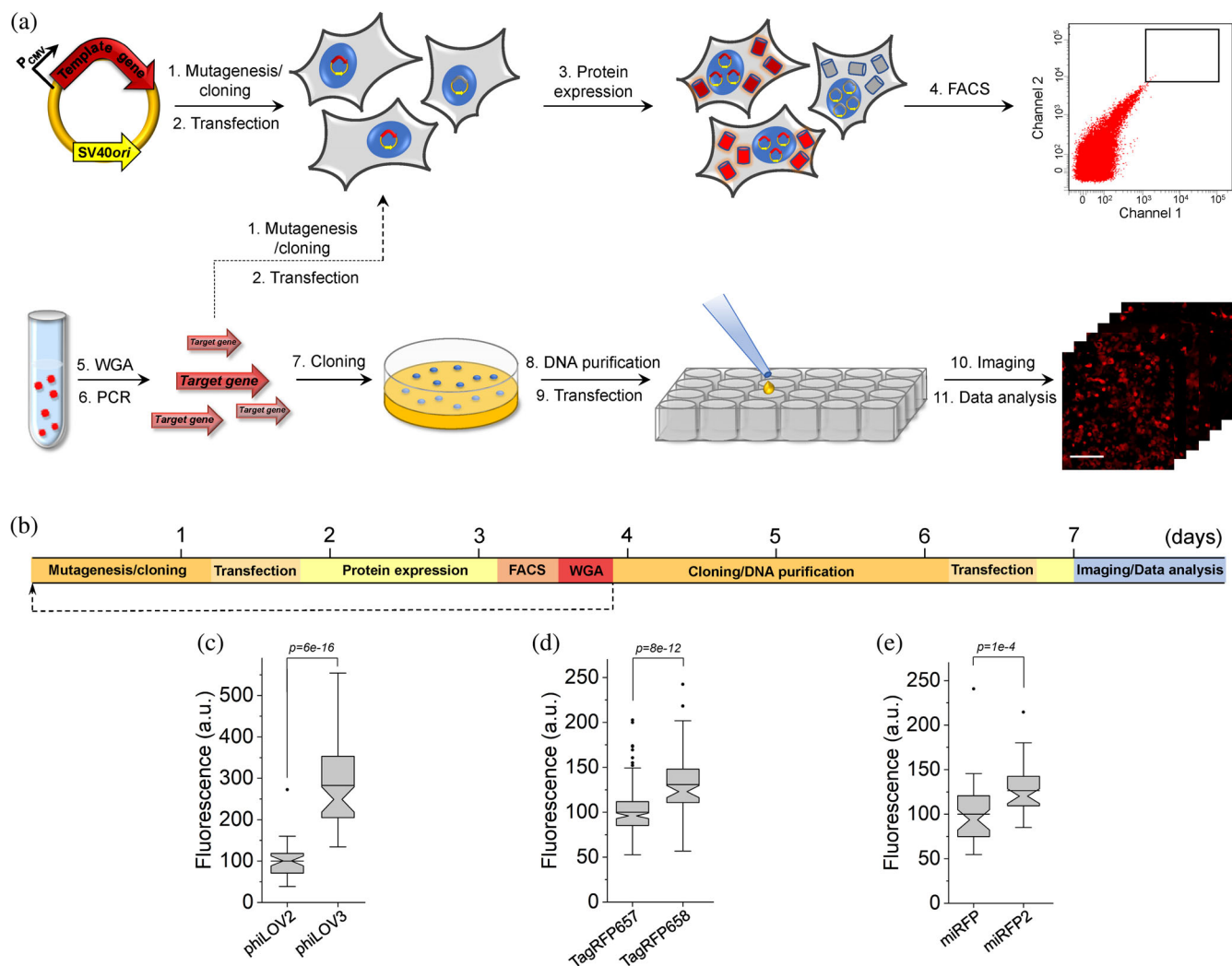
Here, we describe a simple and practical method that overcomes major limitations and drawbacks of currently available methods for directed molecular evolution of FPs in mammalian cells. The developed method allowed us to perform optimization of intracellular brightness in the course of 4 days per iteration by generation and screening of large (up to  $2 \times 10^7$ ) and highly diverse (10–15 mutations per kilobase pair) libraries of FPs in mammalian cells, and an extra 3 days for isolation and identification of individual variants. Moreover, our method allows rapid target gene recovery from very small pools of collected cells (10–100 cells), enabling genotyping and detailed phenotyping of rare cell populations common for large random libraries. We employed the described method to develop a set of green and near-infrared (NIR) FPs, named phiLOV3, TagRFP658, and miRFP2, with enhanced intracellular brightness. The improvements of intracellular brightness for phiLOV3 and miRFP2 did not correlate with molecular brightness when compared to respective parental proteins, thus highlighting the importance of protein optimization in a specific cellular context. Using rational design, we generated an enhanced version of miRFP2 with greatly increased intracellular brightness, called emiRFP2. We demonstrated applicability of the developed proteins for in vivo imaging of neurons in *Caenorhabditis elegans*, *Drosophila*, zebrafish, and intact mouse brain tissue. Combination of the developed FPs with other common green and red FPs enabled crosstalk-free multicolor imaging using standard optical configurations. Moreover, we performed dual-color NIR imaging of subcellular structures using a combination of emiRFP2 with blue-shifted NIR FP mCardinal. The developed method has a great potential to be adopted by protein engineers for optimization of FPs in mammalian cells. We also believe that the new enhanced FPs will find wide application for in vivo imaging of small model organisms.

## 2 | RESULTS

### 2.1 | Protein optimization in mammalian cells

The rapid protein optimization via directed molecular evolution is based on a simple and scalable method for expression of large gene libraries in mammalian cells in combination with high-throughput live cell screening techniques, for example, fluorescence-activated cell sorting (FACS). The workflow includes six major hands-on steps: (a) preparation of gene library; (b) transfection

of mammalian cells with gene library in bulk; (c) screening and collection of individual cells; (d) cloning of the target genes from selected cells into expression vector; (e) transfection of the cloned plasmids into mammalian cells; and (f) imaging and selection of individual clones (Figure 1a). One iteration of directed molecular evolution can be carried out in about 8 days (Figure 1b). In the case of the bulk selection, when pools of selected genes are subjected to the next round of evolution, mutagenesis and screening can be performed with a period of  $\sim 3.5$  days (Figure 1b). To validate the proposed approach, we decided to enhance intracellular brightness of the biochemically and spectrally distinct FPs. We chose a set of FPs that originated from the four evolutionary different classes of chromoproteins and those that are diverse in their later synthetic evolution conditions. Namely, we chose a green FP phiLOV2.1 engineered from the flavin-binding light, oxygen, or voltage (LOV) sensing domain,<sup>31</sup> a naturally occurring green FP UnaG cloned from freshwater eel *Anguilla japonica*,<sup>4</sup> an engineered far-red fluorescent GFP-like FP TagRFP657,<sup>32</sup> and an NIR FP miRFP derived from PAS-GAF domains of the *RpBphP1* bacteriophytochrome.<sup>33</sup> By selecting already optimized FPs as templates, our goal was to demonstrate the great potential of the proposed approach to further enhance desired characteristics in mammalian cells. First, human-codon optimized genes of the selected proteins were subjected to error-prone PCR and cloned into the mammalian expression vector containing an SV40 origin of replication. The SV40 origin of replication significantly increases the expression levels of transgene under transient transfection conditions due to episomal replication of the plasmid within a host cell that expresses the SV40 large T-antigen,<sup>34</sup> which is crucial under one plasmid per cell transfection conditions. The generated random libraries, containing around  $10^6$ – $10^7$  independent clones, were transfected into HEK293FT cells using the modified calcium phosphate method, which was optimized for single gene copy per cell delivery.<sup>35</sup> After plasmid replication and protein expression for 48 hr, we used FACS to collect  $\sim 100$  cells with the highest fluorescence intensity for each library. Next, the target genes, recovered from the pools of collected cells, were either subjected to the next round of directed evolution or directly cloned into the mammalian expression vector (Figure 1a). In the latter case, several hundreds of randomly picked clones were individually transfected into HEK cells to compare their brightness to the corresponding parental proteins. Overall, one iteration of directed evolution was carried out within  $\sim 3.5$  days and additional  $\sim 4$  days were required to clone, express, assess, and sequence individual mutants (Figure 1b). We carried out two rounds of directed evolution for each



**FIGURE 1** Rapid directed molecular evolution of fluorescent proteins in mammalian cells. (a) Workflow for directed evolution of proteins in HEK293FT cells using single-gene delivery via modified calcium phosphate transfection, fluorescence-activating cell sorting (FACS), and automated fluorescence imaging; WGA, whole-genome amplification. (b) Timeline of directed molecular evolution. (c) Green fluorescence of phiLOV2 and phiLOV3 expressed in HEK293FT cells ( $n = 40$  and  $59$  cells, respectively, from two independent transfections for each protein; one-way analysis of variance [ANOVA] is used throughout this figure). Imaging conditions: excitation 475/34 nm from an LED, emission 527/50 nm. Box plots with notches are used throughout this paper, when  $n > 9$  (the narrow part of notch, median; top and bottom of the notch, 95% confidence interval for the median; top and bottom horizontal lines, 25 and 75% percentiles for the data; whiskers extend  $1.5 \times$  the interquartile range from the 25th and 75th percentiles; horizontal line, mean; dots represent outliers, data points which are less than the 25th percentile or greater than the 75th percentile by more than 1.5 times the interquartile range). (d) Near-infrared fluorescence of TagRFP657 and TagRFP658 expressed in HEK293FT cells ( $n = 181$  and  $76$  cells, respectively, from two independent transfections for each protein). Imaging conditions: excitation 631/28 nm from an LED, emission 664LP. (e) Near-infrared fluorescence of miRFP and miRFP2 expressed in HEK293FT cells ( $n = 40$  and  $41$  cells, respectively, from two independent transfections for each protein). Imaging conditions as in (d)

template followed by screening of individually picked clones expressed in HEK cells under fluorescence microscope (see Supplementary Table 1 for the screening conditions and parameters). Assessment of the clones selected from the UnaG library did not identify variants with improved intracellular brightness, although sequencing of the brightest selected variants revealed amino acid substitutions in the structurally important

regions (Supplementary Figures 1 and 2). Therefore, we did not perform further characterization of the UnaG mutants. For phiLOV2.1 and TagRFP657, we identified multiple variants that outperformed corresponding parental proteins in terms of intracellular brightness and in case of miRFP, we identified only one mutant with improved brightness (Supplementary Figure 1). We selected the brightest variant from each library for

further evaluation. However, during imaging we found out that the brightest TagRFP657 variant had very limited photostability, we decided to choose the second brightest variant, which was confirmed to have higher than the parental protein photostability (Supplementary Figure 1). To confirm that the observed improvements were statistically significant, we repeated the measurements for each selected variant on several biological replicates in HEK cells. Indeed, all variants showed biologically and statistically significant improvements in intracellular brightness. The phiLOV2.1 variant showed 2.8-fold higher brightness over parental protein, the TagRFP657 and miRFP variants showed 30 and 27% over parental proteins, respectively (Figure 1c–e). Sequence analysis of the selected variants revealed two amino acid mutations in phiLOV2.1, six in TagRFP657, and nine in miRFP (Supplementary Figures 3–5). Correspondingly, we named the identified variants as phiLOV3, TagRFP658, and miRFP2, and used them for further detailed characterization. As a result, we were able to complete two iterations of directed molecular evolution in mammalian cells within 8 days enhancing intracellular brightness of the selected FPs (Figure 1).

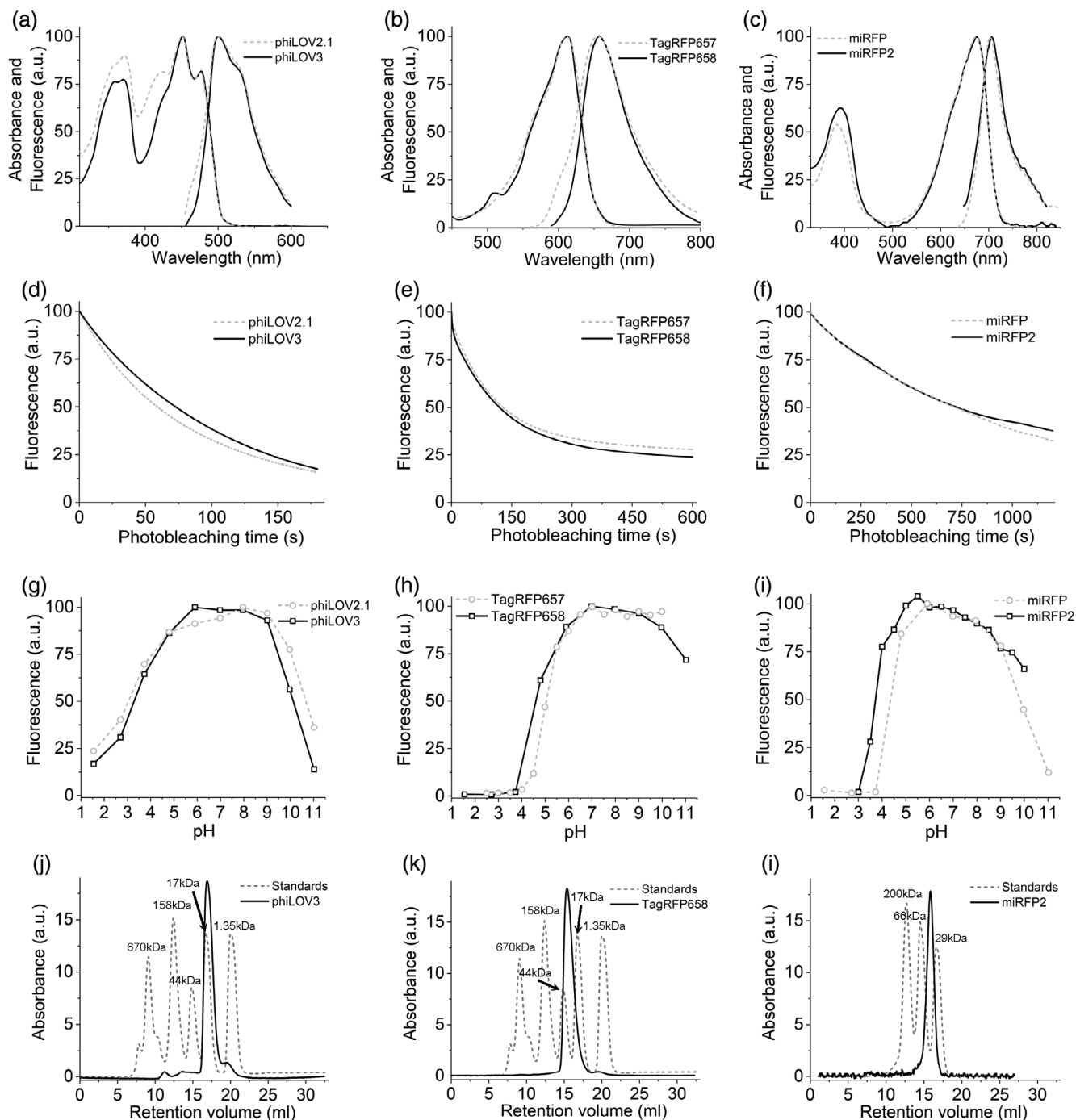
## 2.2 | Spectral and biochemical characterization of the optimized FPs

The introduced mutations did not alter the spectral properties of the developed proteins when compared to the corresponding parental proteins, showing only insignificant shifts in the maxima of the major bands (Figure 2a–c, Tables 1 and 2). Measured in solution using proteins purified from *Escherichia coli*, the molecular brightness of phiLOV3 and TagRFP658 was about 29 and 21% higher than that of their precursors, respectively, while molecular brightness of miRFP2 was 3.7-fold lower than that of miRFP (Tables 1 and 2). Improvement of the TagRFP568 intracellular brightness corresponded to increase in molecular brightness over its precursor. However, relative intracellular brightness of phiLOV3 and miRFP2 in HEK cells was several fold higher than their molecular brightness when compared to phiLOV2.1 and miRFP, respectively (Tables 1 and 2). The TagRFP658 and miRFP2 proteins exhibited similar intracellular photostability compared to the corresponding parental proteins, while phiLOV3 showed about 23% improvement in photostability halftime over its precursor (Figure 2d–f). The pH-stability of fluorescence for phiLOV3, TagRFP658, and miRFP2 was characterized by  $pK_a$  values of 3.3, 4.6, and 3.7, respectively, which were similar to that of the corresponding parental proteins (Figure 2g–i, Tables 1 and 2). Size-exclusion chromatography

demonstrated that the developed proteins preserved monomeric state even at a high concentration in solution (Figure 2j–l, see Supplementary Figure 6 for calibration plots). We also measured two-photon cross section for TagRFP658 to explore its utility for two-photon microscopy. TagRFP658 exhibited a major peak at 1230 nm in two-photon spectrum coinciding with one-photon peak (at double wavelength) rather well and exhibited a strong feature in the region corresponding to the  $S_0 \rightarrow S_n$  transitions with strong absorption only at wavelength below 830 nm (Supplementary Figure 7).

To investigate the significant difference in relative brightness of miRFP and miRFP2 in vitro and in cell culture, we evaluated BV binding efficiency in HEK cells by measuring intracellular brightness under different concentrations of exogenously administrated BV. We used miRFP703 as an additional reference since it shares the highest amino acid identity (~93%) with miRFP2 and was reported to have higher-than-miRFP2 molecular brightness<sup>20,36</sup> (Table 2). Addition of BV at 62.5  $\mu$ M resulted in 3.2- and 3.7-fold increases in intracellular brightness of miRFP and miRFP703 while miRFP2 showed only a 1.8-fold increase (Supplementary Figure 8). The data suggested that miRFP2 has higher BV binding affinity and as a result a larger fraction of miRFP2 exists in the BV-bound state in HEK cells, which potentially can account for its higher intracellular brightness over miRFP and miRFP703. Overall, characterization in vitro and in cell culture showed that enhancement of intracellular brightness was not due to changes in spectral properties or oligomeric state of the proteins but rather due to improving either molecular brightness in case of TagRFP658 or cofactor binding affinity in case of miRFP2.

Next, we compared intracellular brightness of TagRFP658 and miRFP2 in NIH3T3 mouse embryonic fibroblasts and PAC2 zebrafish embryonic fibroblasts under identical imaging conditions. While finalizing this study, it was reported that swapping the N-terminus of the *RpBphP1*-derived miRFPs with that of the *RpBphP2* protein can significantly improve intracellular brightness in mammalian cells without affecting molecular brightness.<sup>20</sup> Following the described strategy, we generated an enhanced version of miRFP2, or emiRFP2 for short, and used it for side-by-side comparison with TagRFP658 and miRFP2. To perform expression-level independent quantification of intracellular brightness, the NIR FPs were co-expressed with mClover3 under the EF1 $\alpha$ :2xCMV:EF1 $\alpha$  bidirectional promoter (Figure 3a,c). Under transient expression in NIH3T3 cells, emiRFP2 exhibited 1.4- and 4.8-fold higher fluorescence than TagRFP658 and miRFP2, respectively, when quantified using Cy5-to-green fluorescence ratio (Figure 3b; here and below we



**FIGURE 2** Spectroscopic and biochemical characterization of phiLOV3, TagRFP658, and miRFP2 in comparison to their precursors. (a–c) Absorption and fluorescence emission spectra of (a) phiLOV2.1 (dashed line) and phiLOV3 (solid line), (b) TagRFP657 (dashed line) and TagRFP658 (solid line), and (c) miRFP (dashed line) and miRFP2 (solid line). (d) Photobleaching curves of phiLOV2.1 (dashed line) and phiLOV3 (solid line) expressed in HEK293FT cells ( $n = 18$  cells from 1 transfected sample, each). Photobleaching conditions: excitation 475/34 nm from an LED at  $\sim 5$  mW/mm<sup>2</sup>. (e) Photobleaching curves of TagRFP657 (dashed line) and TagRFP658 (solid line) expressed in HEK293FT cells ( $n = 16$  cells from 1 transfected sample, each). Photobleaching conditions: excitation 628/31BP from an LED at 95 mW/mm<sup>2</sup>. (f) Photobleaching curves of miRFP (dashed line) and miRFP2 (solid line) expressed in primary cultured mouse neurons ( $n = 25$  and 61 cells, respectively, from one culture, each). Photobleaching conditions: excitation 628/31BP from an LED at 88 mW/mm<sup>2</sup>. (g–i) pH dependence of fluorescence for (g) phiLOV2.1 (dashed line) and phiLOV3 (solid line), (h) TagRFP657 (dashed line; data from Reference 32) and TagRFP658 (solid line), and (i) miRFP (dashed line; data from Reference 35) and miRFP2 (solid line). (j–l) Size-exclusion chromatography of (j) phiLOV3 at a concentration of 3.8 mg/ml, (k) TagRFP658 at a concentration of 8 mg/ml, and (l) miRFP2 at a concentration of 8.8 mg/ml (solid lines) and the indicated molecular weight standards (dashed lines)

TABLE 1 Properties of the FMN-binding phiLOV2.1 and phiLOV3 FPs

Protein	Abs (nm)	Em (nm)	EC ( $M^{-1} cm^{-1}$ )	QY (%)	Molecular brightness <sup>a</sup>	$pK_a$	Brightness in HEK cells (%) <sup>b</sup>	Intracellular photostability (s) <sup>c</sup>
phiLOV2.1	451, 476	501	13,500	20	2,700	3.0	100	59
phiLOV3	452, 477	502	15,800	22	3,480	3.3	283	73

Abbreviations: Abs, absorbance peak; Em, fluorescence emission peak; EC, extinction coefficient; ND, not determined; QY, quantum yield.

<sup>a</sup>Molecular brightness is a product of extinction coefficient and fluorescence emission QY.

<sup>b</sup>Determined as mean green fluorescence relative to phiLOV2.1.

<sup>c</sup>Measured under continuous 470/25 nm wide-field illumination.

reported fluorescence intensities without normalization to excitation and emission efficiencies for spectrally distinct FPs, unless otherwise indicated, to provide direct comparison of intracellular brightness in real experimental settings rather than intrinsic properties of the FPs). Similar results were obtained in PAC2 fibroblasts with emiRFP2 being 1.5- and 4.1-fold brighter than TagRFP658 and miRFP2, respectively (Figure 3d). Administration of exogenous BV at 25  $\mu$ M in PAC2 cells further increased fluorescence of miRFP2 and emiRFP2 by 2- and 2.3-fold, respectively (Figure 3d). Consistently with the previously reported results,<sup>20</sup> swapping the N-terminus of miRFP2 significantly improved its intracellular brightness. The emiRFP2 protein also outperformed TagRFP658 in terms of intracellular brightness in Cy5 channel.

It is important to validate if the FPs optimized in mammalian cells outperform their counterparts evolved in a bacterial system. Therefore, we decided to compare intracellular brightness and photostability of TagRFP658, miRFP2, and emiRFP2 to spectrally similar GFP-like and BphP-based NIR-FPs in HEK cells using a standard Cy5 filter set with a wide band-pass emission filter (730/140 nm), which allows efficient collection of NIR fluorescence. Based on the literature search, we selected a set of monomeric and dimeric NIR FPs including mCardinal,<sup>11</sup> iRFP670,<sup>37</sup> miRFP680,<sup>20</sup> iRFP682,<sup>37</sup> mRhubarb713,<sup>38</sup> and miRFP720,<sup>39</sup> which were reported to have high performance in cultured mammalian cells. To account for the expression level during transient transfection of HEK cells, the NIR FPs were co-expressed with EGFP using the self-cleavage P2A peptide.<sup>40</sup> The intracellular brightness of emiRFP2 was higher than all other assessment BphP-based NIR FPs, but about 1.4-fold lower than mCardinal (Figure 3e, see Table 2 for the quantification of all proteins). It should be noted that the expression level of the FPs as assessed by fluorescence intensity of GFP was comparable for all expressed constructs. Intracellular photostability of TagRFP658, miRFP2, and emiRFP2 measured under continuous wide-field illumination was  $\sim$ 2, 3, and 3-fold higher than

that of mCardinal, respectively, however, lower than that for miRFP680, mRhubarb, and miRFP720 (Figure 3f and Table 2). Based on the assessed characteristics, emiRFP2 and miRFP680 exhibited the best combination of brightness and photostability among the tested NIR FPs. However, as for the majority of in vivo applications, intracellular brightness is usually the most crucial property, we decided to use mCardinal as a major reference for further characterization of the developed NIR FPs in neurons. Moreover, mCardinal outperformed other recently published NIR FPs, such as mMaroon,<sup>41</sup> and miRFP670nano,<sup>19</sup> when compared side-by-side in HEK cells (Supplementary Figure 9).

### 2.3 | Characterization of TagRFP658 and (e)miRFP2 in neurons

We evaluated the utilization of TagRFP658, miRFP2 and emiRFP2 as NIR fluorescent probes for neuronal labeling under one photon microscopy using standard Cy5 and Cy5.5 filter sets. As a reference, we chose to use mCardinal due to its superior intracellular brightness compared to all other tested NIR-FPs (Table 2). First, we expressed TagRFP658 in primary hippocampal mouse neurons and in a subset of neurons in live zebrafish larvae. In both preparations, the fluorescence of TagRFP658 was evenly distributed within the cytosol, individual dendrites, and nucleus of live neurons without any aggregation or nonspecific localization (Figure 4a–c). In cultured neurons, TagRFP658 was about 28% brighter and twice more photostable than mCardinal (photobleaching half-time for TagRFP658 and mCardinal was 164 and 80 s, respectively; Figure 4d,e). Using whole-cell patch clamp recordings, we showed that TagRFP658 expression did not alter membrane resistance, membrane capacitance, or the resting potential of cultured mouse neurons (Supplementary Figure 10).

Next, we assessed performance of emiRFP2 in cultured neurons side-by-side with mCardinal and TagRFP658. Since emiRFP2 outperformed miRFP2 in

TABLE 2 Spectral and biochemical properties of the selected NIR FPs

Protein	Abs (nm)	Em (nm)	EC ( $M^{-1} cm^{-1}$ )	QY (%)	Molecular brightness <sup>a</sup>	pK <sub>a</sub>	Brightness in HEK cells (%) <sup>b</sup>	Photostability (s) <sup>c</sup>	Initial photobleaching rate (%/s) <sup>d</sup>	Ref
mCardinal	603	651	79,000	18	14,220	5.3	290	340	0.24	63
mMaroon1	609	657	80,000	11	8,800	6.2	125 <sup>e</sup>	ND	ND	41
iRFP670	643	670	114,000	11	15,540	4.0	154	245	0.56	37
miRFP670nano	645	670	95,000	10.8	10,260	3.7	8 <sup>e</sup>	ND	ND	19
miRFP680	661	680	94,000	14.5	13,630	4.5	160	ND	0.06	20
iRFP682	663	682	90,000	11	9,900	4.5	140	904	0.09	37
miRFP703	674	703	90,900	8.6	7,820	4.5	48	ND	ND	36
mRhubard713	690	713	113,500	7.6	8,630	ND	33	ND	0.02	38
miRFP720	702	720	98,000	6.1	5,980	4.5	61	ND	0.03	39
TagRFP657	611	657	34,000	10	3,400	5.0	61 <sup>f</sup>	ND	ND	32
TagRFP658	611	658	41,200	10	4,120	4.6	79	640	0.12	This study
miRFP	674	703	92,400	9.7	8,960	4.3	79 <sup>g</sup>	ND	ND	35
miRFP2	676	706	55,600	4.3	2,390	3.7	100	995	0.08	This study
emiRFP2	ND	ND	ND	ND	ND	ND	205	1,050	0.08	This study

Abbreviations: Abs, absorbance peak; Em, fluorescence emission peak; EC, extinction coefficient; FPs, fluorescent proteins; ND, not determined; NIR, near-infrared; QY, quantum yield.

<sup>a</sup>Molecular brightness is a product of extinction coefficient and fluorescence emission QY.

<sup>b</sup>Determined as a Cy5-to-green fluorescence intensity ratio for live HEK cells relative to miRFP2 unless otherwise state.

<sup>c</sup>Measured in live HEK cells under continuous wide-field illumination with 631 nm laser at 66 mW/mm<sup>2</sup>.

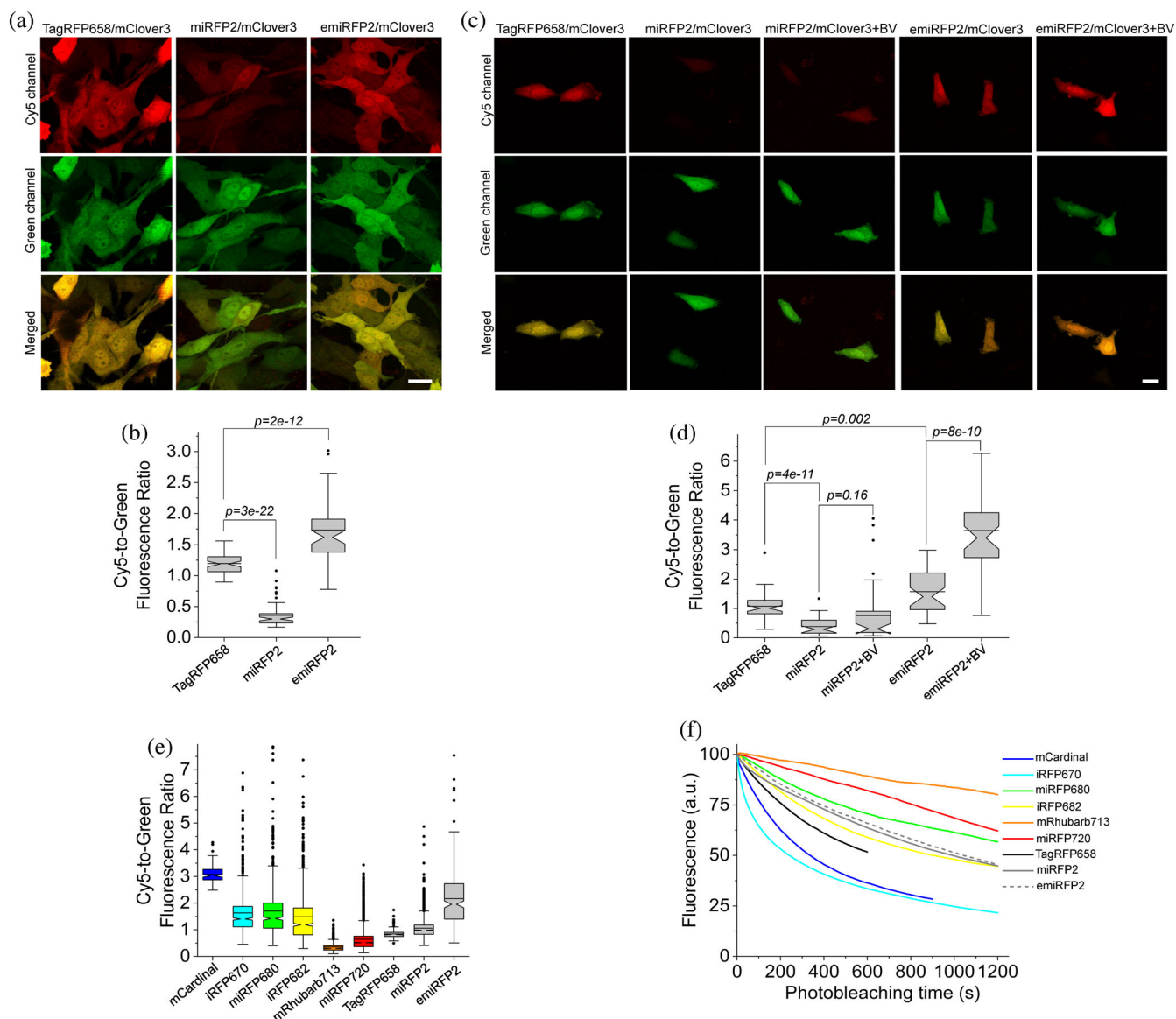
<sup>d</sup>Determined as a liner fit of the initial segment of photobleaching curve.

<sup>e</sup>Determined as mean Cy5 fluorescence intensity in live HEK cells relative to mCardinal.

<sup>f</sup>Determined as mean Cy5 fluorescence intensity in live HEK cells relative to TagRFP658.

<sup>g</sup>Determined as mean Cy5 fluorescence intensity in live HEK cells relative to miRFP2.





**FIGURE 3** Fluorescence imaging of TagRFP658, miRFP2, and emiRFP2 expressed in mammalian and zebrafish cell cultures. (a) Representative fluorescence images of live NIH3T3 mouse embryonic fibroblasts co-expressing TagRFP658, miRFP2, and emiRFP2 with mClover3 under the EF1 $\alpha$ :2xCMV:EF1 $\alpha$  bidirectional promoter ( $n = 67, 62,$  and  $59$  cells, respectively, from one transfection each; the dynamic range of the images for the same channel was kept constant throughout the Panels (b) and (d)). Imaging conditions: Cy5 channel, excitation  $633$  nm from a laser, emission  $660$ – $785$  nm; green channel, excitation  $488$  nm from a laser, emission  $495$ – $530$  nm). (b) Cy5-to-green fluorescence ratio of live NIH3T3 fibroblasts shown in (a) ( $n = 67, 62,$  and  $59$  cells for TagRFP658, miRFP2, and emiRFP2, respectively, from one transfection each; Kruskal–Wallis analysis of variance [ANOVA]). Box plots with notches are used in this figure (see Figure 1c for the description). (c) Representative fluorescence images of live Pac2 zebrafish embryonic fibroblasts co-expressing TagRFP658, miRFP2, and emiRFP2 with mClover3 under the EF1 $\alpha$ :2xCMV:EF1 $\alpha$  bidirectional promoter with and without addition of  $25$   $\mu$ M BV for  $3$  hr before imaging ( $n = 49, 36, 41, 39,$  and  $39$  cells for TagRFP658, miRFP2, miRFP2 + BV, emiRFP2, and emiRFP2 + BV, respectively, from one transfection each). Imaging conditions as in (a). (d) Cy5-to-green fluorescence ratio of live Pac2 fibroblasts shown in (c) with and without addition of  $25$   $\mu$ M BV for  $3$  hr before imaging ( $n = 49, 36, 41, 39,$  and  $39$  cells for TagRFP658, miRFP2, miRFP2 + BV, emiRFP2, and emiRFP2 + BV, respectively, from one transfection each; Kruskal–Wallis ANOVA). Imaging conditions as in (a). Scale bars,  $20$   $\mu$ m. (e) Cy5-to-green fluorescence ratio for near-infrared-fluorescent proteins (NIR-FPs) co-expressed with GFP via the P2A peptide in live HEK cells ( $n = 144; 775; 612; 837; 446; 2,937; 226; 1,981;$  and  $271$  cells for mCardinal, iRFP670, miRFP680, iRFP682, mRhubarb713, miRFP720, TagRFP658, miRFP2, and emiRFP2, respectively, from two independent transfections each). Box plots with notches are used in this figure (see Figure 1c for the full description). Imaging conditions: Cy5 fluorescence, excitation  $635/22$  nm from  $637$  nm laser, emission  $730/140$  nm; green fluorescence, excitation  $478/24$  nm for an LED; emission  $535/46$  nm. (f) Photobleaching curves for NIR-FPs expressed in live HEK cells under continuous wide-field illumination from  $637$  nm laser at  $66$   $\text{mW}/\text{mm}^2$  ( $n = 55, 32, 10, 43, 32, 48, 56, 38,$  and  $37$  cells for mCardinal, iRFP670, miRFP680, iRFP682, mRhubarb713, miRFP720, TagRFP658, miRFP2, and emiRFP2, respectively, from two independent transfections each)

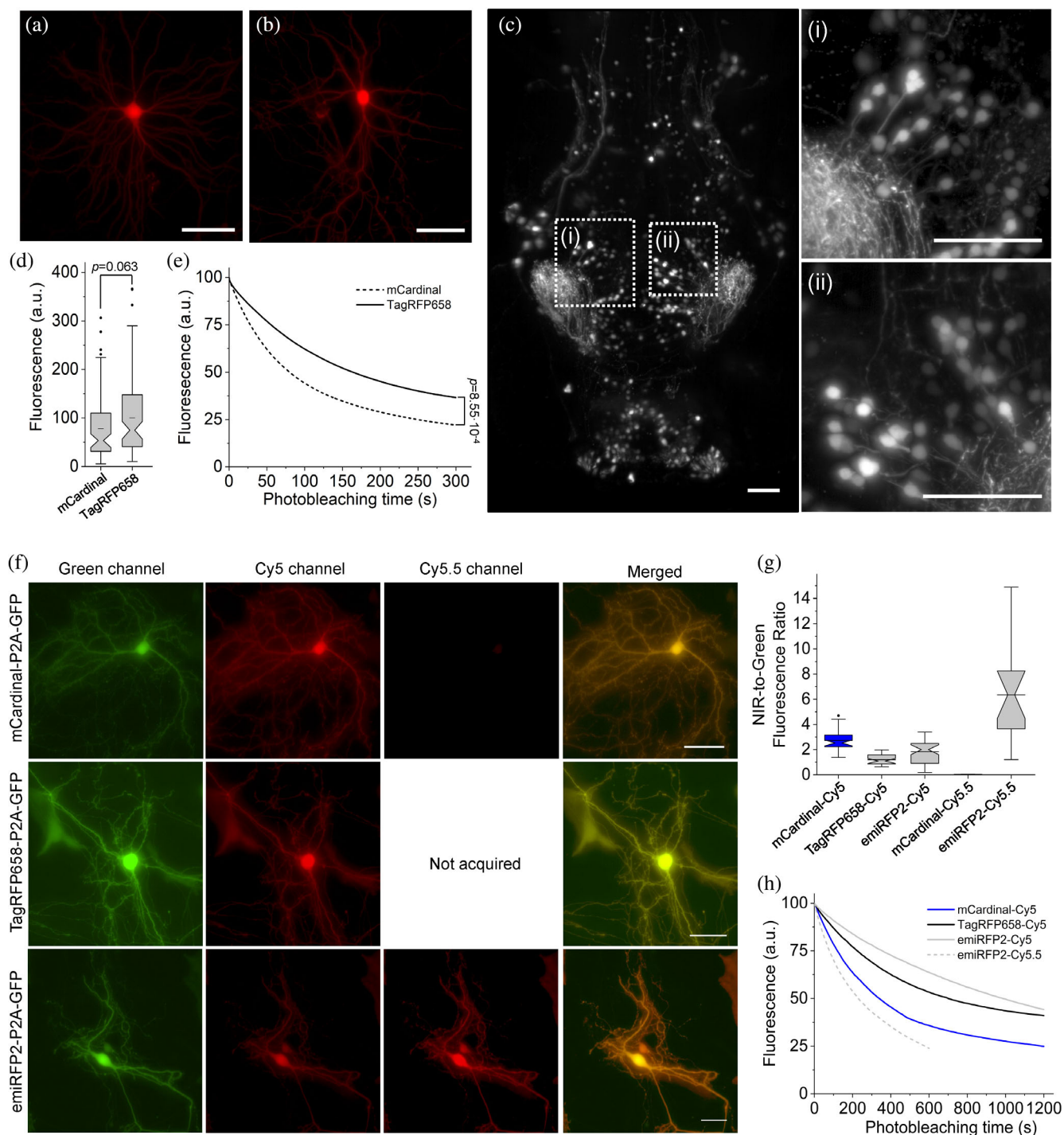


FIGURE 4 Legend on next page.

earlier experiments it was not used for quantitative imaging in neurons, although miRFP2 can be efficiently expressed and imaged in cultured neurons both under transient transfection and rAAV transduction (Supplementary Figure 11). The fluorescence of the NIR FPs co-expressed with EGFP via P2A peptide was evenly distributed within the cytosol, individual dendrites, and

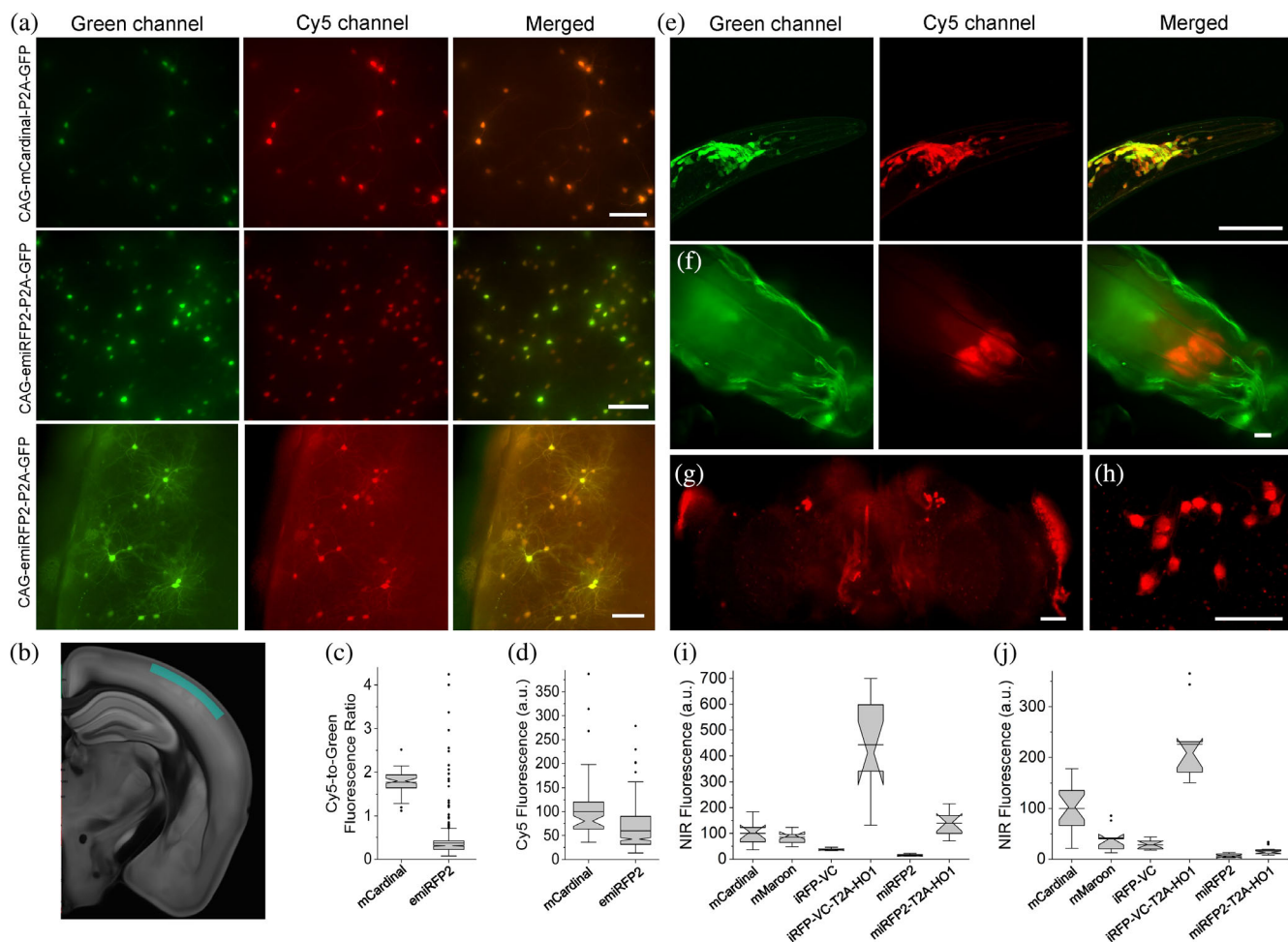
nucleus of live cultured neurons without any aggregation or nonspecific localization (Figure 4f). Quantification of fluorescence intensity revealed that emiRFP2 were about 3.5-times brighter in Cy5.5 channel than in Cy5 channel, while the mCardinal fluorescence in Cy5.5 channel was almost undetectable (Figure 4g). When quantified by Cy5-to-green fluorescence ratio, mCardinal was 2.2- and

1.5-fold brighter than TagRFP658 and emiRFP2, respectively. However, when raw mean fluorescence values in Cy5 channel were compared, TagRFP658 was about 20% brighter than mCardinal, similarly to the results shown in Figure 4c (note comparable fluorescence intensity for the representative images of mCardinal and TagRFP658 in Cy5 channel, but significantly dimmer GFP fluorescence in case of mCardinal-P2A-GFP construct; Supplementary Figure 12). Intracellular photostability of emiRFP2 was four-fold lower under Cy5.5 illumination compared to Cy5 excitation (photobleaching half-times were 230 and 990 s under Cy5.5 and Cy5 illumination, respectively; Figure 4h). At the same time, emiRFP2 exhibited superior intracellular photostability compared to mCardinal and TagRFP658 in the Cy5 channel, which were characterized by photobleaching halftime of 340 and 700 s, respectively, closely matching corresponding values obtained in HEK cells (Figure 4h and Table 2). Similar correlations of intracellular brightness and photostability in Cy5 and Cy5.5 channels were also observed in live HEK cells (Supplementary Figure 12). These results demonstrated that Cy5.5 channel provided high efficiency of the emiRFP2 fluorescence imaging; however, the gain in brightness came at the cost of reduced photostability.

High performance of miRFP2 and emiRFP2 in cultured neurons encouraged us to express them in vivo in model organisms such as mice, *C. elegans*, and *Drosophila melanogaster*. First, we co-expressed mCardinal and emiRFP2 with GFP in cortical neurons in mice via in

utero *electroporation* and performed imaging of the expressed proteins in acute brain slices at around P28. The emiRFP2 protein expressed in vivo showed even distribution in cell bodies and processes without aggregation (Figure 5a,b). Quantitative imaging showed that mCardinal had 4.8-fold higher Cy5-to-green fluorescence ratio than emiRFP2; however, mean fluorescence intensity of mCardinal was only 1.7-fold higher than that of emiRFP2 (Figure 5c,d; note significantly lower green fluorescence intensity in the representative image for CAG-mCardinal-P2A-GFP construct compared to CAG-emiRFP2-P2A-GFP in Figure 5a). In addition, values for Cy5-to-green fluorescence ratio in the case of emiRFP2 exhibited significant variability ranging from 0.07 to 4.2 versus only 1.1 to 2.5 for mCardinal. Next, we expressed the codon-optimized genes of miRFP2 using pan-neuronal expression systems in transgenic *C. elegans* and in *Drosophila* fruit flies. In case of *C. elegans*, we did not observe any specific NIR signal, while NIR fluorescence in larvae and adult fruit flies was clearly detectable although its intensity was several times lower than in cultured neurons under the same imaging conditions (Supplementary Figure 13). Reduced fluorescence of miRFP2 can be due to the insufficient concentration of the BV cofactor in worms and flies. Previously, it was shown that co-expression of heme oxygenase-1 (HO1) in worms and flies can enable fluorescence of the BphP-based FPs<sup>42</sup> and biosensors.<sup>43</sup> To optimize conditions for miRFP2 expression, we constructed two bicistronic vectors using IRES2 and a viral 2A cleavage sequence to co-

**FIGURE 4** Intracellular brightness and photostability of mCardinal, TagRFP658, and emiRFP2 in live cultured hippocampal mouse neurons and in vivo in zebrafish larvae. (a,b) Representative fluorescence images of primary cultured mouse hippocampal neurons expressing TagRFP658 at (a) 14 and (b) 23 days in vitro (DIV;  $n = 53$  and 33 neurons, respectively, from two independent cultures). Imaging condition: excitation 631/28 nm from an LED, emission 664LP. (c) Representative light sheet image of head of zebrafish larvae at 4 days postfertilization expressing TagRFP658 in neurons ( $n = 10$  fish from two independent injections). Imaging conditions: excitation 638 nm from a laser, emission 665LP. (i, ii) High-magnification images of the respective regions shown in white boxes in e. Scale bars, 50  $\mu\text{m}$ . (d) Relative fluorescence of cultured mouse hippocampal neurons expressing mCardinal and TagRFP658 ( $n = 78$  and 85 neurons, respectively, from two independent cultures for each protein; one-way analysis of variance [ANOVA]). Imaging conditions as in (a). Box plots with notches are used in this figure (see Figure 1c for the full description). (e) Raw photobleaching curves for mCardinal (dashed line) and TagRFP658 (solid line) in primary cultured mouse hippocampal neurons ( $n = 9$  and 7 neurons, respectively, from one culture each; one-way ANOVA). Imaging condition: excitation 631/28 nm from an LED at 70  $\text{mW}/\text{mm}^2$ , emission 664LP. (f) Representative fluorescence images of cells transfected with pAAV-CAG-mCardinal-P2A-GFP (top), pAAV-TagRFP658-P2A-GFP (middle), and pAAV-emiRFP2-P2A-GFP ( $n = 39, 33,$  and 41 neurons from three, two, and three independent transfections from one culture each for mCardinal, TagRFP658, and emiRFP2, respectively, for Cy5 channel and  $n = 15$  and neurons from one independent transfection from one culture each for mCardinal and emiRFP2, respectively, for Cy5.5 channel). Imaging conditions: Cy5 channel: excitation 635/22 nm from 637 nm laser, emission 730/140 nm; Cy5.5 channel: excitation 680/13 nm from 680 nm laser, emission 710 LP; GFP channel: excitation 478/24 nm from an LED; emission 535/46 nm. Images in Cy5 and Cy5.5 were taken under matching excitation intensity (66  $\text{mW}/\text{mm}^2$ ) and the same exposure time (100 ms). The dynamic range of fluorescence intensity in Cy5 and Cy5.5 channels are identical across all images. Scale bar, 20  $\mu\text{m}$ . (g) Near-infrared (NIR)-to-green fluorescence ratio for mCardinal, TagRFP658, and emiRFP2 for the experiment shown in (f). (h) Intracellular photostability of mCardinal, TagRFP658, and emiRFP2 in Cy5 and Cy5.5 channels ( $n = 8, 7,$  and 9 neurons from three, two, and three independent transfections from one culture each for mCardinal, TagRFP658, and emiRFP2, respectively, under Cy5 excitation and  $n = 5$  neurons from one transfection from one culture for emiRFP2 under Cy5.5 excitation). Imaging conditions the same as in (a)



**FIGURE 5** In vivo imaging and characterization of miRFP2 and emiRFP2. (a) Representative fluorescence images of live brain slices expressing mCardinal-P2A-GFP (top) and emiRFP2-P2A-GFP (middle and bottom;  $n = 6$  slices from three mice each; L1 and L2/3 of somatosensory cortex, see Panel (b) for brain map). Imaging conditions: green channel: excitation 478/24 nm for an LED, emission 535/46 nm; Cy5 channel: excitation 635/22 nm from 637 nm laser, emission 730/140 nm. Top and middle rows of images are presented using the same dynamic range to facilitate visual comparison of mCardinal and emiRFP2 expressing neurons, bottom row images have adjusted LUTs to visualize processes of the emiRFP2 positive neurons. Scale bars, 100  $\mu\text{m}$ . (b) Brain section indicating imaged region highlighted in marine blue (image from [kimlab.io/brain-map](http://kimlab.io/brain-map)). (c) Cy5-to-Green Fluorescence Ratio of live neurons shown in (a) ( $n = 67$  and 803 neurons from three mice each for mCardinal and emiRFP2, respectively). Imaging conditions same as in (a). (d) Cy5 mean fluorescence of live neurons shown in (a) ( $n = 67$  and 803 neurons from three mice each for mCardinal and emiRFP2, respectively). Imaging conditions same as in (a). (e) Representative fluorescence images of the *Caenorhabditis elegans* head co-expressing codon-optimized genes of miRFP2-T2A-HO1 and jGCamp7b in neurons ( $n = 15$  worms from two independent microinjections). Imaging conditions: green channel, excitation 488 nm from a laser, emission 500–650 nm; Cy5 channel, excitation 631 nm from a laser, emission 645–700 nm. Scale bar, 50  $\mu\text{m}$ . (f) Representative fluorescence images of live intact third instar *Drosophila* larva expressing miRFP2-T2A-HO1 ( $n = 10$  larvae from two transgenic lines). Imaging conditions: green channel, excitation 475/34 nm from LED, emission 527/50 nm (green fluorescence corresponds to autofluorescence); Cy5 channel, excitation 631/28 nm from LED, emission 665LP. Scale bar, 100  $\mu\text{m}$ . (g) Representative low-magnification fluorescence image of brain explant from adult *Drosophila* fly expressing codon-optimized gene of miRFP2-T2A-HO1 in R84C10-GAL4 line ( $n = 5$  brains from one transgenic lines). Imaging conditions: excitation 635/22 nm from 637 nm laser, emission 665LP. Scale bar, 50  $\mu\text{m}$ . (h) Representative high-magnification fluorescence image of brain explant from adult *Drosophila* fly expressing codon-optimized miRFP2-T2A-HO1 in R84C10-GAL4 line ( $n = 5$  brains from one transgenic lines). Imaging conditions: excitation 631 nm from a laser, emission 645–700 nm. Scale bar, 50  $\mu\text{m}$ . (i) Relative near-infrared (NIR) fluorescence of mCardinal, mMaroon, iRFP-VC, iRFP-VC-T2A-HO1, miRFP2, and miRFP2-T2A-HO1 expressed pan-neuronally in *Drosophila* third instar larvae ( $n = 10, 12, 7, 11, 9,$  and 10 larvae from one transgenic line each, respectively). Imaging conditions same as in (h). (j) Relative Cy5 fluorescence of mCardinal, mMaroon, iRFP-VC, iRFP-VC-T2A-HO1, miRFP2, and miRFP2-T2A-HO1 expressed pan-neuronally in adult *Drosophila* fly ( $n = 19, 10, 11, 11, 11,$  and 21 flies from one transgenic line each, respectively). Imaging conditions same as in (h)

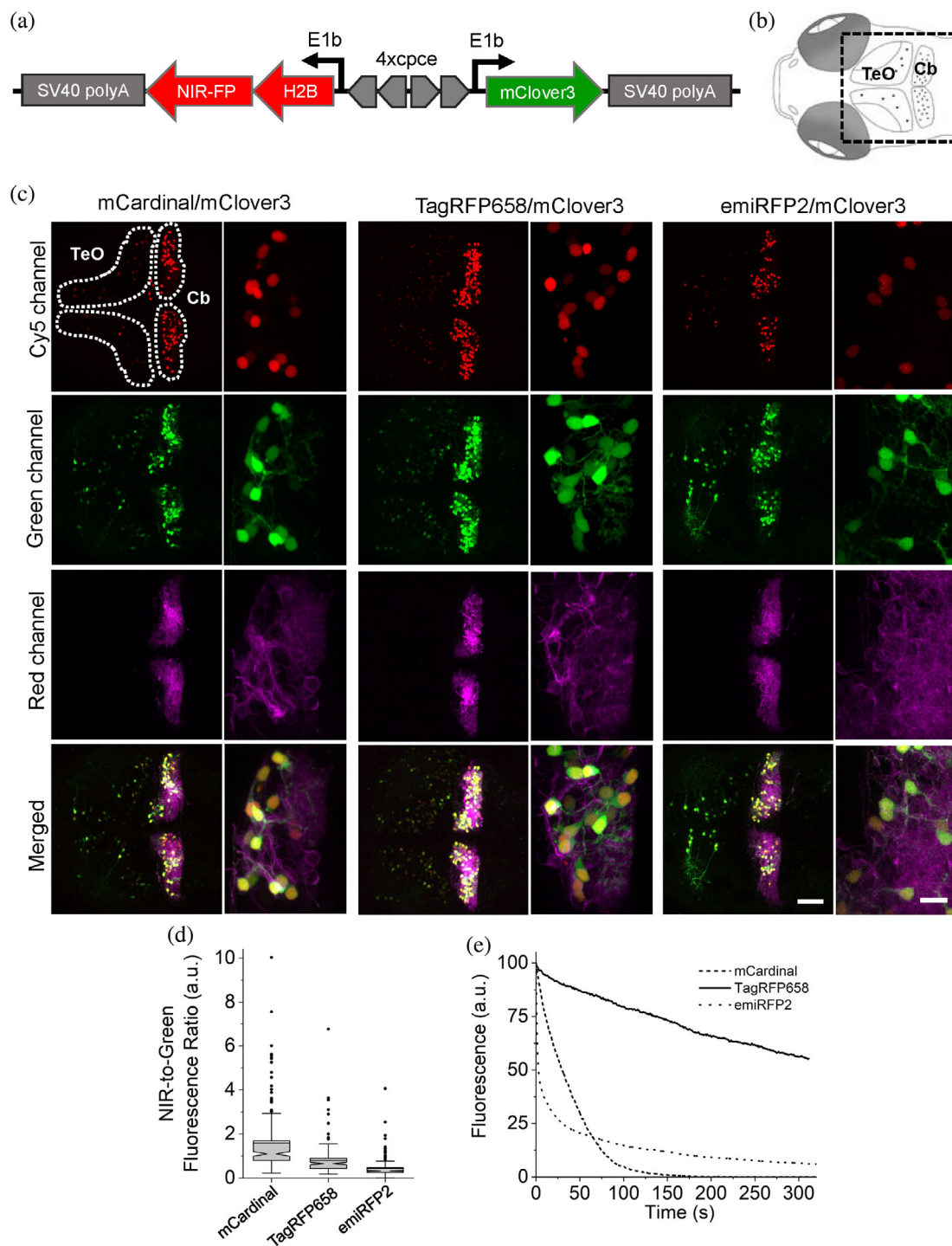
express miRFP2 and HO1 and transfected them into HEK cells. Quantification of NIR fluorescence in HEK cells revealed that HO1 co-expression via IRES2 and P2A improves miRFP2 brightness by 40 and 83%, respectively (Supplementary Figure 14). Therefore, we decided to use the viral P2A cleavage sequence to co-express codon-optimized genes of miRFP2 and HO1 in neurons in worms and flies. Indeed, co-expression of HO1 enabled bright NIR fluorescence of miRFP2 in *C. elegans* and in larvae and adult fruit flies allowing visualization of individual neurons (Figure 5e–h). To validate miRFP2 performance in fruit flies we decided to compare its fluorescence to that of other GFP-like and BphP-derived NIR FPs under identical expression conditions. We establish transgenic lines with pan-neuronal expression of *Drosophila* codon optimized genes of mCardinal, mMaroon, iRFP-VC (aka iRFP713/V256C), and iRFP-VC-P2A-HO1 and quantified their Cy5 fluorescence in live intact third instar larva and adult fruit flies (Figure 5i,j). Co-expression of HO1 significantly enhanced brightness of miRFP2 and iRFP-VC both in adult flies and in larvae, although increase of miRFP2 brightness in adult flies was less pronounced than that in larva, 2.6-fold versus 9-fold. In the case of iRFP-VC, fluorescence enhancement in larva and adult flies was comparable, about 12- and 8-fold, respectively. While miRFP2-T2A-HO1 outperformed mCardinal and mMaroon in terms of brightness in third instar larva, it was noticeably dimmer in adult fruit flies (Figure 5i,j). At the same time, co-expression of iRFP-VC with HO1 resulted in the brightest fluorescence among the tested proteins both in larva and adult flies. Thus, we demonstrated that miRFP2 is a suitable NIR FP for imaging neurons in culture and in vivo in small model organisms, like worms and flies, and outperformed other high performing NIR FPs under certain conditions, for example, in fruit fly larva. Our data also showed that co-expression of HO1 in worms and flies are essential for achieving sufficient brightness of BphP-based NIR FPs in *C. elegans* and *Drosophila* flies.

## 2.4 | Multicolor imaging in cell culture and in vivo

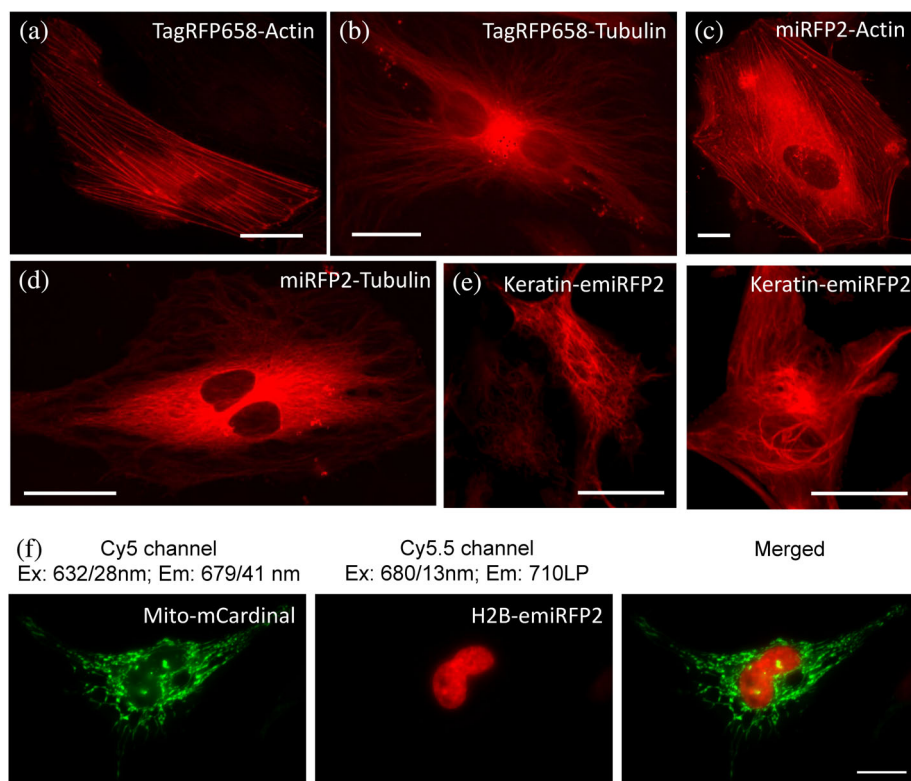
We tested the feasibility of TagRFP658 and miRFP2 in multicolor neuronal imaging in combination with green and red FPs expressed in larval zebrafish cerebellar Purkinje cells (PCs). First, we generated a set of the bidirectional expression constructs (Figure 6a) using zebrafish codon-optimized genes of NIR FPs allowing to co-express a nuclear-localized H2B histone fused to either mCardinal, TagRFP658, miRFP2, or emiRFP2 with

cytoplasmic green FP mClover3 and transfected them into NIH3T3 cells stably expressing trans-Golgi network protein 46 (TGN46) fused to red FP mScarlet. Imaging using standard filter configurations under confocal microscope allowed for crosstalk-free triple color imaging for each combination of the selected FPs (Supplementary Figure 15). Next, we expressed the selected FPs in specific neuronal subpopulations of zebrafish larvae, especially targeting cerebellar PCs using the same constructs but carrying a PC-specific bidirectional promoter<sup>44</sup> instead of the ubiquitous CMV-EF1 promoter. As the enhanced variant of miRFP2, emiRFP2, exhibited higher intracellular brightness both in NIH3T3 cells and zebrafish PAC2 fibroblast than that of the original miRFP2 (Figure 3b,d), it was selected for expression in zebrafish larva. Using a standard confocal microscope, we easily visualized NIR fluorescence in PC nucleus using Cy5 channel together with mClover3 distributed throughout entire PC's cytoplasm using green channel, while imaging in red channel provided clear visualization of the PCs axonal and/or dendritic structures with membrane tagged mScarlet fluorescence whose expressions are predominantly detected in PCs together with slight expression in tectal neurons (Figure 6b,c). Thus, NIR fluorescence of the tested NIR FPs were easily distinguishable from mScarlet fluorescence and thus can be used to label multiple neuronal compartments in conjunction with additional blue, or green FPs. We further quantified the brightness and photostability of mCardinal, TagRFP658, and emiRFP2 fluorescence by expressing corresponding constructs in PCs of less pigmented brass embryos (Supplementary Figure 16). To account for expression level of the FPs, we calculated NIR-to-green fluorescence ratio for single PCs in zebrafish embryos. Fluorescence quantification revealed that mCardinal fluorescence was about twice higher than that of TagRFP658, whereas emiRFP2 exhibited 3.6-fold lower brightness than mCardinal (Figure 6d). Photobleaching experiments performed under identical excitation power for all selected FPs demonstrated 10-times higher photostability of TagRFP658 over mCardinal. However, emiRFP2 fluorescence exhibited rapid decay with half-time about 3 s making emiRFP2 about 10-fold less photostable than mCardinal at least in this zebrafish model (Figure 6e). Despite the lower intracellular brightness of TagRFP658 than mCardinal, it can be a more practical fluorescent tag for live imaging due to significantly higher photostability.

Earlier we demonstrated that (e)miRFP2 imaging in Cy5.5 channel is several times more efficient than in Cy5 channel, while mCardinal fluorescence is not detectable in Cy5.5 channel (Figure 4f,g and Supplementary Figure 12). Therefore, we decided to explore a possibility for dual-color NIR imaging of subcellular structures



**FIGURE 6** Three-color in vivo imaging of zebrafish larvae. (a) Schematic drawing of a bidirectional PC-specific expression system using 4x PC specific regulatory element (cpce) to co-express near-infrared (NIR)-FP-H2B fusion protein with mClover3. (b) Schematic drawings of dorsal view of 4 days postfertilization (dpf) larval zebrafish brain region delineating the optic tectum (TeO) and the cerebellum (Cb). The region of interest enclosed by a square was recorded in images shown below. (c) Representative confocal images of 4 dpf zebrafish larvae transiently co-expressing H2B-NIR-FP with mClover3 (expression of this construct occurs in a mosaic manner) in stable transgenic background with PC specific membrane targeted mScarlet. The TeO and the cerebellar region (Cb) are enclosed by the white dashed lines. Each subset of images (left, whole overview of tectum and cerebellar region; right, higher magnified images depicting PCs) shows the expression of each H2B-NIR-FP fusion (left, mCardinal; middle, TagRFP658; right, emiRFP2), cytoplasmic mClover3, membrane targeted mScarlet, and the overall merged image (upper to lower;  $n = 4$  fish for each subset). Imaging conditions, NIR channel: excitation 633 nm laser, emission 722/63 nm; green channel: excitation 488 nm laser, emission 513/17 nm; red channel: excitation 561 nm laser, emission 585/15 nm. Scale bars: 50  $\mu\text{m}$  (overviewed images), 10  $\mu\text{m}$  (higher magnified images) (d) NIR-to-green fluorescence ratio for NIR-FP expressing PCs ( $n = 175, 161, 192$  cells for mCardinal, TagRFP658, and emiRFP2, respectively, from four fish each; imaging conditions as in (c)). Box plots with notches are used (see caption of Figure 1c for the full description). (e) Photostability curves for mCardinal (dashed line), TagRFP658 (solid line), and emiRFP (dotted line) expressed in PCs ( $n = 4$  cells for each from one fish each)



**FIGURE 7** Fluorescence imaging of TagRFP658, miRFP2, and emiRFP2 fusions expressed in HeLa cells. (a–e) Representative fluorescence images of live HeLa cells transfected with (a) TagRFP658- $\alpha$ -Tubulin ( $n = 11$  cells from two independent transfections) and (b) TagRFP658- $\beta$ -actin ( $n = 19$  cells from two independent transfections). (c) miRFP2- $\alpha$ -Tubulin ( $n = 15$  cells from two independent transfections), (d) miRFP2- $\beta$ -actin ( $n = 21$  cells from two independent transfections), and (e) Keratin-emiRFP2 ( $n = 8$  cells from two independent transfections). Imaging conditions: (a–d) excitation 631/28 nm from an LED, emission 664LP under wide-field microscope. (e) Excitation 640 nm from laser, emission 670–750 nm under confocal microscope. (f) Dual-color near-infrared (NIR) imaging of live HeLa cells co-expressing Mito-mCardinal and H2B-emiRFP2 ( $n = 12$  cells from two independent transfection)

using combination blue-shifted and red-shifted NIR FPs, for example, mCardinal and emiRFP2. First, we verified that the developed NIR FPs can be properly localized in fusions with structured proteins in mammalian cells. Indeed, fusions of TagRFP658 and (e)miRFP2 with  $\alpha$ -tubulin,  $\beta$ -actin, and keratin demonstrated proper localization in cultured mammalian cells (Figure 7a–f). Next, we co-expressed Mito-mCardinal and H2B-emiRFP2 fusions in HeLa cells and acquired images in Cy5 and Cy5.5 channels. To further improve spectral separation, we swapped the wide band-pass emission filter in our standard Cy5 channel with a narrower emission filter 679/41 nm. This optical setup enabled crosstalk-free imaging of mCardinal and emiRFP2 (Figure 7g). The blue-shifted and red-shifted NIR FPs can further increase spectral multiplexing of fluorescence imaging in combination with other common FPs.

### 3 | DISCUSSION

Here, we described a practical and simple approach for rapid optimization of FPs via directed molecular evolution in mammalian cells. The presented method does not require generation of stable cell lines or viral vector preparation and can be carried out using commonly available high-throughput cell sorting hardware such as FACS. The simplicity and accessibility of the presented method

are the major advantages compared to previously described methods, which either involve the generation of stable lines<sup>29,45</sup> or require custom-built and/or expensive hardware such as high precision DMDs,<sup>46</sup> microfluidic sorters,<sup>47</sup> and micraft arrays.<sup>48</sup> Furthermore, the present method is also faster than alternative ones. One iteration of directed evolution accomplished with selecting individual variants can be performed in a course of  $\sim 8$  days, which is on average five to eight times faster than alternative methods involving stable cell line generation<sup>26,28,29</sup> or viral vector production.<sup>27,47,49</sup> The faster timeline compared to previously reported strategies was mainly achieved due to the utilization of transient transfection of large gene libraries in combination with fast target gene recovery from small pools of collected cells ( $\sim 50$ – $100$  cells). Besides a faster timeline, transient transfection provides additional advantages over traditional methods of gene library expression. It enables the evolution of proteins for which it is impossible or very hard to establish stable cell lines, such as, for example, opsins.<sup>33</sup> In addition, it is more accessible than other methods as it only requires very standard expression vectors containing SV40ori, widely used HEK293T cells, and cheap transfection reagents. Besides clear advantages, there are also other experimental considerations that should be taken into account. Conditions of gene delivery are optimized so that most of the transfected cells ( $\sim 50\%$ ) receive a single gene copy at total transfection efficiency of about

~5%, which is similar to lentiviral transduction at a multiplicity of infection of ~0.1 that also used single gene copy delivery.<sup>27,50,51</sup> Therefore, in the case of large libraries, it may significantly extend screening time as 95% of cells are negative. To reduce screening time using instrumental methods, negative cells can be removed using selective markers of choice, which can be included in the expression vector, although it will extend the timeline as antibiotic selection usually takes 3–4 days or so. Furthermore, about 50% of transfected cells contain more than one gene,<sup>33</sup> which correspondingly requires screening of about twice more individual clones (Steps 7–9 of workflow in Figure 1a) than cells that were collected during screening in order to ensure coverage of all diversity of recovered genes. However, it should be noted that lentiviral vectors provide similar efficiency of single gene copy delivery as gene delivery via calcium phosphate transfection, and viral transduction is a stochastic process with the same probability distribution. The developed method was tested only in HEK293T cells because this cell line provides sufficient expression level upon single-copy plasmid transfection due to episomal replication of DNA facilitated by the SV40*ori* sequence in the expression vector. Besides HEK293T, there are several other commercially available cell lines expressing the large T antigen required for episomal replication (available via ATCC) as well as self-replicating plasmids,<sup>52</sup> which can provide a high enough expression level under required transfection conditions. However, applicability of other cell lines and expression vectors remains to be tested and verified. In this study, we utilized FACS for high-throughput screening of mammalian cells, however, any other method for fluorescence cell screening can be integrated into the described workflow. For example, the SPOTlight approach,<sup>46</sup> microfluidic sorter,<sup>47</sup> photostick technique,<sup>53</sup> or robotic cell picker<sup>33</sup> can be implemented following FACS enrichment of transfected cells, thus enabling phenotyping of single cells for a diversity of biochemical and photophysical properties of expressed mutated proteins, for example, subcellular localization, pH stability, photostability, quantum yield, and so forth.

The developed approach was validated by enhancing intracellular brightness of green and NIR FPs derived from evolutionary distinct naturally occurring chromoproteins. The starting template proteins, phiLOV2.1, TagRFP657, and miRFP, which were already highly optimized via directed molecular evolution either in *E. coli* or in mammalian cells, were significantly enhanced just in two iterations of directed evolution. We demonstrated that oligomerization or alternation of fluorescence spectra cannot account for the increase in intracellular brightness. In the case of TagRFP657, the enhancement of intracellular brightness can be most

likely explained by improvements in chromophore maturation as TagRFP658 is characterized by a higher extinction coefficient compared to its precursor. However, for phiLOV2.1 and miRFP, the increase in intracellular brightness does not correspond to changes in molecular brightness. In contrast to GFP-like FPs, like TagRFP657, phiLOV2.1 and miRFP incorporate exogenous chromophores available in the cytoplasm of mammalian cells. Therefore, their intracellular brightness, besides of course intrinsic photophysical properties, will depend on efficiency of chromophore incorporation and chromophore availability. We revealed that miRFP2 has a higher affinity toward BV than miRFP (Supplementary Figure 8); however, it is hard to precisely quantify for which fraction of intracellular brightness enhancement the increase in affinity is responsible. The exact molecular basis of intracellular brightness increase observed for miRFP2 and phiLOV3 will require further investigations, which are beyond the scope of the present study. Overall, this study provided additional evidence, aligning with many previous reports, that relative molecular brightness of FPs expressed in *E. coli* does not always match intracellular brightness and the observed difference is particularly pronounced for proteins that require exogenous chromophores. The presented results also justify the need for optimization of FPs in mammalian cells for higher performance in vivo.

To date, only four FPs were reported to be developed via directed molecular evolution in mammalian cells. Three of them are GFP-like FPs, mPlum,<sup>26</sup> Kriek,<sup>27</sup> and mCrispRed,<sup>28</sup> and one is the BphP-based NIR FP miRFP.<sup>33</sup> A major rationale for evolving FPs in mammalian cells is to optimize their performance for in vivo applications; however, only mPlum and miRFP were tested in vivo. Crucially, side-by-side comparison of mPlum with other GFP-like FPs did not reveal any advantages of mPlum for in vivo imaging,<sup>54,55</sup> while miRFP was not compared to other BphP-based FPs either in cultured cells or in vivo. Therefore, it is important to demonstrate that the developed directed evolution approach is capable of yielding practical FPs that outperform other FPs developed in *E. coli*. First, assessment of the developed FPs in HEK cells showed miRFP2 has comparable intracellular brightness to that of the best performing BphP-based NIR FPs reported to date, such as iRFP670, miRFP680, and iRFP682. However, it should be noted that iRFP670, miRFP680, and iRFP682 although evolved in *E. coli* were actually selected in HeLa cells.<sup>36,37</sup> In turn, rationally enhanced version of miRFP2 significantly outperformed other BphP-based NIR FPs in HEK cells and can be considered as the brightest BphP-based NIR FP in HEK cells (Table 2). Thus, the presented method enabled development of one of the best



performing NIR FPs in HEK cells in the class in just five rounds of directed evolution (three rounds to select miRFP and two rounds to select miRFP2). However, further assessment of miRFP2 in vivo in model organisms reveals its inconsistent performance across various preparations. For example, miRFP2 co-expressed with HO1 was brighter mCardinal in fruit fly larvae, while dimmer in adult flies (Figure 5i,j). Fluorescence of miRFP2 in *C. elegans* can be enabled only by HO1 co-expression. The intracellular brightness of emiRFP2 was 1.5-fold dimmer than mCardinal in cultured neurons, but about 4.8-fold dimmer in brain tissue (Figures 4g and 5c). Furthermore, emiRFP2 was brighter than TagRFP658 in cultured neurons, while dimmer in zebrafish (Figures 4g and 6d). We also observed significantly greater variation of Cy5-to-green fluorescence ratio for miRFP2 expressed in neurons in vivo in mice in comparison to cultured neurons that is most likely due to higher variability of BV concentration in vivo. We suggest that high variability of miRFP2 performance in vivo in model organisms is determined by heme metabolism, which is required for BV production. For example, in wild-type *C. elegans*, heme utilization is not going via the BV intermediate as in vertebrates,<sup>56,57</sup> and thus expression of miRFP2 itself did not result in NIR fluorescence. However, we demonstrated that co-expression of HO1 is sufficient to enable BV production in *C. elegans*. Similarly, HO1 co-expression in *Drosophila* boosted fluorescence intensity of BphP-based NIR FPs expressed in neurons. Overall, these results indicated that high intracellular brightness of BphP-based NIR FPs in cultured cells may not be retained in vivo in model organisms, which complicates selection of the right NIR FP for in vivo applications and thus may require testing several candidates in real settings.

## 4 | MATERIALS AND METHODS

### 4.1 | Molecular cloning and mutagenesis

The mammalian codon-optimized genes of phiLOV2.1 and UnaG were synthesized de novo by GenScript based on the amino acid sequences reported in the original publications.<sup>4,31</sup> The TagRFP657 and miRFP genes were acquired from Addgene (TagRFP657 plasmid#31959; miRFP plasmid#108409). Synthetic DNA oligonucleotides used for cloning were purchased from Integrated DNA Technologies. PrimeStar Max master mix (Clontech) was used for high-fidelity PCR amplifications. Restriction endonucleases were purchased from New England BioLabs and used according to the manufacturer's protocols. Ligations were performed using T4

DNA ligase (Fermentas) or InFusion HD kits (Clontech). Small-scale isolation of plasmid DNA was performed with Mini-Prep kits (Qiagen); large-scale DNA plasmid purification was done with GenElute HP Endotoxin-Free Plasmid Maxiprep Kits (Sigma-Aldrich). Random libraries of mutants were prepared using Mutazyme II DNA polymerase (Agilent) under high mutation rate conditions (9–16 mutations per kilobase pair) and subcloned into the pN1 vector (Clontech). Obtained gene libraries in expression vectors were electroporated into NEB10- $\beta$  *E. coli* host cells (New England BioLabs). Serial dilutions ( $10^{-4}$  and  $10^{-5}$ ) of the electroporated cells were plated on LB/agar medium supplemented with  $100 \text{ mg ml}^{-1}$  kanamycin to estimate electroporation and cloning efficiency. For each library, 20 randomly selected clones were sequenced to confirm ligation efficiency, that is, fraction of the clones containing target genes, and mutation rate. The remainder of the cells was grown overnight in LB medium supplemented with  $100 \text{ mg ml}^{-1}$  of kanamycin for subsequent plasmid DNA isolation. Library transfection into HEK293FT cells (Invitrogen) was performed as described previously. Briefly, HEK293FT cells were maintained between 10 and 70% confluence at  $37^\circ\text{C}$  with 5%  $\text{CO}_2$  in DMEM medium (Cellgro) supplemented with 10% heat-inactivated FBS (Corning), 1% penicillin/streptomycin (Cellgro), and 1% sodium pyruvate (BioWhittaker). Transfection of HEK293FT cells with gene libraries was performed using a commercially available calcium phosphate (CaPhos) transfection kit (LifeTechnologies) according to the modified protocol using the pUC19 plasmid as “dummy” DNA in weight ratio library DNA: pUC19 1:100. To sort the gene library-transfected HEK293FT cells using flow cytometry, cells were harvested from a culture dish  $\sim 48$  hr after gene library transfection by applying trypsin for 5–10 min (Cellgro) and then washed twice by centrifuging the cell suspension for 5 min at 500 rpm and resuspending cells in PBS (Cellgro). The washed cells were then resuspended in PBS supplemented with 4% FBS (Corning) and 10 mM EDTA at a density of  $1\text{--}2 \cdot 10^6$  cells/ml and filtered through a  $30\text{-}\mu\text{m}$  filter (Falcon) to prevent clogging on the FACS machine. The filtered cells were sorted by FACSAria (BD Biosciences) running BD FACS Diva 8.0 software and equipped with standard 488- and 640-nm solid-state lasers. Debris, dead cells, and cell aggregates were gated out using forward and side scatter before desired fluorescence signals were detected. For each library, several hundred cells exhibiting highest fluorescent intensity in the corresponding channel were collected and subjected to whole genome amplification (WGA) using a commercially available WGA kit (New England BioLabs) followed by PCR amplification. Amplicons with a size corresponding to that of the target

gene were purified by agarose gel electrophoresis and cloned into the pN1 vector. Obtained colonies were individually picked to confirm the correct insert using PCR with a pair of primers that anneal to CMV promoter and the target genes. Colonies with correct insert were cultured in LB with 100 mg ml<sup>-1</sup> of kanamycin for plasmid purification. Purified plasmids were individually transfected into HEK293FT cells using TransIT-X2 reagent (Mirus). In 24 hr post transfection, HEK cells were imaged using a wide-field Nikon Eclipse Ti equipped with 10× NA 0.45 and 20× NA 0.75 objective lenses, a SPECTRA-X light engine (Lumencor) with 475/28 nm, and 631/28 nm exciters (Semrock), a 5.5 Zyla camera (Andor), and automated stage (Ludl), controlled by NIS-Elements AR software Nikon to assess fluorescence and photostability.

The genes of miRFP720 and miRFP670nano were de novo synthesized by GenScript using sequences reported in the original publications.<sup>19,39</sup> The genes of mCardinal, mMaroon, and miRFP703 were acquired from Addgene as plasmids #54590, #54554, and #79988, respectively. To construct  $\alpha$ -tubulin and  $\beta$ -actin fusions, the phiLOV3, TagRFP658, and miRFP2 genes were PCR amplified and swapped with mClover2 in pmClover2-tubulin-C-18 plasmid (Addgene #56376) and with TagRFP675 in pTagRFP675-actin-C1 (Addgene #44277) using InFusion cloning (Clontech). To construct Keratin fusions, emiRFP2 were PCR amplified and swapped with miRFP670nano in pKeratin-miRFP670nano plasmid (Addgene #127437) using Fusion cloning (Takara). To express TagRFP658 and mCardinal in neurons under CaMKII promoter the corresponding genes were PCR amplified and swapped with the Arch-GFP gene in FCK-Arch-GFP (Addgene #22217) using InFusion cloning. The pAAV-CAG-miRFP720-P2A-GFP and pAAV-CAG-miRFP2-P2A plasmids were constructed by cloning miRFP720-P2A and miRFP2-P2A into pAAV-CAG-GFP (Addgene #37825) in frame with GFP.

For expression in zebrafish larvae, the genes of mCardinal, TagRFP678, miRFP, and miRFP2 were codon-optimized using the online resource at <http://www.bioinformatics.org/>, synthesized de novo by GenScript. The zebrafish codon optimized emiRFP2 gene was cloned by substituting the nucleotides encoding RpBphP1-based N-terminus of zebrafish codon optimized miRFP2 (aa 2–19) with those encoding 13 amino acid preceding the chromophore-binding Cys in RpBphP2 as previously reported.<sup>20</sup>

For co-expression of TagRFP658, miRFP2, and emiRFP2 with mClover3 in NIH3T3 mouse embryonic fibroblasts (DSMZ) and PAC2 zebrafish embryonic fibroblasts, the zebrafish codon-optimized genes of the corresponding NIR FPs and mClover3 were cloned into

an expression vector carrying the bidirectional promoter (p-EF1a:2xCMV:EF1a) consisting of two CpG free mCMV enhancers (Invivogen) each followed by CpG free hEF1alpha promoter (Invivogen) and SV40 polyA sequences. The generated expression cassettes were flanked by Tol2 transposon arms. The p-EF1a:2xCMV:EF1a bidirectional vector allows for expression of two reporter genes with comparable expression levels. For pan-neuronal expression in zebrafish, the zebrafish codon optimized gene of TagRFP658 was cloned into pTol2-UAS-zArchon1-KGC-GFP-ER2 plasmid (Addgene #108427) described before<sup>33</sup> by swapping zArchon1-KGC-GFP-ER2 using SpeI and AscI sites.

For transient expression in cerebellar Purkinje neurons in zebrafish, EF1a:2xCMV:EF1a promoter in the above mentioned bidirectional vector carrying mClover3 and H2B-NIR FP, either of mCardinal, TagRFP658, or emiRFP2 was replaced by Purkinje neuron-specific bidirectional promoter (E1b:4xcpce:E1b) enabling co-expression of the corresponding NIR FPs with mClover3 predominantly in zebrafish cerebellar PCs while also inducing slight expression in tectal neurons of larval zebrafish.<sup>44</sup>

For generation of transgenic flies, the mCardinal, mMarron, iRFP-VC, iRFP-VC-T2A-HO1, miRFP2, and miRFP2-T2A-HO1 genes were codon-optimized for expression in *D. melanogaster* using the online resource at <http://www.bioinformatics.org/>, synthesized de novo and cloned into the 20XUAS-IVS-Jaws-mVenus\_tr plasmid (Addgene #111553) by swapping the Jaws-mVenus gene.

For expression in *C. elegans*, the genes of miRFP2, miRFP2-T2A-HO1, and jGCaMP7s were codon-optimized using SnapGene codon-optimization tool, synthesized de novo by GenScript and cloned into an expression vector under the tag-168 promoter the drives pan-neuronal expression.

## 4.2 | Protein purification and in vitro characterization

For protein purification, the phiLOV2.1, phiLOV3, TagRFP657, and TagRFP658 genes were cloned into pBAD-HisD vector and transformed into TOP10 cells. To express the protein, bacterial cells were grown in RM medium supplemented with ampicillin and 0.002% arabinose for 15–18 hr at 37°C followed by 24 hr at 18°C. Proteins were purified using TALON Metal Affinity Resin (Clontech) and dialyzed overnight against PBS buffer, pH 7.4. To express miRFP720 and miRFP2 the genes were cloned in pBAD-HisB plasmid and transformed into BW25113 bacterial cells, containing pWA23h plasmid

encoding HO1 under rhamnose promoter. The colonies were grown in LB medium supplemented with ampicillin, kanamycin, 0.002% arabinose, and 0.02% rhamnose for 20 hr at 37°C. For protein purification, Ni-NTA agarose (Qiagen) was used. Protein was eluted with PBS containing 100 mM EDTA, followed by dialysis overnight against PBS buffer, pH 7.4.

Spectral properties of the proteins were measured in PBS at pH 7.4. The fluorescence and absorption spectra of phiLOV2.1, phiLOV3, TagRFP657, and TagRFP658 were measured using a Fluorolog3 spectrofluorometer (Jobin Yvon) and a Lambda 35 UV/Vis spectrometer (PerkinElmer), respectively. The fluorescence and absorption spectra of miRFP720 and miRFP2 were measured with a CM2203 spectrofluorometer (Solar, Belarus) and NanoDrop 2000c spectrophotometer (Thermo Scientific), respectively. The extinction coefficient of phiLOV2.1 and phiLOV3 was determined as a ratio between the absorbance value of the peak at 374 nm, which correspond to FMN absorbance with extinction coefficient of  $12,500 \text{ M}^{-1} \text{ cm}^{-1}$ ,<sup>58</sup> and the value of the peak at major band peaked at  $\sim 476 \text{ nm}$ . For determination of the quantum yield, integrated fluorescence of phiLOV3 was compared with that of equally absorbing phiLOV2.1, characterized by quantum yield value of 0.2.<sup>31</sup> To determine extinction coefficients of TagRFP657 and TagRFP658, we relied on measuring the mature chromophore concentrations, using alkali-denatured proteins as previously described.<sup>32</sup> For determination of the quantum yield, integrated fluorescence of TagRFP658 was compared with that of equally absorbing TagRFP657, characterized by quantum yield value of 0.1.<sup>32</sup> The extinction coefficient of miRFP720 and miRFP2 was determined as a ratio between the absorbance value of the peak at Q-band and the value of the peak at Soret band, characterized by extinction coefficient of  $39,900 \text{ M}^{-1} \text{ cm}^{-1}$ .<sup>8</sup> For quantum yield determination, the integrated fluorescence values of purified miRFP2 were compared with equally absorbing miRFP720 (quantum yield 0.061). pH titrations for phiLOV2.1, phiLOV3, TagRFP657, and TagRFP658 were performed in a 96-well black clear bottom plate (Corning) by 1:20 dilution with a series of commercially available pH buffers (HYDRION) using a SpectraMax-M5 plate reader (Molecular Devices) to read out fluorescence. To determine pH stability, miRFP2 and miRFP720 were diluted 1:100 into a series of home-made pH adjusted buffers with NaOH and HCl (30 mM citric acid, 30 mM borax, or 30 mM NaCl) with pH values ranging from 3 to 10 in 0.5 pH units interval in a 96-well black clear bottom plate (Thermo Scientific) using a Modulus II Microplate Reader (Turner Biosystems) to read out fluorescence using 625 nm excitation and 660–720 nm emission filters.

Size-exclusion chromatography for phiLOV3 and TagRFP658 was performed by GenScript on a Superdex 200 10/300 GL column (GE Healthcare Life Sciences) using a gel filtration standard (#1511901; BIO-RAD). Size-exclusion chromatography for miRFP720 and miRFP2 was performed with a Superdex 200 10/300 GL column using GE AKTA Explorer (Amersham Pharmacia, UK) FPLC System.

Two-photon spectrum and cross sections of TagRFP658 were measured in PBS buffer at concentrations  $\sim 1\text{--}5 \cdot 10^{-5} \text{ M}$  in 1 mm glass spectroscopy cuvettes (Starna cells) using an MOM two-photon fluorescent microscope (Sutter Instrument) coupled with an Insight DeepSee (Newport) femtosecond laser, as described before.<sup>33</sup> For the measurement of spectral shape, fluorescence was collected through a combination of 694SP (Semrock) and 630/60 (Chroma) filters for TagRFP658 and through HQ705/100 (Chroma) filter for Styryl 9-M dye (Aldrich) in chloroform used as a reference standard. The two-photon cross section was measured relatively to LDS 698 dye (Exciton) in chloroform<sup>54</sup> at 1,150 nm using a combination of HQ705/100 (Chroma) and 630/60 (Chroma) filters and relatively to Rhodamine 700 in ethanol<sup>59</sup> at 1,200 nm using 675/20 (Chroma) filter. The spectral shape was then scaled to these values. The differences between the peak absolute values obtained with two different standards were within 10%.

### 4.3 | Protein characterization in mammalian cells

All procedures involving animals at MIT and Westlake University were conducted in accordance with the U.S. National Institutes of Health Guide for the Care and Use of Laboratory Animals and approved by the Massachusetts Institute of Technology or Westlake University Committee on Animal Care. HEK293FT (Invitrogen) and HeLa (ATCC CCL-2) cells were maintained between 10 and 70% confluence at 37°C with 5% CO<sub>2</sub> in DMEM medium (Cellgro) supplemented with 10% heat-inactivated FBS (Corning), 1% penicillin/streptomycin (Cellgro), and 1% sodium pyruvate (BioWhittaker). Cells were authenticated by the manufacturer and tested for mycoplasma contamination to their standard levels of stringency and were here used because they are common cell lines for testing new tools. Primary mouse hippocampal neuronal culture was prepared as described previously. HEK293FT and HeLa cells were transiently transfected using TransIT-X2 (Mirus Bio LLC) or calcium phosphate transfection kit (K278001, Invitrogen) according to the manufacturer's protocol and imaged 24–48 hr after transfection. Mouse primary hippocampal

neurons were prepared from postnatal day 0 or 1 Swiss Webster (Taconic) mice (both male and female mice were used) and cultured as previously described.<sup>33</sup> Neuronal cultures were transfected using calcium phosphate transfection kit (Life Sciences) according to the protocol described before.<sup>33</sup> Transduction of neuronal culture was done at four DIV using  $\sim 10^9$  viral particles of rAAV8-hSyn-miRFP2 (Janelia Farm Viral Core, the rAAV genome titer was determined by dot blot) per well of standard 24-well plate (Corning). Imaging of HEK293FT cells and neuronal cultures for Figures 1–3 was done using a Nikon Eclipse Ti inverted microscope equipped with a SPECTRA X light engine (Lumencor) with 475/28 nm and 631/28 nm exciters (Semrock), and a 5.5 Zyla camera (Andor), controlled by NIS-Elements AR software, and using  $10 \times$  NA 0.3 and  $40 \times$  NA 1.15 objective lenses. Imaging of HEK293FT cells for Figure 3 was done under using a Nikon Eclipse Ti2-E inverted microscope equipped with a SPECTRA III light engine (Lumencor) with 475/28 nm and 635/22 nm exciters (Semrock), a 680 nm solid-state laser (MRL-III-680–800 mW, CNI Laser) and 680/13 nm exciter (Semrock), and a Orca Flash4.0v3 camera (Hamamatsu), controlled by NIS-Elements AR software, and using  $20 \times$  NA 0.75 objective lenses. To perform fair comparison of fluorescence intensity in Cy5 (excitation 635/22 nm from 637 nm laser, emission 730/140 nm) and Cy5.5 (excitation 680/13 nm from 680 nm laser, emission 710 LP) channels, images were acquired under matching excitation power ( $66 \text{ mW/mm}^2$ ) and the same exposure time (100 ms).

NIH3T3 cells were grown in DMEM (4.5  $\mu\text{g/L}$  glucose) supplemented with 10% FCS, and maintained in a 5%  $\text{CO}_2$  incubator at  $37^\circ\text{C}$ . The NIH3T3 stable cell line was generated by the transfection of pEF-TGN46-mScarlet-iresPuro, followed by isolation and expansion of a clone grown in DMEM containing Puro-mycin at the concentration of 2  $\mu\text{g/ml}$ . *Zebrafish PAC2 fibroblast cells* were cultured in L-15 medium supplemented with 15% FCS, and maintained in an incubator at  $28^\circ\text{C}$  and atmospheric  $\text{CO}_2$ . Transfection of plasmids was performed onto cells grown on  $\mu$ -slide 4-well glass bottom dish (ibidi) using jetPRIME reagent (Polyplus) according to the manufacturer's protocol. For PAC2 cells, transfection was carried out in L-15 medium supplemented with 5% FCS. Then, 6 hr after transfection, an equal volume of 15% FCS containing medium was added to adjust its final concentration of FCS at 10%. Live NIH3T3 and PAC2 cells were imaged 48 hr post transfection using a laser scanning confocal microscope (TCS SP8, Leica Microsystems, Germany) with  $40 \times$  water (NA 1.1) objective. The BV solution (Sigma) at the final concentration 25  $\mu\text{M}$  was added to PAC2 cells 48 hr after

transfection, followed by recording of their images 3 hr after the BV administration. mClover3 was excited by an argon laser at 488 nm and detection range at 495–530 nm, whereas each of NIR RFs was excited by HeNe 633 laser and detection range at 660–785 nm. For three-color imaging, transfected NIH3T3 cells were fixed with 4%PFA 36 hr after transfection, followed by imaging using a confocal microscope. mScarlet fluorescence was excited by a DPSS 561 nm laser with the detection range at 570–600 nm, followed by simultaneous imaging of mClover3 and the NIR FPs as described above.

#### 4.4 | In utero electroporation and acute slice preparation

Embryonic day (E) 15.5 timed-pregnant female Swiss Webster mice (Taconic) were used for in utero electroporation as described before.<sup>33</sup> The pAAV-CAG-emiRFP2-P2A-GFP or pAAV-CAG-mCardinal-P2A-GFP plasmid at  $\sim 1 \mu\text{g}/\mu\text{l}$  concentration were injected into the lateral ventricle of one cerebral hemisphere of an embryo. Acute brain slices were obtained from mice at P20–30 without regard for sex using standard techniques as described before.<sup>33</sup> The brain slices were imaged using a Nikon Eclipse Ti2-E inverted microscope equipped with a SPECTRA III light engine (Lumencor) with 475/28 nm and 635/22 nm exciters (Semrock), and a Orca Flash4.0v3 camera (Hamamatsu), controlled by NIS-Elements AR software, and using  $10 \times$  NA 0.45 and  $20 \times$  NA 0.75 objective lenses.

#### 4.5 | Zebrafish larvae preparation and imaging

All experiments involving zebrafish at MIT and Technische Universität Braunschweig were conducted in accordance with protocols approved by Massachusetts Institute of Technology Committee on Animal Care following guidelines described in the U.S. National Institutes of Health *Guide for the Care and Use of Laboratory Animals* or by German legislation following European Union guidelines (EU Directive 2010\_63) according to location of the respective experiments. Zebrafish larvae expressing TagRFP658 were prepared and imaged as described previously.<sup>33</sup> Briefly, pTol2-UAS-zTagRFP658 plasmid was co-injected with Tol2 transposase mRNA into homozygous nacre embryos of the pan-neuronal expressing Gal4 line, tg(elav13:GAL4-VP16)nns6.<sup>60</sup> Injected larvae were screened for NIR fluorescence in the brain and spinal cord at 2–3 days postfertilization (dpf) using the Nikon wide-field microscope described above

and subsequently imaged at 3–4 dpf using Zeiss Lightsheet Z.1 microscope. Lightsheets were generated by two illumination objectives (10 $\times$ , NA 0.2), and the fluorescence signal detected by a 20 $\times$  water immersion objective (NA 1.0). The laser line used for excitation was 638 nm. Optical filters used to separate and clean the fluorescence response included a Chroma T647lpxr as a dichroic, and a Chroma ET665lp. Tiled data sets were taken with the Zeiss ZEN Software, and subsequently merged and processed with FIJI, and Arivis Vision4D.

To co-express the mCardinal, TagRFP658, and emiRFP2 H2B fusions with mClover3, one cell stage embryos of the pigmentation-compromised zebrafish *brass* strain and *Tg(4xen.cpce-E1B:gap-mScarlet)* line were microinjected with the corresponding E1b:4xcpce: E1b expression vector described above together with *tol2* mRNA (1.5 nl of injection mix containing 25 ng/ $\mu$ l of both *tol2* and pTol2-plasmid). mClover3 positive larval fish at 4 dpf were subjected to confocal microscopy analysis. The stable transgenic line, *Tg(4xen.cpce-E1B: gap: mScarlet)* was generated by the injection of *tol1* mRNA together with a Toll1-reporter plasmid in which membrane-targeted mScarlet expression regulated by 4xcpce:E1b is restrictively induced in PCs, because ectopically expressed mScarlet in tectal neurons was eliminated by 4 $\times$  miRNA181a target sequence inserted into the 3'UTR of the reporter gene.<sup>44</sup> For membrane targeting, mScarlet was fused N-terminally to the first 20 amino acids encoded by the zebrafish *gap43* gene. Fish larvae exhibiting no fluorescence in the corresponding channels were excluded from further imaging. Fluorescence imaging was performed using a laser scanning confocal microscope (TCS SP8, Leica Microsystems, Germany) with 40 $\times$  water (NA 1.1) objectives. mClover3 (excited by an argon laser at 488 nm, detection range at 495–530 nm) and each of NIR FPs (excited by HeNe 633 nm laser, detection range at 660–785 nm) expressed in PCs were imaged simultaneously. For three-color imaging, first, a DPSS 561 nm laser was used to excite mScarlet fluorescence, which was detected with the range at 570–600 nm, followed by simultaneous imaging of mClover3 and the NIR FPs as described above. Reconstructions and projections from z-stacks of images were generated with the 3D-projection program included in the LAS X software (Leica Microsystems, Germany). Acquired images were processed with FIJI to measure the fluorescent intensity ratio of each of NIR FPs and mClover3 in each PC. In vivo photostability of TagRFP658, emiRFP2, and mCardinal was assessed using PCs continuously exposed to HeNe633 laser set at 70% of the laser power in the software setting. The region of interest (116.25  $\mu$ m  $\times$  116.25  $\mu$ m) was drawn encompassing several PCs labeled with nuclear localized

TagRFP658, or mCardinal, and single plan image (optical section: 3.56  $\mu$ m) was taken every 0.648 s without interval for 5 min. Acquired images were processed with FIJI to measure the fluorescent intensity in each PC in each time point.

#### 4.6 | Preparation and imaging in *drosophila*

Transgenic fly lines with the following genotypes *y1 w67c23; P{y[+t7.7] w[+mC] = UAS-mCardinal} attP40*, *y1 w67c23; P{y[+t7.7] w[+mC] = UAS-mMaroon} attP40*, *y1 w67c23; P{y[+t7.7] w[+mC] = UAS-iRFP-VC} attP40*, *y1 w67c23; P{y[+t7.7] w[+mC] = UAS-iRFP-VC-T2A-HO1} attP40*, *y1 w67c23; P{y[+t7.7] w[+mC] = UAS-miRFP2} attP40*, *y1 w67c23; P{y[+t7.7] w[+mC] = UAS-miRFP2-T2A-HO1} attP40* were generated by Bestgene using the user provided 20XUAS-IVS plasmids described above. Flies were raised on standard cornmeal medium at room temperature. To drive pan-neuronal protein expression, generated transgenic adult male flies were mated with C155-GAL4 (pan-neural, a gift from Littleton lab at MIT) virgin females to generate heterozygous (C155-GAL4/+ or Y; Transgene/+).

Intact third instar larva and 2- to 4-day old adult flies were immobilized on the coverslip for further imaging. To drive protein expression in specific neurons, generated transgenic flies were crossed with R84C10-GAL4 to generate heterozygous progenies. Dissected brains from 5- to 6-day old adult females were used for further imaging. No larvae or adult flies, carrying the genes of target proteins, were excluded from the study. The wide-field Nikon microscope described above was used to acquire low magnification images and a Zeiss LSM 800 confocal microscope with Airyscan equipped with 631 nm laser for excitation for high magnification images.

#### 4.7 | *C. elegans* preparation and imaging

Worms were cultured and maintained following standard protocols.<sup>61</sup> Transgenic worms with extrachromosomal array co-expressing miRFP2 or miRFP2-T2A-HO1 with jGCaMP7s pan-neuronally were generated by co-injection of the two plasmids tag-168::miRFP2 or tag-168::miRFP2-T2A-HO1 with tag-168::NLS-jGCaMP7s into N2 background worms as described before.<sup>62</sup> All plasmids were injected at 10 ng/ $\mu$ l. Worms exhibiting the highest green fluorescence in neurons were picked and mounted on 2% agarose pads on microscope slides, worms without green fluorescence were excluded from further imaging, immobilized with 5 mM tetramisole,

covered by a coverslip, and imaged using a Zeiss LSM 800 confocal microscope with Airyscan equipped with 631 nm laser for excitation.

#### 4.8 | Data analysis and statistics

Data were analyzed offline using NIS-Elements Advance Research software, Excel (Microsoft), OriginPro, ImageJ, the Microscope online application (<https://www.fpbase.org/microscope>), and Arivis Vision4D. Data collection for fluorescence spectra of NIR FPs was done using <https://www.fpbase.org>. All attempts at replication of the experiments were successful. We did not perform a power analysis, since our goal was to create a new technology; and as recommended by the NIH, “In experiments based on the success or failure of a desired goal, the number of animals required is difficult to estimate...” As noted in the aforementioned paper, “The number of animals required is usually estimated by experience instead of by any formal statistical calculation, although the procedures will be terminated [when the goal is achieved].” These numbers reflect our past experience in developing neurotechnologies. All attempts at replication of the experiments were successful. No randomization was used in the study. No blinding was used in the study.

#### ACKNOWLEDGMENTS

The authors thank B. Trout and C. Sudrik from MIT for help with spectroscopic analysis of phiLOV3 and TagRFP658. The authors also thank Shoh Asano from MIT for the help with zebrafish imaging, Cuiyun Bu from TU Braunschweig for the help with zebrafish injections, Shaofeng Wu from Westlake University for the help with transgenic worms preparation, and Yi Sun and Baoyue Liu from Westlake University for the help with fruit flies imaging. We are grateful to S. Flavell from MIT for the *C. elegans* expression vector. This work was supported by start-up funding from the Foundation of Westlake University, National Natural Science Foundation of China grant 32050410298, 2020 BBRF Young Investigator Grant, and MRIC Funding 103536022023 to K. D. P., an internal grant of the National Research Center Kurchatov Institute No. 2752 of 28.10.2021 to F. V. S., RFBR grant No. 19-04-00395 to O. M. S., RSF grant No. 17-14-01256 to T. V. K., the German Research Foundation (DFG, K1949/7-2) Project 241961032 to R. W. K., by Lisa Yang, John Doerr, HHMI, and grants NIH R01DA029639, NIH R01MH12297101, NIH R01DA045549, NIH R01MH114031, NIH R43MH109332, NIH R01GM104948, and NSF Grant 1734870 grants to E. S. B., and NIH BRAIN program U24NS109107 grant to M. D.

#### CONFLICT OF INTEREST

The authors declare no potential conflict of interests.



#### AUTHOR CONTRIBUTIONS

**Kiryl D. Piatkevich** and **Edward S. Boyden**: Initiated the project and made high-level designs and plans. **Kiryl D. Piatkevich** and **Erica E. Jung**: Developed the rapid directed molecular evolution approach. **Kiryl D. Piatkevich** and **Erica E. Jung**: Developed phiLOV3 and TagRFP658. **Kiryl D. Piatkevich** and **Siranush Babakhanova**: Developed miRFP2. **Kazuhiko Namikawa**: Designed and cloned emiRFP2. **Kiryl D. Piatkevich**: Characterized phiLOV3 and TagRFP658 in vitro and together with **Erica E. Jung** and **Hanbin Zhang** performed their characterization in mammalian cells. **Mikhail Drobizhev**: Measured two-photon absorption properties in solution. **Lea Eisenhard** and **Kazuhiko Namikawa**: Performed protein characterization in NIH3T3 and PAC2 cells. **Erica E. Jung** and **Kiryl D. Piatkevich**: Performed electrophysiology experiments in cultured neurons. **Oksana M. Subach**, **Dmitry A. Korzhenevskiy**, **Tatiana V. Rakitina**, and **Fedor V. Subach**: Characterized miRFP2 in vitro. **Kiryl D. Piatkevich**, **Hanbin Zhang**, and **Siranush Babakhanova**: Characterized miRFP2 and emiRFP2 in mammalian cells. **Yangdong Wang**: Performed IUE. **Kiryl D. Piatkevich** and **Hanbin Zhang**: Performed acute brain slice imaging. **Kiryl D. Piatkevich**, **Erica E. Jung**, **Lea Eisenhard**, **Kazuhiko Namikawa**, and **Reinhard W. Köster**: Prepared zebrafish and performed their imaging. **Jing Shi** and **Hongyun Tang** prepared transgenic worms and **Xian Xiao** imaged them. **Kiryl D. Piatkevich**, **Siranush Babakhanova**, **Demian Park**, and **Wenjing Wang**: Prepared transgenic flies and performed their imaging. **Kiryl D. Piatkevich**: Performed the statistical analysis. **Kiryl D. Piatkevich**: Oversaw all aspects of the project, analyzed and interpreted the data, wrote the manuscript with contributions from all of the authors.

#### DATA AVAILABILITY STATEMENT

All other data generated or analyzed during this study are available from the corresponding authors on reasonable request. The nucleotide sequences of the reported proteins are available at Genbank at the following accession codes: MZ682637 phiLOV3; MZ682638 TagRFP658; MZ682639 miRFP2; MZ682640 emiRFP2.

#### ORCID

**Fedor V. Subach**  <https://orcid.org/0000-0003-2720-7821>  
**Kiryl D. Piatkevich**  <https://orcid.org/0000-0002-7777-9468>

## REFERENCES

- Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA. Fluorescent proteins and their applications in imaging living cells and tissues. *Phys Rev.*, 2010;p. 90, 1103–1163.
- Prasher DC, Eckenrode VK, Ward WW, Prendergast FG, Cormier MJ. Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene*. 1992;111:229–233.
- Alieva NO, Konzen KA, Field SF, et al. Diversity and evolution of coral fluorescent proteins. *PLoS One*. 2008;3:e2680.
- Kumagai A, Ando R, Miyatake H, et al. A bilirubin-inducible fluorescent protein from eel muscle. *Cell*. 2013;153:1602–1611.
- Losi A. Flavin-based blue-light photosensors: A photobiophysics update. *Photochem Photobiol*. 2007;83:1283–1300.
- Rottwinkel G, Oberpichler I, Lamparter T. Bathy phytochromes in rhizobial soil bacteria. *J Bacteriol*. 2010;192:5124–5133.
- Narikawa R, Nakajima T, Aono Y, et al. A biliverdin-binding cyanobacteriochrome from the chlorophyll d-bearing cyanobacterium *Acaryochloris marina*. *Sci Rep*. 2015;5:1–10.
- Filonov GS, Piatkevich KD, Ting L-M, Zhang J, Kim K, Verkhusha VV. Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat Biotechnol*. 2011;29:757–761.
- Piatkevich KD, Subach FV, Verkhusha VV. Far-red light photoactivatable near-infrared fluorescent proteins engineered from a bacterial phytochrome. *Nat Commun*. 2013;4:1–10.
- Rodriguez EA, Tran GN, Gross LA, et al. A far-red fluorescent protein evolved from a cyanobacterial phycobiliprotein. *Nat Methods*. 2016;13:763–769.
- Chu J, Haynes RD, Corbel SY, et al. Non-invasive intravital imaging of cellular differentiation with a bright red-excitabile fluorescent protein. *Nat Methods*. 2014;11:572–578.
- Shaner NC, Lambert GG, Chammas A, et al. A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nat Methods*. 2013;10:407–409.
- Ai H-W, Baird MA, Shen Y, Davidson MW, Campbell RE. Engineering and characterizing monomeric fluorescent proteins for live-cell imaging applications. *Nat Protoc*. 2014;9:910–928.
- Bindels DS, Postma M, Haarbosch L, van Weeren L, Gadella TWJ. Multiparameter screening method for developing optimized red-fluorescent proteins. *Nat Protoc*. 2020;15:450–478.
- Subach FV, Piatkevich KD, Verkhusha VV. Directed molecular evolution to design advanced red fluorescent proteins. *Nat Methods*. 2011;8:1019–1026.
- Piatkevich KD, Suk H-J, Kodandaramaiah SB, et al. Near-infrared fluorescent proteins engineered from bacterial phytochromes in neuroimaging. *Biophys J*. 2017;113:2299–2309.
- Heppert JK, Dickinson DJ, Pani AM, et al. Comparative assessment of fluorescent proteins for in vivo imaging in an animal model system. *Mol Biol Cell*. 2016;27:3385–3394.
- Botman D, de Groot DH, Schmidt P, Goedhart J, Teusink B. In vivo characterisation of fluorescent proteins in budding yeast. *Sci Rep*. 2019;9:1–14.
- Oliinyk OS, Shemetov AA, Pletnev S, Shcherbakova DM, Verkhusha VV. Smallest near-infrared fluorescent protein evolved from cyanobacteriochrome as versatile tag for spectral multiplexing. *Nat Commun*. 2019;10:1–13.
- Matlashov ME, Shcherbakova DM, Alvelid J, et al. A set of monomeric near-infrared fluorescent proteins for multicolor imaging across scales. *Nat Commun*. 2020;11:1–12.
- Piatkevich KD, Subach FV, Verkhusha VV. Engineering of bacterial phytochromes for near-infrared imaging, sensing, and light-control in mammals. *Chem Soc Rev*. 2013;42:3441–3452.
- Pourmir A, Johannes TW. Directed evolution: Selection of the host organism. *Comput Struct Biotechnol J*. 2012;2:e201209012.
- Hendel SJ, Shoulders MD. Directed evolution in mammalian cells. *Nat Methods*. 2021;18:346–357.
- Manna P, Hung ST, Mukherjee S, et al. Directed evolution of excited state lifetime and brightness in FusionRed using a microfluidic sorter. *Integr Biol*. 2018;10:516–526.
- Arakawa H, Kudo H, Batrak V, et al. Protein evolution by hypermutation and selection in the B cell line DT40. *Nucleic Acids Res*. 2008;36:1–11.
- Wang L, Jackson WC, Steinbach PA, Tsien RY. Evolution of new nonantibody proteins via iterative somatic hypermutation. *Proc Natl Acad Sci U S A*. 2004;101:16745–16749.
- Dean KM, Lubbeck JL, Davis LM, et al. Microfluidics-based selection of red-fluorescent proteins with decreased rates of photobleaching. *Integr Biol*. 2015;7:263–273.
- Erdogan M, Fabritius A, Basquin J, Griesbeck O. Targeted in situ protein diversification and intra-organelle validation in mammalian cells. *Cell Chem Biol*. 2020;27:610–621.e5.
- Chen H, Liu S, Padula S, et al. Efficient, continuous mutagenesis in human cells using a pseudo-random DNA editor. *Nat Biotechnol*. 2019;38:165–168.
- English JG, Olsen RHJ, Lansu K, et al. VEGAS as a platform for facile directed evolution in mammalian cells. *Cell*. 2019;178:748–761.e17.
- Christie JM, Hitomi K, Arvai AS, et al. Structural tuning of the fluorescent protein iLOV for improved photostability. *J Biol Chem*. 2012;287:22295–22304.
- Morozova KS, Piatkevich KD, Gould TJ, Zhang J, Bewersdorf J, Verkhusha VV. Far-red fluorescent protein excitable with red lasers for flow cytometry and superresolution STED nanoscopy. *Biophys J*. 2010;99:L13–L15.
- Piatkevich KD, Jung EE, Straub C, et al. A robotic multi-dimensional directed evolution approach applied to fluorescent voltage reporters article. *Nat Chem Biol*. 2018;14:352–360.
- Mahon MJ. Vectors bicistronically linking a gene of interest to the SV40 large T antigen in combination with the SV40 origin of replication enhance transient protein expression and luciferase reporter activity. *Biotechniques*. 2011;51:119–126.
- Piatkevich KD, Jung EE, Straub C, et al. A robotic multi-dimensional directed evolution approach applied to fluorescent voltage reporters. *Nat Chem Biol*. 2018;14:352–360.
- Shcherbakova DM, Balaban M, Emelyanov AV, Brenowitz M, Guo P, Verkhusha VV. Bright monomeric near-infrared fluorescent proteins as tags and biosensors for multiscale imaging. *Nat Commun*. 2016;7:1–12.
- Shcherbakova DM, Verkhusha VV. Near-infrared fluorescent proteins for multicolor in vivo imaging. *Nat Methods*. 2013;10:751–754.
- Rogers OC, Johnson DM, Firnberg E. mRhubarb: Engineering of monomeric, red-shifted, and brighter variants of iRFP using structure-guided multi-site mutagenesis. *Sci Rep*. 2019;9:1–8.

39. Shcherbakova DM, Cox Cammer N, Huisman TM, Verkhusha VV, Hodgson L. Direct multiplex imaging and optogenetics of Rho GTPases enabled by near-infrared FRET article. *Nat Chem Biol.* 2018;14:591–600.
40. Kim JH, Lee SR, Li LH, et al. High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One.* 2011;6:1–8.
41. Bajar BT, Lam AJ, Badiie RK, et al. Fluorescent indicators for simultaneous reporting of all four cell cycle phases. *Nat Methods.* 2016;13:993–996.
42. Yu D, Baird MA, Allen JR, et al. A naturally monomeric infrared fluorescent protein for protein labeling in vivo. *Nat Methods.* 2015;12:763–765.
43. Qian Y, Cosio DMO, Piatkevich KD, et al. Improved genetically encoded near-infrared fluorescent calcium ion indicators for in vivo imaging. *PLoS Biol.* 2020;18:e3000965.
44. Namikawa K, Dorigo A, Zagrebelsky M, et al. Modeling neurodegenerative spinocerebellar ataxia type 13 in zebrafish using a Purkinje neuron specific tunable coexpression system. *J Neurosci.* 2019;39:3948–3969.
45. Wang L, Tsien RY. Evolving proteins in mammalian cells using somatic hypermutation. *Nat Protoc.* 2006;1:1346–1350.
46. Lee J, Liu Z, Suzuki PH, et al. Versatile phenotype-activated cell sorting. *Sci Adv.* 2020;6:eabb7438.
47. Fiedler BL, Van Buskirk S, Carter KP, et al. Droplet microfluidic flow cytometer for sorting on transient cellular responses of genetically-encoded sensors droplet microfluidic flow cytometer for sorting on transient cellular responses of genetically-encoded sensors. *Anal Chem.* 2016;89:711–719.
48. Wheeler EC, Vu AQ, Einstein JM, et al. Pooled CRISPR screens with imaging on microrraft arrays reveals stress granule-regulatory factors. *Nat Methods.* 2020;17(17):636–642.
49. Davis LM, Lubbeck JL, Dean KM, Palmer AE, Jimenez R. Microfluidic cell sorter for use in developing red fluorescent proteins with improved photostability. *Lab Chip.* 2013;13:2320–2327.
50. Chen KC, Wu CH, Chang CY, et al. Directed evolution of a lysosomal enzyme with enhanced activity at neutral pH by mammalian cell-surface display. *Chem Biol.* 2008;15:1277–1286.
51. Chen CP, Hsieh YT, Prijovich ZM, et al. ECSTASY, an adjustable membrane-tethered/soluble protein expression system for the directed evolution of mammalian proteins. *Protein Eng Des Sel.* 2012;25:367–375.
52. Black J, Vos JM. Establishment of an oriP/EBNA1-based episomal vector transcribing human genomic  $\beta$ -globin in cultured murine fibroblasts. *Gene Ther.* 2002;9:1447–1454.
53. Chien M-P, Werley CA, Farhi SL, Cohen AE. Photostick: A method for selective isolation of target cells from culture. *Chem Sci.* 2015;6:1701–1705.
54. Yamaoka N, Kawasaki Y, Xu Y, et al. Establishment of in vivo fluorescence imaging in mouse models of malignant mesothelioma. *Int J Oncol.* 2010;37:273–279.
55. Caysa H, Hoffmann S, Luetzkendorf J, et al. Monitoring of xenograft tumor growth and response to chemotherapy by non-invasive in vivo multispectral fluorescence imaging. *PLoS One.* 2012;7:e47927.
56. Sinclair J, Hamza I. Lessons from bloodless worms: Heme homeostasis in *C. elegans*. *Biomaterials.* 2015;28:481–489.
57. Liu Y, Kaval KG, van Hoof A, Garsin DA. Heme peroxidase HPX-2 protects *Caenorhabditis elegans* from pathogens. *PLoS Genet.* 2019;15:1–25.
58. Macheroux P. UV-visible spectroscopy as a tool to study flavoproteins. *Methods Mol Biol.* 1999;131:1–7.
59. Makarov NS, Campo J, Hales JM, Perry JW. Rapid, broadband two-photon-excited fluorescence spectroscopy and its application to red-emitting secondary reference compounds. *Opt Mater Express.* 2011;1:551.
60. Kimura Y, Satou C, Higashijima S-I. V2a and V2b neurons are generated by the final divisions of pair-producing progenitors in the zebrafish spinal cord. *Development.* 2008;135:3001–3005.
61. Brenner S. The genetics of *Caenorhabditis elegans*. *Genetics.* 1974;77:71–94.
62. Evans TC. (2006). Transformation and microinjection. *WormBook*, ed. The *C. elegans* Research Community, *WormBook*, doi/10.1895/wormbook.1.108.1, <http://www.wormbook.org>
63. Cranfill PJ, Sell BR, Baird MA, et al. Quantitative assessment of fluorescent proteins. *Nat Methods.* 2016;13:557–562.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Babakhanova S, Jung EE, Namikawa K, Zhang H, Wang Y, Subach OM, et al. Rapid directed molecular evolution of fluorescent proteins in mammalian cells. *Protein Science.* 2022;31:728–51. <https://doi.org/10.1002/pro.4261>



# Rapid Directed Molecular Evolution of Fluorescent Proteins in Mammalian Cells

Siranush Babakhanova<sup>1,2,3,4+</sup>, Erica E. Jung<sup>5+</sup>, Kazuhiko Namikawa<sup>6+</sup>, Hanbin Zhang<sup>7,8,9+</sup>, Yangdong Wang<sup>7,8,9</sup>, Oksana M. Subach<sup>10</sup>, Dmitry A. Korzhenevskiy<sup>10</sup>, Tatiana V. Rakitina<sup>10,11</sup>, Xian Xiao<sup>7,8,9</sup>, Wenjing Wang<sup>7,8,9</sup>, Jing Shi<sup>7,8,9</sup>, Mikhail Drobizhev<sup>12</sup>, Demian Park<sup>1,4</sup>, Lea Eisenhard<sup>6</sup>, Hongyun Tang<sup>7,8,9</sup>, Reinhard W. Köster<sup>6</sup>, Fedor V. Subach<sup>10</sup>, Edward S. Boyden<sup>1,4,13,14,15,16\*</sup>, Kiryl D. Piatkevich<sup>7,8,9\*</sup>

<sup>1</sup>Media Arts and Sciences, Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts (MA), USA.

<sup>2</sup>Department of Physics, MIT, Cambridge, MA, USA.

<sup>3</sup>Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, USA.

<sup>4</sup>MIT McGovern Institute for Brain Research, MIT, Cambridge, MA, USA.

<sup>5</sup>Department of Mechanical and Industrial Engineering, The University of Illinois at Chicago, Chicago, IL, USA

<sup>6</sup>Division of Cellular and Molecular Neurobiology, Zoological Institute, Technische Universität Braunschweig, Germany.

<sup>7</sup>School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China

<sup>8</sup>Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, China

<sup>9</sup>Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, Zhejiang, China

<sup>10</sup>National Research Center “Kurchatov Institute”, Moscow, Russian Federation

<sup>11</sup>Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, RAS, Moscow, Russian Federation

<sup>12</sup>Department of Microbiology and Immunology, Montana State University, Bozeman, MT, USA

<sup>13</sup>Department of Biological Engineering, MIT, Cambridge, MA, USA.

<sup>14</sup>Department of Brain and Cognitive Sciences, MIT, Cambridge, MA, USA.

<sup>15</sup>Howard Hughes Medical Institute, Cambridge, Massachusetts, USA

<sup>16</sup>Koch Institute, MIT, Cambridge, Massachusetts, USA

<sup>+</sup>These authors contributed equally to this work.

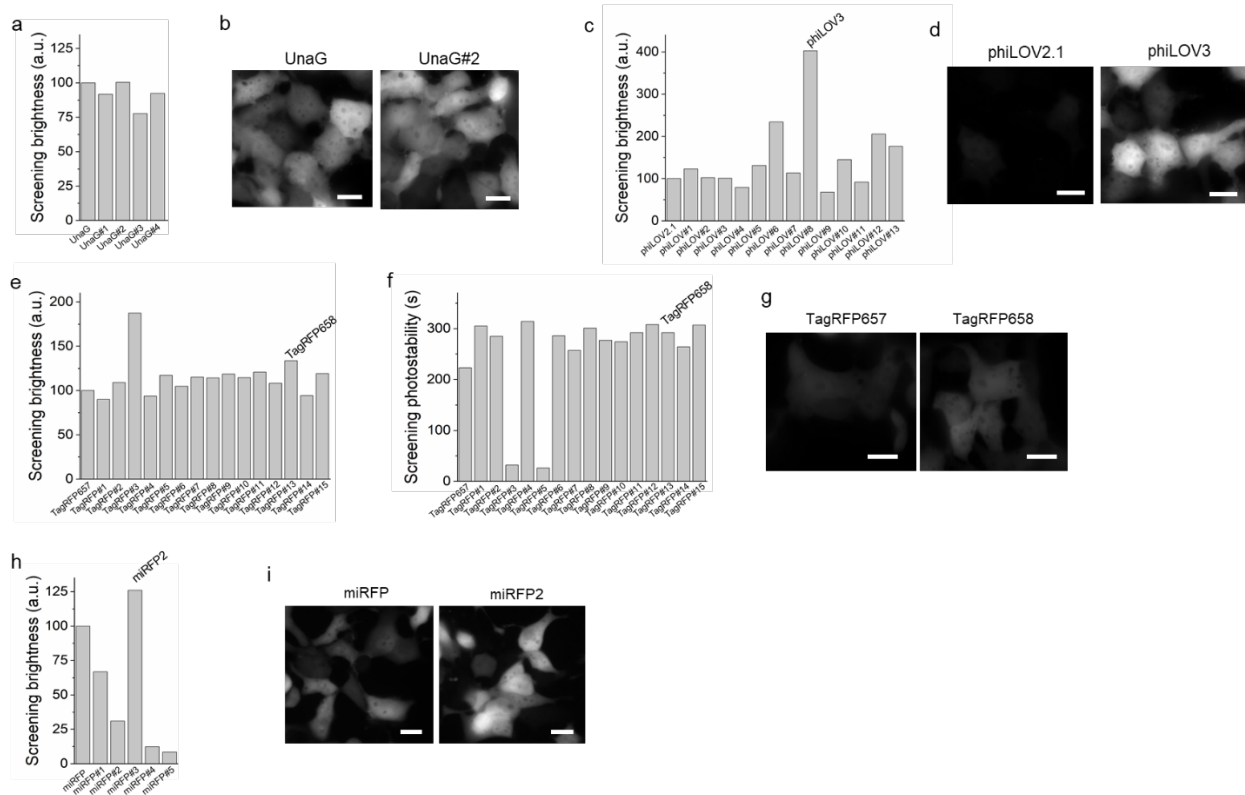
Correspondence to: [edboyden@mit.edu](mailto:edboyden@mit.edu) and [kiryl.piatkevich@westlake.edu.cn](mailto:kiryl.piatkevich@westlake.edu.cn)

**Supplementary Table 1.** Screening conditions for the selected FPs.

<b>Protein template</b>	<b>Library size (independent clones)</b>	<b>Fraction of clones with target genes (%)</b>	<b>FACS screening conditions</b>	<b>Fluorescence imaging conditions</b>
phiLOV2.1	$7.8 \cdot 10^5$	92	Ex: 488 nm; Em: 515/20BP	10x 0.45NA; Ex: 475/34BP; Em: 527/50BP;
	$2.3 \cdot 10^6$	100		
UnaG	$2.2 \cdot 10^6$	83	Ex: 488 nm; Em: 515/20BP	10x 0.45NA; Ex: 475/34BP; Em: 527/50BP;
	$4.0 \cdot 10^6$	100		
TagRFP657	$1.1 \cdot 10^6$	75	Ex: 640 nm; Em: 670/30BP	10x 0.45NA; Ex: 628/31BP; Em: 664LP
	$1.01 \cdot 10^7$	100		
miRFP	$2.01 \cdot 10^7$	95	Ex: 640 nm; Em: 710/50BP	10x 0.45NA; Ex: 628/31BP; Em: 664LP
	$1.8 \cdot 10^6$	60		

Ex – excitation wavelength; Em – emission wavelength; BP – bandpass; LP – longpass.

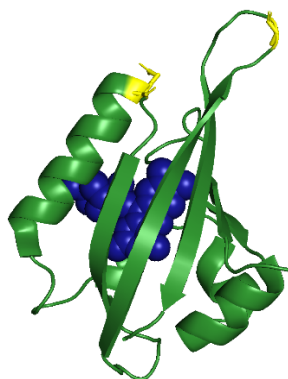
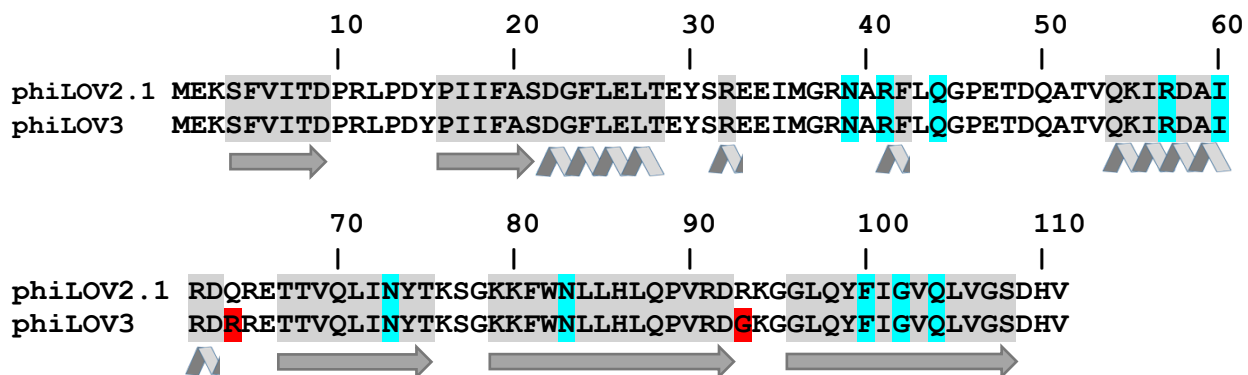
**Supplementary Figure 1.** Screening of the mutants selected in the final round of directed molecular evolution of UnaG, phiLOV2.1, TagRFP657, and miRFP.



To assess fluorescence and photostability, the selected mutants were expressed in HEK cells under CMV promoter and imaged under wide-field microscope. **(a)** Screening fluorescence brightness of the top four mutants in comparison to their parental protein UnaG. **(b)** Fluorescent images of live HEK cells expressing UnaG and UnaG#2 (dynamic range of the images are adjusted to the same values). **(c)** Screening fluorescence brightness of the selected phiLOV2.1 mutants. **(d)** Fluorescent images of live HEK cells expressing phiLOV2 and phiLOV3 (dynamic range of the images are adjusted to the same values). **(e)** Screening fluorescence brightness of the selected TagRFP657 mutants. **(f)** Screening photostability of the selected TagRFP657 mutants. **(g)** Fluorescent images of live HEK cells expressing TagRFP657 and TagRFP658 (dynamic range of the images are adjusted to the same values). **(h)** Screening fluorescence brightness of the selected miRFP mutants. **(i)** Fluorescent images of live HEK cells expressing miRFP and miRFP2 (dynamic range of the images are adjusted to the same values). Scale bars, 10 μm.

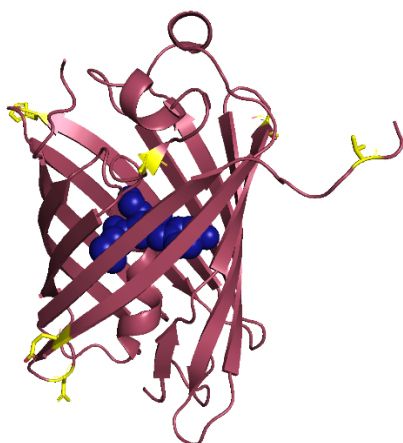
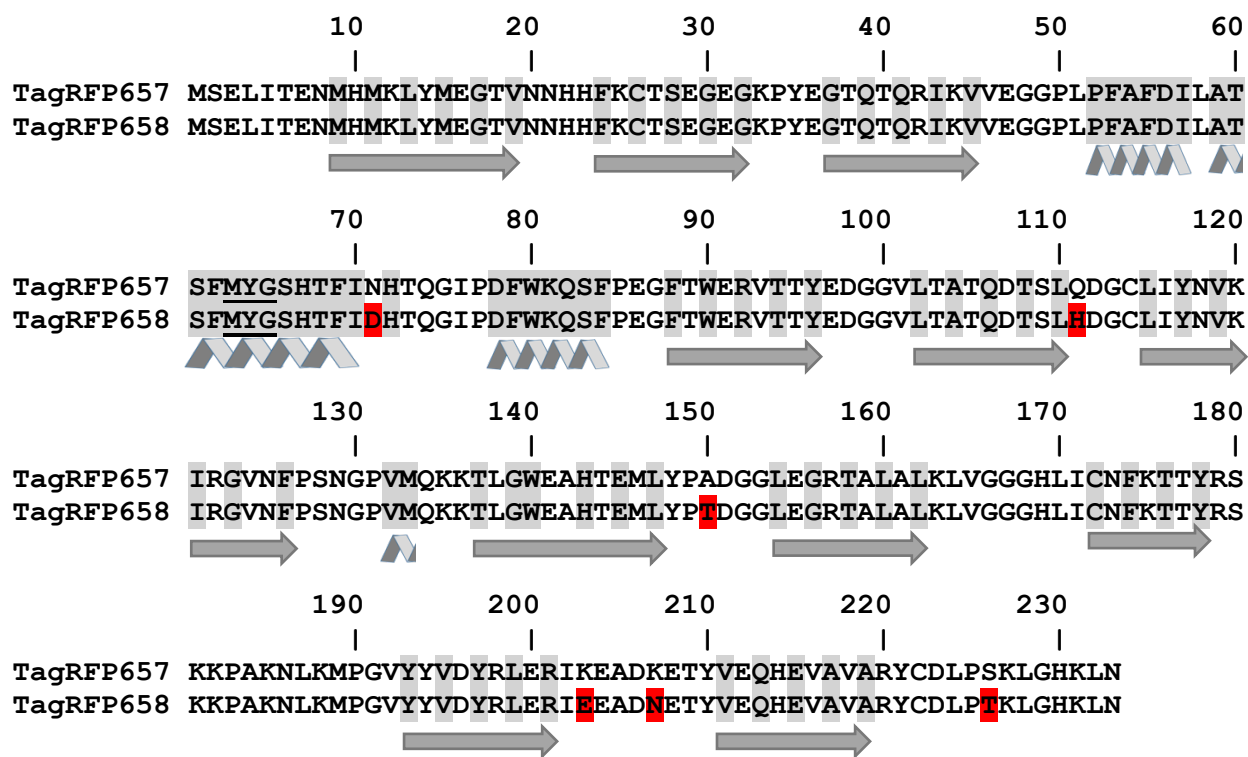


**Supplementary Figure 3.** Alignment of amino acid sequences of phiLOV2.1 and phiLOV3.



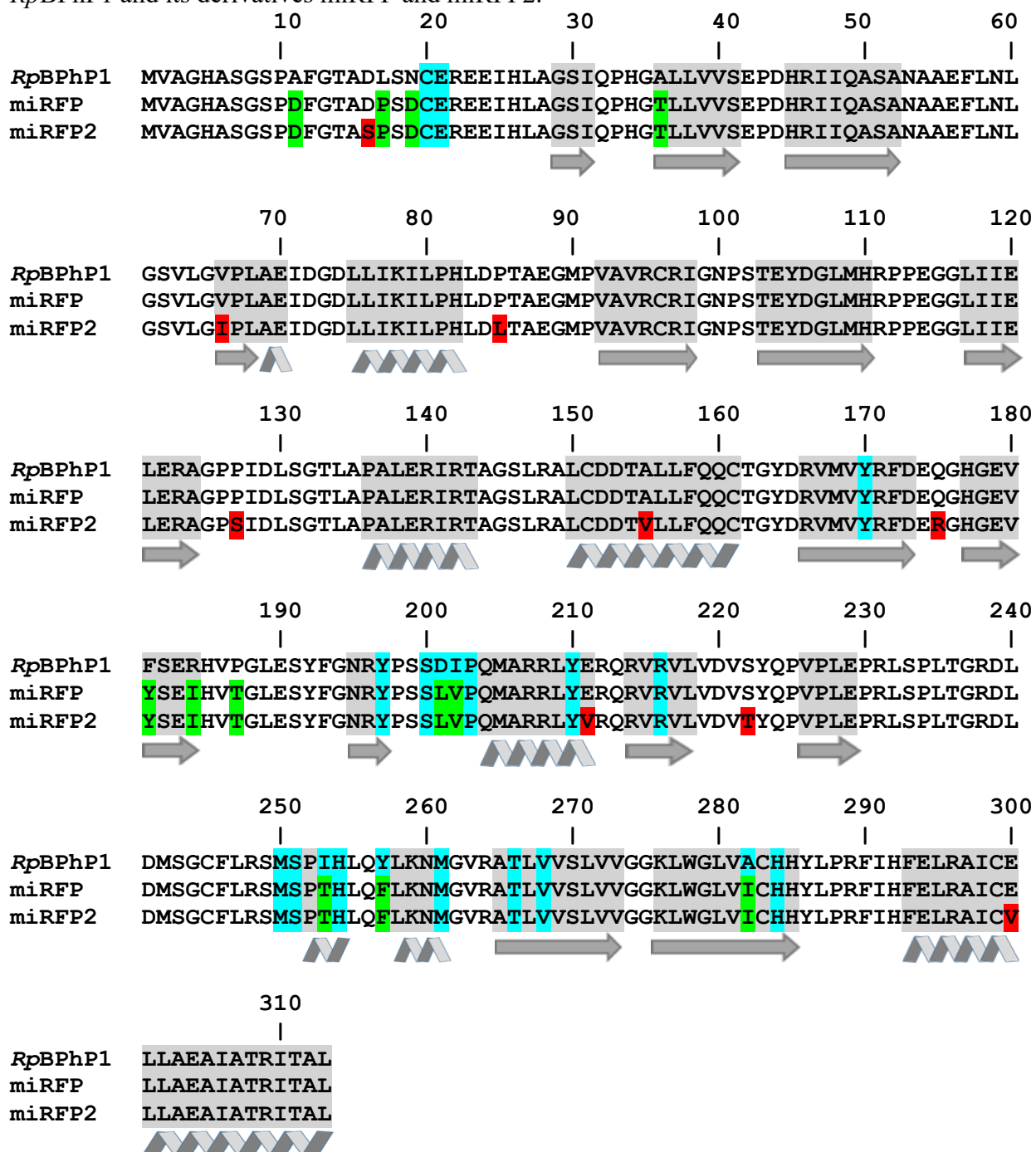
(Top) The residues surrounding the chromophore within 3.5 Å are highlighted in cyan (PDB: 4EEU). Mutations resulting in the conversion of parental phiLOV2.1 into the phiLOV3 variant are highlighted in red. The  $\beta$ -sheet-forming regions and  $\alpha$ -helices are shaded and denoted with arrows and ribbons, respectively. (Bottom) 3D protein structure of phiLOV2.1 (PDB 4EEU), chromophore shown in blue, amino acids mutated in phiLOV3 are shown in yellow.

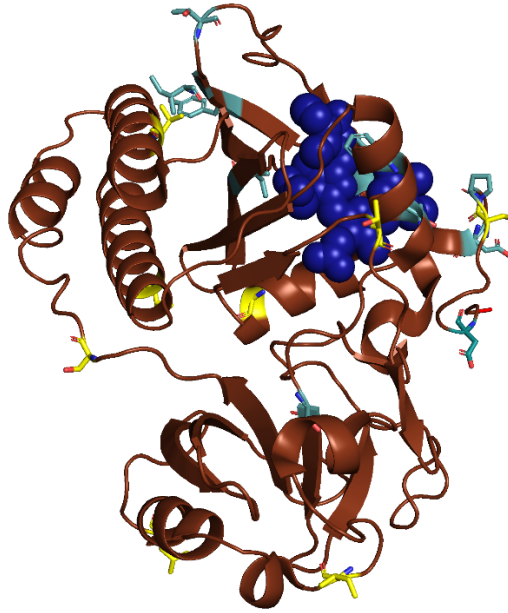
**Supplementary Figure 4.** Alignment of amino acid sequences of TagRFP657 and TagRFP658.



(*Top*) Internal amino acids are shaded, and the chromophore-forming residues are underlined. Mutations resulting in the conversion of TagRFP657 into TagRFP658 are highlighted in red. The  $\beta$ -sheet-forming regions and  $\alpha$ -helices are shaded and denoted with arrows and ribbons, respectively. (*Bottom*) 3D protein structure of TagRFP (PDB 3M22, precursor of TagRFP657), chromophore shown in blue, amino acids mutated in TagRFP658 are shown in yellow.

**Supplementary Figure 5.** Alignment of amino acid sequences of PAS-GAF domains of the *RpBPhP1* and its derivatives miRFP and miRFP2.

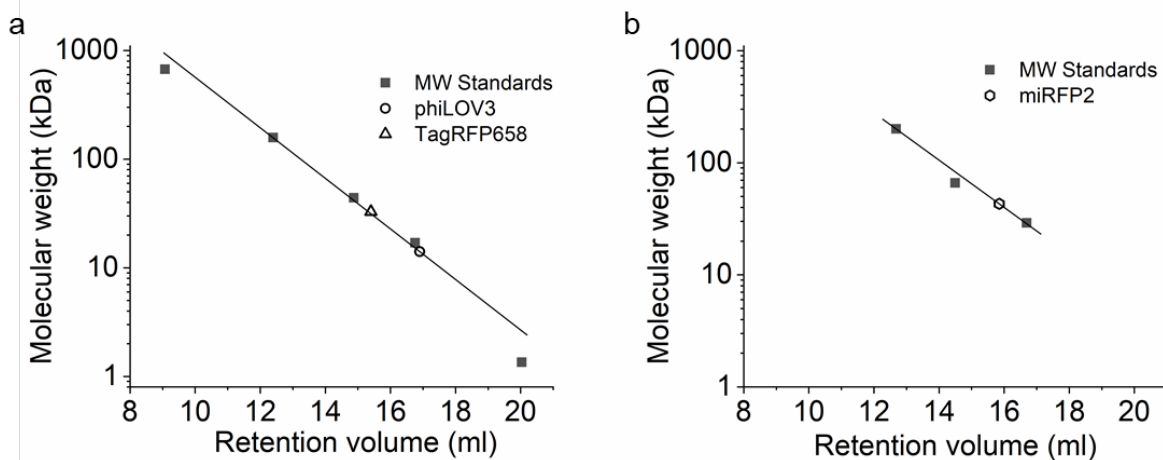




(*Top*) The residues surrounding the chromophore within 4.0 Å in *RpBphP1* are highlighted in cyan (PDB: 4GW9). Mutations resulting in the conversion of parental *RpBphP1* into miRFP are highlighted in green. Mutations resulting in the conversion of miRFP into miRFP2 variant are highlighted in red. The  $\beta$ -sheet-forming regions and  $\alpha$ -helices are shaded and denoted with arrows and ribbons, respectively. (*Bottom*) 3D protein structure of *RpBphP1* photosensory core domain (PDB 5OY5; precursor of miRFP), chromophore shown in blue, amino acids mutated in miRFP are shown in cyan, amino acids mutated in miRFP2 are shown in yellow.

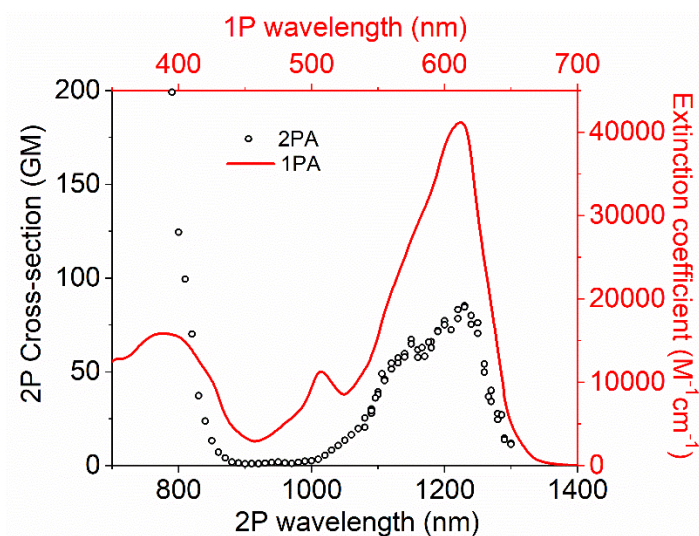


**Supplementary Figure 6.** Size exclusion chromatography calibration plots.



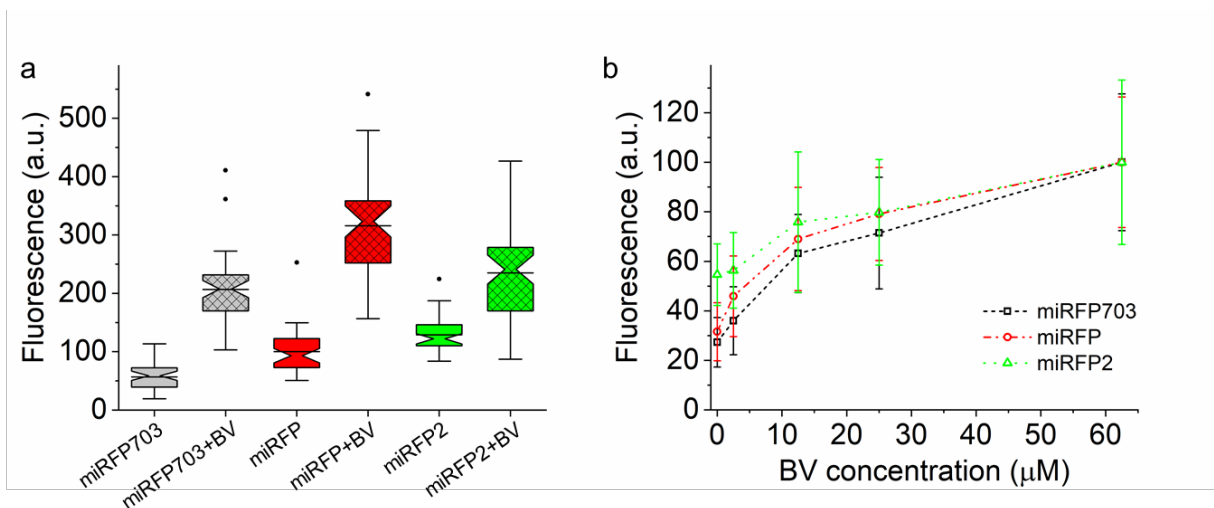
Size exclusion chromatography calibration plots showing the relative retention volumes of protein molecular weight standards and proteins of interest (see Figure 2 for size exclusion chromatography profiles). **(a)** Calibration plot showing the relative retention volumes of protein molecular weight standards (black squares; Gel Filtration Standard, Bio-Rad;  $n = 1$  technical replicate) and phiLOV3 (circle) and TagRFP658 (triangle). **(b)** Calibration plot showing the relative retention volumes of protein molecular weight standards (black squares;  $n = 1$  technical replicate) and miRFP2 (hexagon).

**Supplementary Figure 7.** Two-photon cross-section spectrum of TagRFP658.



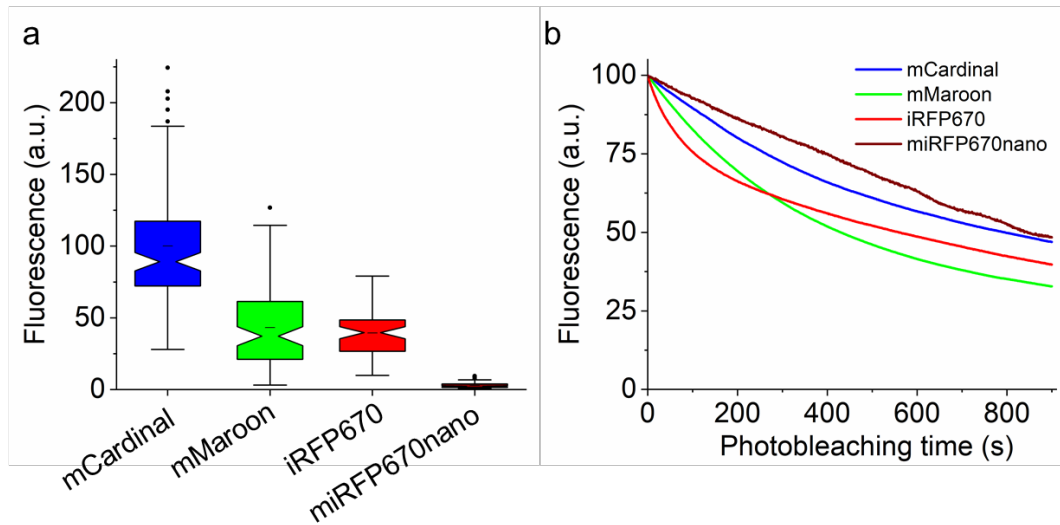
One-photon absorption (solid line) and two-photon cross-section (open circles) absorption spectra of TagRFP658. The two-photon cross-section is presented versus laser wavelength used for excitation. GM, Goepfert-Mayer units. Previously we showed that EGFP and bacterial phytochrome photoreceptor (BphP)-derived FPs can be efficiently co-excited under two-photon excitation at 880 nm due to an overlap of their corresponding first  $S_0 \rightarrow S_1$  and higher  $S_0 \rightarrow S_n$  (ref.<sup>1</sup>) electronic transitions. These results suggest that two-photon spectral overlap of TagRFP658 with EGFP is less substantial than that for iRFPs and might not be practical for single-wavelength co-excitation with EGFP.

**Supplementary Figure 8.** Effect of exogenous biliverdin on NIR fluorescence of miRFP703, miRFP, and miRFP2.



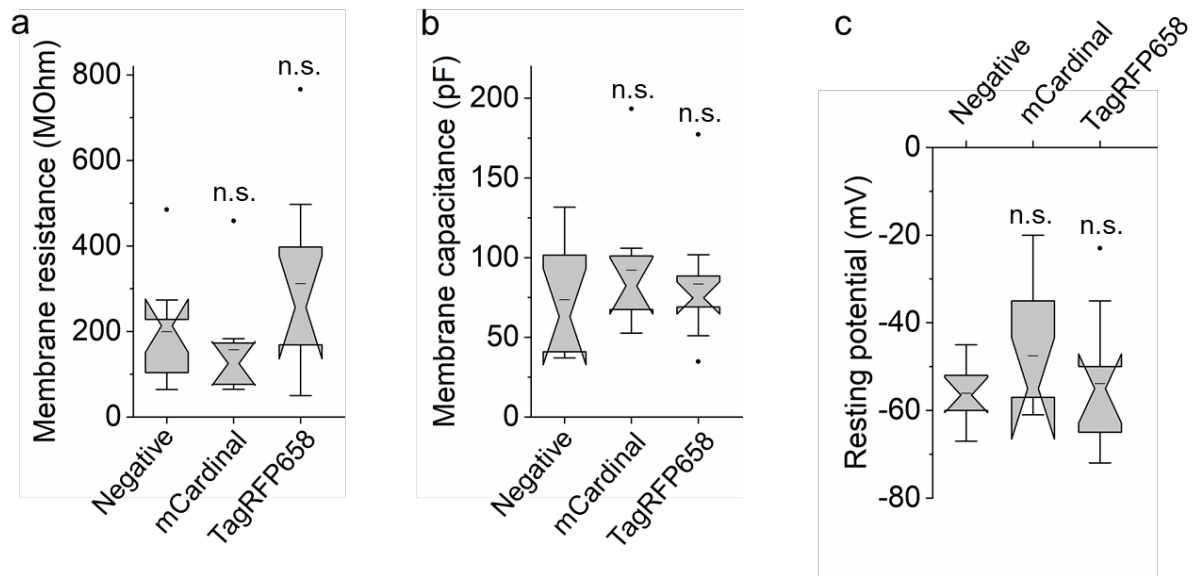
**(a)** Relative normalized fluorescence of HEK cells expressing miRFP703 (gray), miRFP (red), and miRFP2 (green) under CMV promoter with (cross-hatched boxes) and without (open boxes) addition of 62.5 μM for 3 h before imaging ( $n = 44, 45, 40, 40, 41,$  and  $48$  cells for miRFP703, miRFP703+BV, miRFP, miRFP+BV, miRFP2, and miRFP2+BV, respectively, from one independent transfection each). Imaging conditions: excitation 631/28 nm from an LED, emission 664LP. Box plots with notches are used (see caption for **Figure 1c** for description). **(b)** Exogenous BV binding to miRFP703 (black dashed line), miRFP (red dashed line), and miRFP2 (green dashed line) proteins expressed in HEK cells. Cells were incubated with the respective concentration of BV for 3 h before imaging ( $n = \sim 40$  cells for each protein under each condition from one independent transfection). Open symbols, mean; error bars, standard deviation. Imaging conditions same as in **(a)**.

**Supplementary Figure 9.** Quantitative assessment of the selected near-infrared FPs in HEK cells.



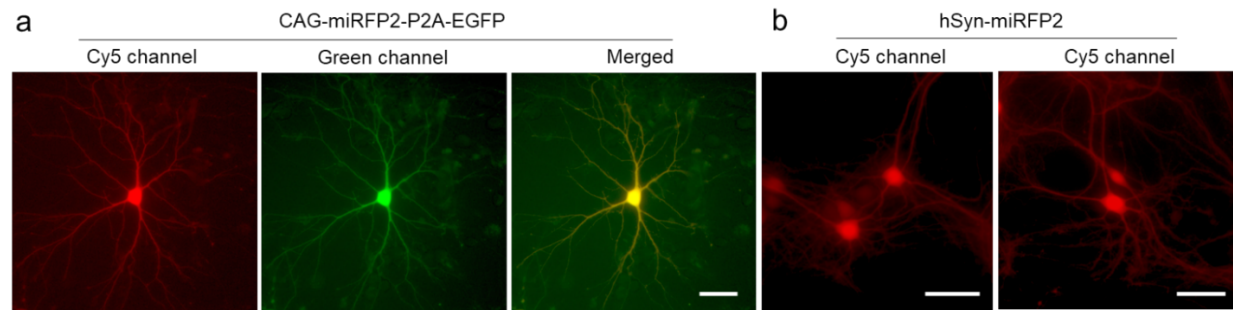
The selected FPs were expressed in HEK cells under CMV promoter and imaged 48 h post transfection. **(a)** Relative normalized fluorescence and **(b)** raw photobleaching curves for mCardinal (blue), mMaroon (green), iRFP670 (red), and miRFP670nano (deep red) expressed in HEK cells (for fluorescence  $n=127$ ,  $93$ ,  $65$ , and  $119$  cells, respectively, from one culture; for photobleaching  $n=19$ ,  $12$ ,  $38$ , and  $10$  cells, respectively, from one culture; imaging conditions: wide-field excitation  $631/28$  nm from LED at  $25$  mW/mm<sup>2</sup>, emission 664LP).

**Supplementary Figure 10.** Membrane properties of cultured primary mouse hippocampal neurons expressing mCardinal and TagRFP658.



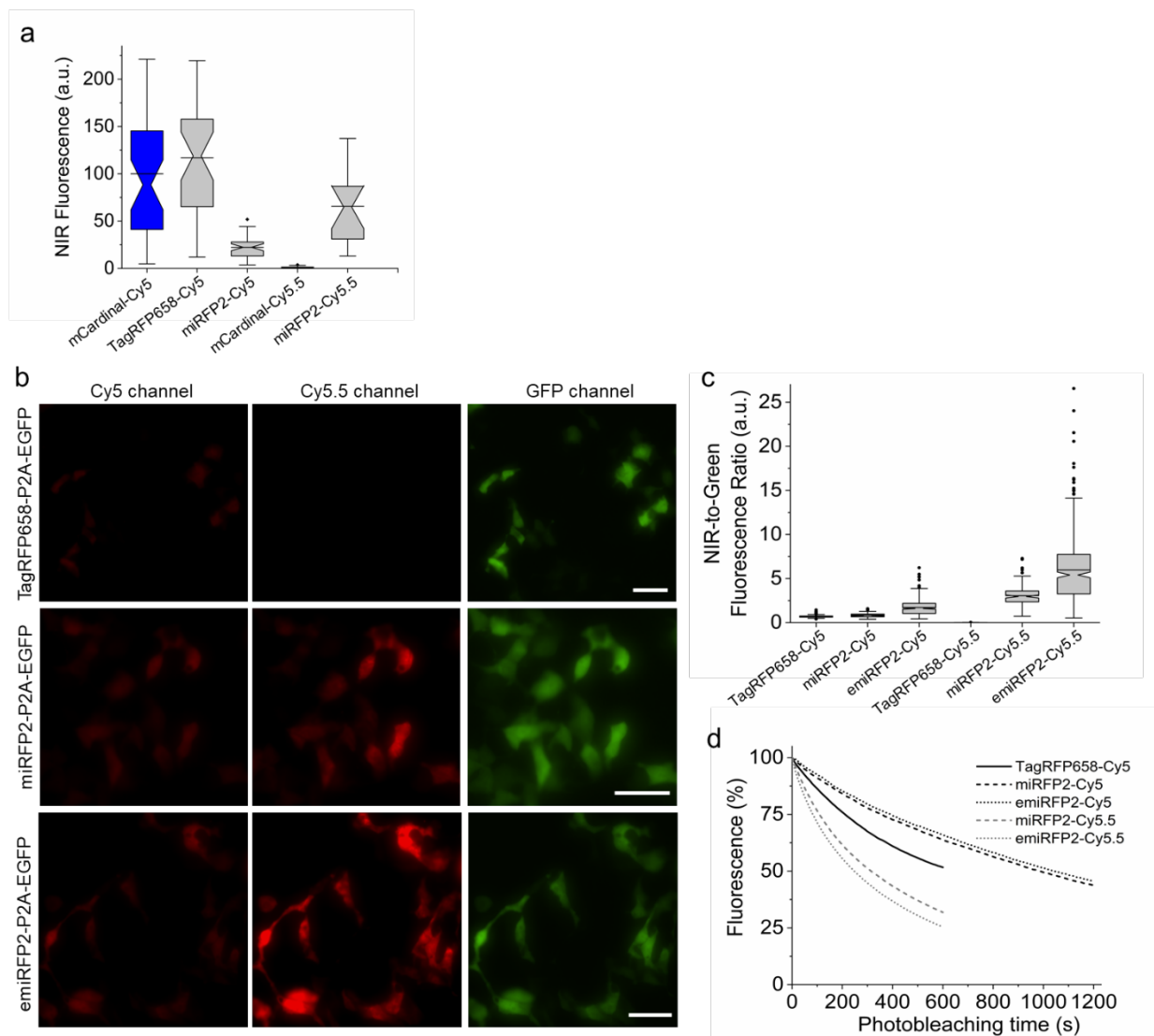
Cultured hippocampal neurons expressing mCardinal ( $n=9$  cells from one culture) and TagRFP658 ( $n=9$  cells from one culture) under CaMKII promoter were patched to compare membrane properties to non-transfected neurons (negative control,  $n=10$  cells from two cultures). **(a)** Membrane resistance.  $P > 0.05$ , not significant (n.s.) Kruskal–Wallis one-way analysis of variance throughout all panels of this figure. **(b)** Membrane capacitance. **(c)** Resting potential. Throughout this figure, box plots with notches are used; narrow part of notch, median; top and bottom of the notch, 95% confidence interval for the median; top and bottom horizontal lines, 25% and 75% percentiles for the data; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles; horizontal line, mean. For datasets with  $n < 10$ , open circles represent individual data points; data points which are less than the 25th percentile or greater than the 75th percentile by more than 1.5 times the interquartile range are also represented as open circles.

**Supplementary Figure 11.** Expression of miRFP2 in primary mouse hippocampal neurons.



**(a)** Representative image of calcium-phosphate transfected mouse hippocampal neuron expressing miRFP2-P2A-EGFP under CAG promoter (n=40 neurons from one culture). Imaging conditions: NIR channel, excitation 631/28 nm from an LED, emission 664LP; green channel: excitation 475/34BP from an LED, emission 527/50 nm. **(b)** Representative images of rAAV-transduced mouse hippocampal neurons expressing miRFP2 under hSyn promoter (n=120 neurons from one culture). Imaging conditions: NIR channel, excitation 631/28 nm from an LED, emission 664LP.

**Supplementary Figure 12.** Intracellular brightness and photostability of mCardinal, TagRFP658, miRFP2, and emiRFP2 in live neurons and HEK cells in Cy5 and Cy5.5 channels.

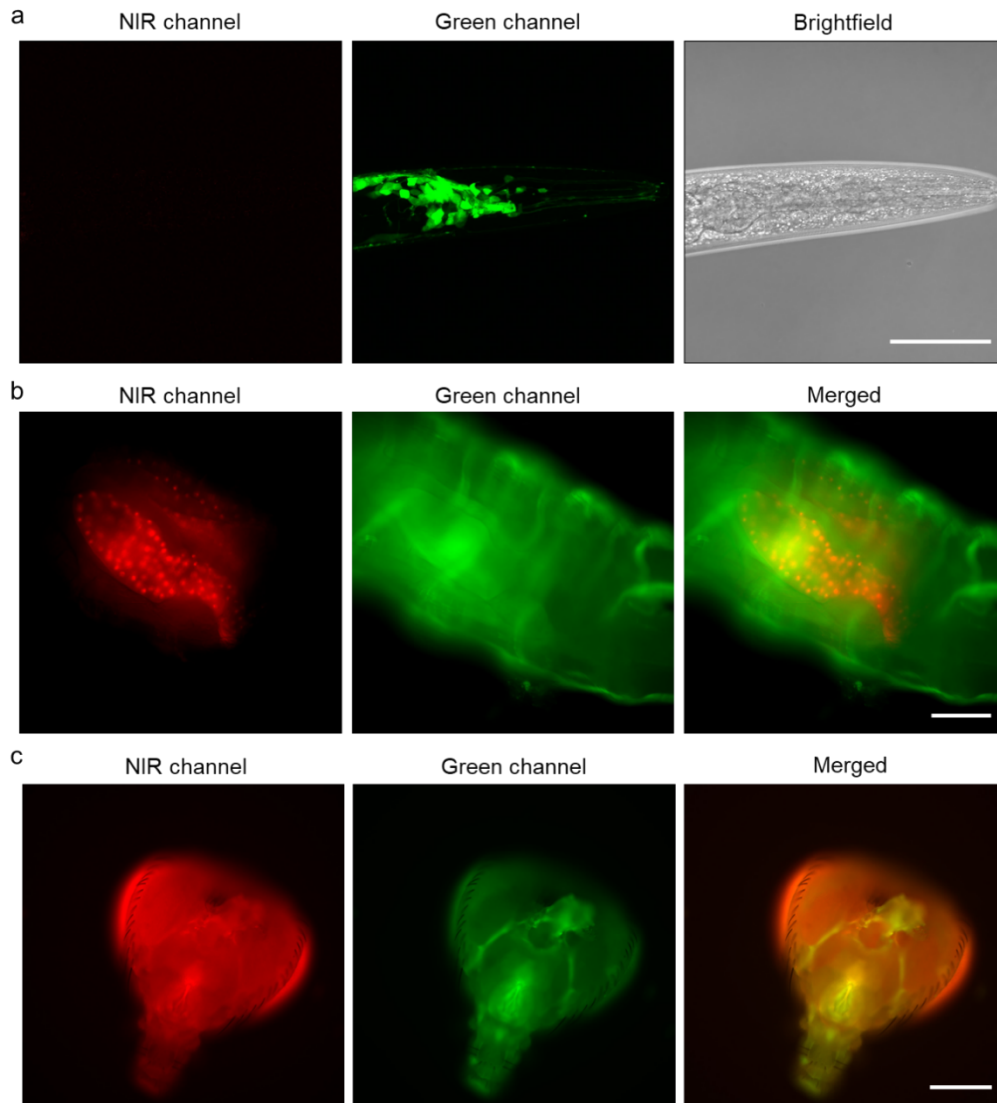


Intracellular brightness and photostability of TagRFP658, miRFP2, and emiRFP2 in live neurons and HEK cells. **(a)** Normalized NIR fluorescence intensity of mCardinal, TagRFP658, miRFP2 expressed in live primary cultured neurons in Cy5 and Cy5.5 channels ( $n = 39, 33,$  and  $41$  neurons from  $3, 2,$  and  $3$  independent transfections from one culture each for mCardinal, TagRFP658, and emiRFP2, respectively, for Cy5 channel and  $n = 15$  and neurons from one independent transfection from one culture each for mCardinal and emiRFP2, respectively, for Cy5.5 channel). Imaging conditions: Cy5 channel: excitation  $635/22$  nm from  $637$  nm laser, emission  $730/140$  nm; Cy5.5 channel: excitation  $680/13$  nm from  $680$  nm laser, emission  $710$  LP. Images in Cy5 and Cy5.5

were taken under matching excitation intensity (66 mW/mm<sup>2</sup>) and the same exposure time (100 ms). Box plots with notches are used in this figure (see **Fig. 1c** for the description). **(b)** Representative fluorescence images of cells transfected with pAAV-CAG-TagRFP658-P2A-EGFP (top), pAAV-miRFP2-P2A-EGFP (middle), and pAAV-emiRFP2-P2A-EGFP (bottom; n = 226, 131, and 475 cells from 2, 4, and 2 independent transfections, respectively). Imaging conditions: Cy5 and Cy5.5 channels same as in panel **a**; GFP channel: excitation 478/24 nm for an LED; emission 535/46 nm. The dynamic range of fluorescence intensity in Cy5 and Cy5.5 channels are identical across all images. Scale bar, 50  $\mu$ m. **(c)** NIR-to-green fluorescence ratio for TagRFP658, miRFP2, and emiRFP2 in live HEK cells shown in **a** (n = 226, 131, and 475 cells from 2, 4, and 2 independent transfections, respectively). **(d)** Intracellular photostability of TagRFP658, miRFP2, and emiRFP2 in Cy5 and Cy5.5 channels (n = 56, 57, and 37 cells for TagRFP658, miRFP2, and emiRFP2; 131, and 475 cells from 2 independent transfections each under Cy5 excitation; n = 44 and 78 from 2 and 3 independent transfection for miRFP2 and emiRFP2, respectively). Imaging conditions the same as in **a**.

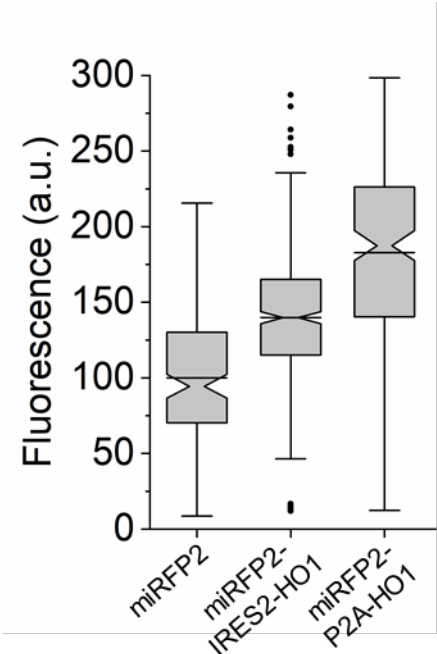


**Supplementary Figure 13.** Imaging of miRFP2 in transgenic *C.elegans* worms and *Drosophila melanogaster*.



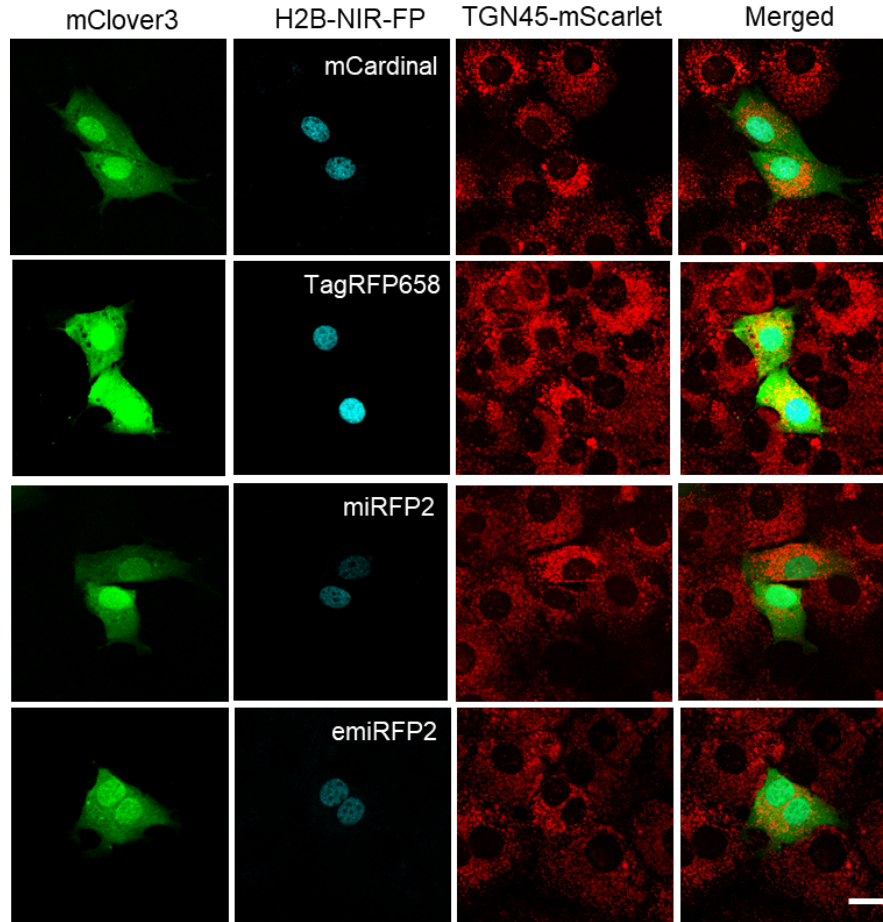
Live transgenic *C.elegans* worms and *Drosophila* larvae with pan-neuronal expression of corresponding codon-optimized genes of miRFP2 were imaged in NIR and green channels. **(a)** Representative fluorescence and brightfield images of the *C.elegans* head co-expressing codon-optimized genes of miRFP2 and jGCaMP7b in neurons (n=15 worms from two independent microinjections). Imaging conditions: NIR channel, excitation 631 nm from a laser, emission 645-700 nm; green channel, excitation 488 nm from a laser, emission 500-650 nm. Scale bar, 50  $\mu$ m. **(b)** Representative fluorescence images of live intact 3<sup>rd</sup> instar *Drosophila* larva expressing miRFP2 (n= 10 larvae from two transgenic lines). Imaging conditions: NIR channel, excitation 631/28 nm from LED, emission 665LP; green channel, excitation 475/34 nm from LED, emission 527/50 nm (green fluorescence correspond to autofluorescence). Scale bar, 250  $\mu$ m. **(c)** Representative fluorescence images of adult *Drosophila* fly head expressing miRFP2 (n= 11 flies from one transgenic line). Imaging conditions the same as in **b**. Scale bar, 250  $\mu$ m.

**Supplementary Figure 14.** Influence of HO1 co-expression on fluorescence of miRFP2 in live HEK cells.



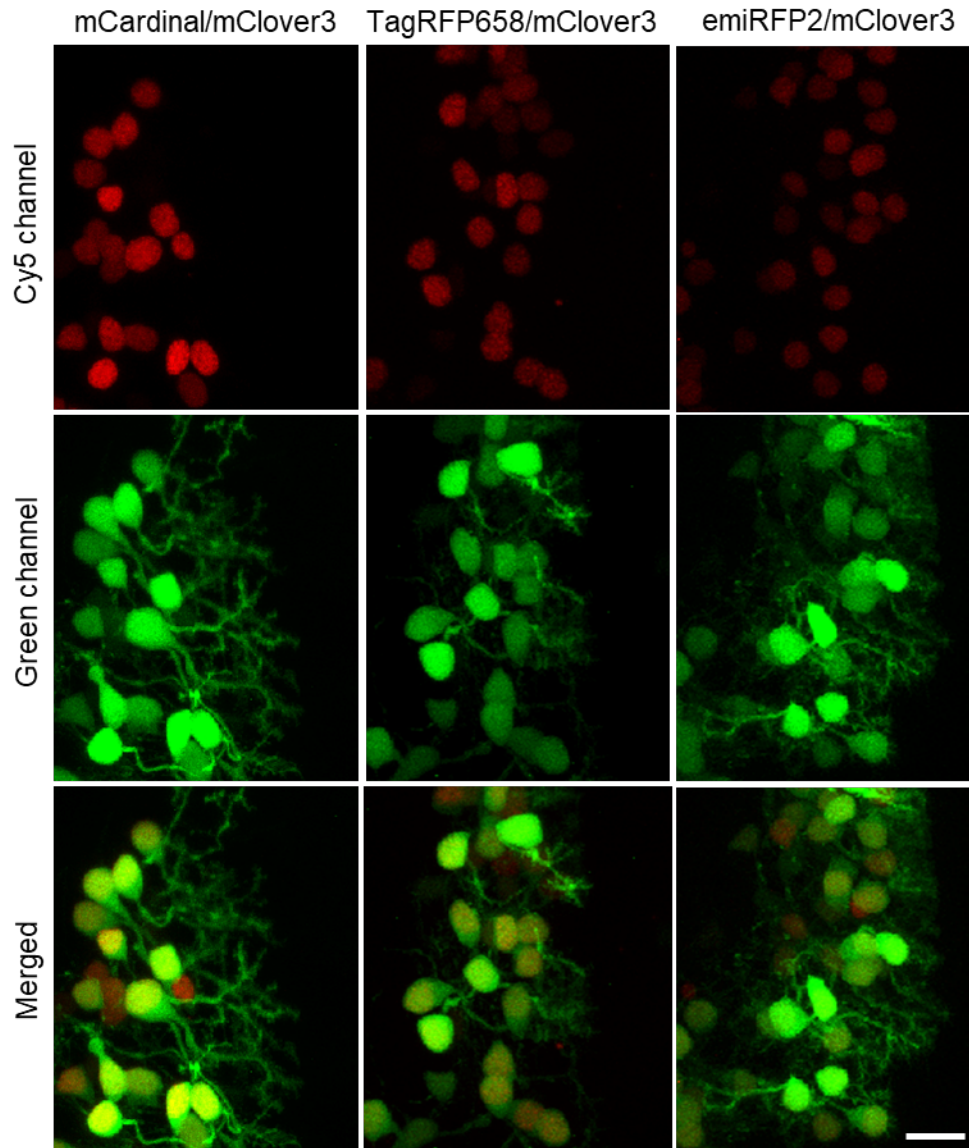
Relative fluorescence of live HEK cells expressing miRFP2, miRFP2-IRES2-HO1, and miRFP2-P2A-HO1 under CMV promoter (n= 142, 394, and 187 cells, respectively, from two independent transfections each).

**Supplementary Figure 15.** Triple color imaging of fixed NIH3T3 cells.



Each bidirectional expression construct carrying either of H2B fused with NIR-FP together with mClover3 was transfected into NIH3T3 cells stably expressing mScarlet N-terminally fused to TGN46 located in TGN and the vesicles cycling between TGN and the cell surface. Transfected cells were fixed by 4% PFA, followed by imaging using confocal microscope in green, red, and Cy5 channels. Imaging conditions: green channel, excitation 488 nm from an argon laser, emission 495-530 nm; Red channel, excitation 561 nm from a laser, emission 570-600 nm. Cy5 channel, excitation 633 nm from a laser, emission 660-785 nm. Scale bar, 20  $\mu$ m.

**Supplementary Figure 16.** Dual-color *in vivo* imaging of wild-type zebrafish larvae co-expressing NIR FPs with mClover3 in PCs.



Representative confocal images of 4 dpf wild type zebrafish larvae transiently co-expressing H2B-NIR FPs with mClover3. Each subset of images (top to bottom) shows the co-expression of each H2B-NIR-FP fusion (left, mCardinal; middle, TagRFP658; right, emiRFP2) with cytoplasmic mClover3 (n=4 fish for each subset). Imaging conditions, Cy5 channel: excitation 633 nm laser, emission 722/63 nm; green channel: excitation 488 nm laser, emission 513/17 nm. Scale bar, 10  $\mu$ m.

## References

1. Piatkevich, K. D. K. D. *et al.* Near-Infrared Fluorescent Proteins Engineered from Bacterial Phytochromes in Neuroimaging. *Biophys. J.* **113**, 2299–2309 (2017).