



Published in final edited form as:

Dev Biol. 2024 October ; 514: 109–116. doi:10.1016/j.ydbio.2024.06.011.

Tissue-specific and endogenous protein labeling with split fluorescent proteins

Gloria D. Liguas^{1,2,*}, German Paniagua^{1,*}, Jesselynn LaBelle^{1,2}, Adela Ramos-Martinez¹, Kyle Shen¹, Emma H. Gerlt¹, Kaddy Aguilar¹, Ngoc Nguyen¹, Stefan C. Materna^{1,2,3}, Stephanie Woo^{1,2,3,†}

¹Department of Molecular Cell Biology, University of California, Merced, CA USA

²Quantitative and Systems Biology Graduate Group, University of California, Merced, CA USA

³Health Sciences Research Institute, University of California, Merced, CA USA

Abstract

The ability to label proteins by fusion with genetically encoded fluorescent proteins is a powerful tool for understanding dynamic biological processes. However, current approaches for expressing fluorescent protein fusions possess drawbacks, especially at the whole organism level. Expression by transgenesis risks potential overexpression artifacts while fluorescent protein insertion at endogenous loci is technically difficult and, more importantly, does not allow for tissue-specific study of broadly expressed proteins. To overcome these limitations, we have adopted the split fluorescent protein system mNeonGreen_{21-10/11} (split-mNG2) to achieve tissue-specific and endogenous protein labeling in zebrafish. In our approach, mNG2₁₋₁₀ is expressed under a tissue-specific promoter using standard transgenesis while mNG2₁₁ is inserted into protein-coding genes of interest using CRISPR/Cas-directed gene editing. Each mNG2 fragment on its own is not fluorescent, but when co-expressed the fragments self-assemble into a fluorescent complex. Here, we report successful use of split-mNG2 to achieve differential labeling of the cytoskeleton genes *tubb4b* and *krt8* in various tissues. We also demonstrate that by anchoring the mNG2₁₋₁₀ component to specific cellular compartments, the split-mNG2 system can be used to manipulate protein localization. Our approach should be broadly useful for a wide range of applications.

Keywords

Zebrafish; CRISPR-Cas; protein tagging; split fluorescent protein

Introduction

Protein labeling by fusion with genetically encoded fluorescent proteins has been a powerful tool for studying biological processes, allowing scientists to visualize and track proteins of interest in live cells. Fluorescent protein labeling has been especially useful for investigating the dynamic processes that occur during embryonic development.

[†]Author for correspondence: swoo6@ucmerced.edu.

*These authors contributed equally

However, traditional methods for generating and expressing fluorescent fusion proteins, especially in multicellular organisms, have several drawbacks. In zebrafish and other model organisms, expression of fusion proteins can be achieved by injection of *in vitro* transcribed mRNA (Rosen et al., 2009), which is ubiquitous, or by transgenesis, which utilizes gene regulatory elements to drive spatiotemporal restricted expression (Clark et al., 2011). These approaches, however, run the risk of producing overexpression artifacts, in which proteins may not function or localize correctly when expressed at higher than wild-type levels (Simiczjew et al., 2014). An alternative approach is to knock in fluorescent protein coding sequences into the genetic locus of that protein of interest (Albadri et al., 2017; Auer and Del Bene, 2014; Kimura et al., 2014). Although this approach has the advantage of preserving endogenous regulation of that protein's expression, many proteins are expressed broadly; issues arise when there is a need to study a broadly expressed protein in a specific tissue. Thus, there is a need for tissue-specific and endogenous tagging of proteins.

Split fluorescent proteins (split-FPs) are self-complementing protein fragments that only fluoresce when bound together. Split-FPs have been successfully used to visualize and quantify cell-cell interactions (Feinberg et al., 2008), signaling pathway activation (Harvey and Smith, 2009), and subcellular protein localization (Cho et al., 2022). One commonly used split-FP system is based on the yellow-green fluorescent protein monomeric NeonGreen2 (mNG2) in which strands 1–10 of the mNG2 beta-barrel (mNG2_{1–10}) and strand 11 (mNG2₁₁) are expressed as independent protein fragments (Feng et al., 2017). On their own, the fragments are nonfluorescent, but when present in the same cell, they will self-assemble into a bimolecular complex with similar spectral properties to the intact, full-length fluorescent protein. The split-mNG2 system has been demonstrated to function in several different organisms and cell types (Cho et al., 2022; Kesavan et al., 2021; O'Hagan et al., 2021). Here, we adapt it for use in zebrafish to achieve tissue-specific and endogenous protein labeling. In our approach, mNG2_{1–10} is expressed under the control of a tissue-specific promoter using standard zebrafish transgenesis techniques. Because the mNG2₁₁ fragment is only 16 amino acids long, its short sequence can be easily inserted into endogenous genetic loci by CRISPR/Cas-directed gene editing. In this way, the mNG2₁₁-tagged protein will continue to be expressed at endogenous levels, but fluorescent signal will only be detected in tissues in which mNG2_{1–10} is co-expressed (Fig. 1A).

Results

mNG2_{1–10} and mNG2₁₁ can assemble fluorescent complexes in zebrafish embryos

To assess the viability of our protein labeling strategy, we first determined if split-FP fragments could self-assemble in zebrafish embryos to form functional fluorescent complexes (Fig. 1B–D). We tested two different FP_{1–10/11}-type systems, split-GFP (Kamiyama et al., 2016) and split-mNG2 (Feng et al., 2017). We injected mRNAs encoding GFP_{1–10} and GFP₁₁-H2B (GFP₁₁ fused to histone 2B) or mNG2_{1–10} and mNG2₁₁-H2B (mNG2₁₁ fused to histone 2B) into zebrafish embryos. For both systems, expression of the FP_{1–10} or FP₁₁ fragments alone did not produce fluorescence. However, when both fragments were co-expressed, we could detect nuclear-localized fluorescent signals by 6 hours post-fertilization (hpf) using confocal fluorescence microscopy (Fig. 1B–C). We

observed that embryos expressing split-mNG2 fragments (Fig. 1C) were brighter than those expressing the split-GFP fragments (Fig. 1B). Over time, split-mNG2 fluorescence remained brighter than split-GFP, which is consistent with a previous study showing that split-mNG2 can produce stronger fluorescence with less background compared to split-GFP (Feng et al., 2017). By 24 hpf split-mNG2 fluorescence was bright enough to be detected by a fluorescence stereomicroscope (Fig. 1D–E). Split-mNG2 fluorescence could still be detected after paraformaldehyde fixation even with some loss of brightness (Fig. S1). Based on these observations, we only used the split-mNG2 system for further experiments.

Generating mNG2_{1–10} transgenic lines

We next determined whether transgene-driven expression of mNG2_{1–10} could be used to spatially restrict fluorescence (Fig. 2). We generated multiple transgenic zebrafish lines that express mNG2_{1–10} under control of various promoters representing a broad range of tissue types including *fezf2* (brain and eye) (Berberoglu et al., 2009), *myl7* (myocardium) (Huang et al., 2003), and *ubb* (ubiquitous expression) (Mosimann et al., 2011). To verify that these transgenic lines were functional, we injected transgenic embryos with mNG2₁₁-H2B mRNA, which would be distributed ubiquitously, and qualitatively assessed fluorescence patterns at 24 or 48 hpf. We found that uninjected mNG2_{1–10} transgenic embryos exhibited no detectable fluorescence (Fig. S2). In contrast, transgenic embryos injected with mNG2₁₁-H2B mRNA exhibited fluorescence in spatially restricted patterns consistent with the promoter used to drive mNG2_{1–10} expression (Fig. 2A–F). Compared to embryos expressing full-length, intact GFP under control of the same tissue-specific promoters, we found that GFP and split-mNG2 fluorescence were present in the same tissues and regions (Fig. 2G–L). In some cases, we observed minor differences in brightness that may be due to slight differences in staging or insertion-specific differences in transgene expression, but the overall pattern of tissue restriction was comparable between split-mNG2 and intact GFP lines.

mNG2₁₁ tagging by CRISPR/Cas-directed gene editing

We next determined whether proteins of interest could be tagged with mNG2₁₁ at their endogenous genetic loci by CRISPR/Cas-guided homology directed repair (Fig. 3). Previous reports have suggested that split-FP tagging works best for highly expressed genes (Goudeau et al., 2021; O’Hagan et al., 2021). Therefore, we targeted three genes that are highly expressed with relatively broad patterns — *tubb4b*, which codes for Beta-tubulin 4b; *krt8*, which codes for Keratin 8; and *h2az2b*, which codes for histone H2A. We designed guide RNAs (gRNAs) targeting each gene just downstream of the start (*tubb4b*) or upstream of the stop (*krt8*, *h2aza2b*) codon to generate, respectively, N- or C-terminal mNG2₁₁ tags. We injected gRNAs together with Cas9 mRNA and a repair template that contained the coding sequence for mNG2₁₁ and a short linker (Fig. 3A); the repair template consisted of double-stranded DNA with single-stranded homology arms of 30 bp at each end (Liang et al., 2017). To verify that the knock-in was successful, we pooled injected embryos and performed insert-specific PCR that amplified the mNG2₁₁ insertion but not the unedited wild-type (Fig. 3B).

For *tubb4b*, we estimated the knock-in efficiency using quantitative PCR. To determine mNG2₁₁ prevalence, we pooled and extracted DNA from 30 injected F₀ embryos at 24 hpf. We amplified mNG2₁₁ using insert-specific primers and amplified the untargeted, single-copy gene *prox1a* for comparison; we obtained a Ct of 5 cycles between the two. As zebrafish are diploid, *prox1a* is present in two copies per cell, but each mNG2₁₁ knock-in likely occurred only in one *tubb4b* allele per cell. We thus estimated that roughly 1 in every 16 cells in our pooled sample carried the knock-in allele, corresponding to a knock-in efficiency of about 6%, although not necessarily in-frame nor equally distributed among embryos. This knock-in efficiency is on par with other reports of CRISPR-guided knock-in in zebrafish (Auer and Del Bene, 2014; Zhang et al., 2023).

To establish stable, germline-transmitted lines for the mNG2₁₁ insertions, we raised injected F₀ fish to adulthood and identified several founders representing multiple alleles for each gene. Some alleles contained indel mutations at the insertion junctions or within the insertion itself. For example, both alleles recovered for *h2az2b* contained mutations within the mNG2₁₁ sequence and produced very dim fluorescence (Fig. S3). Therefore, we chose to propagate only alleles with precise integration of the mNG2₁₁ sequence, resulting in establishment of one line each for *tubb4b* (*tubb4b^{ucm131}*, referred to here as *mNG2₁₁-tubb4b*) and *krt8* (*krt8^{ucm132}*, referred to here as *krt8-mNG2₁₁*).

To confirm that the mNG2₁₁ tag is functional and does not alter endogenous expression patterns, we injected embryos with mNG2₁₋₁₀ mRNA and qualitatively assessed fluorescence. For *mNG2₁₁-tubb4b*, we observed strong fluorescence at 24 hpf that was especially prominent in the eye and brain (Fig. 3D) and along the neural tube (Fig. 3E). For *krt8-mNG2₁₁*, fluorescence appeared restricted to the skin epidermis at 24 hpf (Fig. 3G, H). These fluorescence patterns are consistent with the reported expression patterns for both *tubb4b* (Thisse and Thisse, 2008; Zhuo et al., 2012) and *krt8* (Fischer et al., 2014; Thisse and Thisse, 2008). At the subcellular level, we observed that fluorescence for both genes was enriched at the cell periphery and excluded from the nucleus, which would be expected for cytoskeletal filaments. For both genes, we observed no fluorescence in uninjected embryos (Fig. 3F, I)

Combinatorial expression of tissue-specific mNG2₁₋₁₀ and mNG2₁₁-tagged proteins

After successfully generating mNG2₁₋₁₀ transgenic lines and mNG2₁₁ insertions, we next determined whether these lines could be combined to achieve tissue-specific protein labeling (Fig. 4A). We crossed each of our mNG2₁₁-tagged lines — *mNG2₁₁-tubb4b* and *krt8-mNG2₁₁* — with each of our mNG2₁₋₁₀ transgenic lines — *fezf2:mNG2₁₋₁₀*, *myl7:mNG2₁₋₁₀*, and *ubb:mNG2₁₋₁₀*. For *mNG2₁₁-tubb4b*, crossing to *ubb:mNG2₁₋₁₀* produced fluorescence broadly throughout the head (Fig. 4B–B'), enabling timelapse analysis of tubulin dynamics in the otic vesicle and surrounding region (Video 1). This fluorescence pattern is similar to mNG2₁₋₁₀ mRNA injection and to the reported expression pattern for *tubb4b* (Thisse and Thisse, 2008; Zhuo et al., 2012). In contrast, crossing to *fezf2:mNG2₁₋₁₀* resulted in fluorescence restricted to the brain and eye (Fig. 4C–C'), consistent with the known expression pattern for *fezf2* (Jeong et al., 2006). Finally, crossing to *myl7:mNG2₁₋₁₀* resulted in no observable fluorescence (Fig. 4D). This result is consistent

with the reported expression pattern for *tubb4b*, which has not been reported to be expressed in the heart.

The results we obtained for *krt8-mNG2₁₁* similarly demonstrated retention of endogenous expression patterns. Crossing to *ubb:mNG2₁₋₁₀* resulted in fluorescence primarily in the skin at 24 hpf (Fig. 4E–E'), similar to mNG2₁₋₁₀ mRNA injection. We also observed fluorescence in cells of the enveloping layer at 10 hpf (Fig. 4F), consistent with the reported expression pattern for *krt8* (Fischer et al., 2014; Thisse and Thisse, 2008). Crossing *krt8-mNG2₁₁* to *fezf2:mNG2₁₋₁₀* or *myl7:mNG2₁₋₁₀* resulted in no observable fluorescence (Fig. 4G–H), which is expected as *krt8* has not been reported to be expressed in either cardiac or neural tissues.

Altogether, our results show that combining transgenic mNG2₁₋₁₀ expression and mNG2₁₁ tagging can achieve tissue-specific fluorescent protein labeling that preserves endogenous expression patterns.

Directing protein localization with split-mNG2

Given that split-mNG2 fragments self-assemble, it may be possible to use mNG2₁₋₁₀ as a “bait” to direct mNG2₁₁-tagged proteins to specific subcellular locations. To determine the feasibility of this application, we fused mNG2₁₋₁₀ to a localization signal for the outer mitochondrial membrane (mito-mNG2₁₋₁₀) (Bear et al., 2000) (Fig. 5A). We then injected mRNA for mito-mNG2₁₋₁₀ into *krt8-mNG2₁₁* embryos. Compared to control embryos injected with untagged mNG2₁₋₁₀ (Fig. 5B–D), embryos injected with mito-mNG2₁₋₁₀ exhibited qualitatively different fluorescence localization patterns that co-localized with the mitochondrial dye MitoTracker (Fig. 5E–G). These results suggest that mito-mNG2₁₋₁₀ is indeed directing mNG2₁₁-tagged Keratin 8 to the mitochondria. Thus, by anchoring mNG2₁₋₁₀ to specific cellular compartments, the split-mNG2 system can be used to manipulate protein localization.

Discussion

In this study, we describe using the mNG2_{1-10/11} split fluorescent protein system to achieve tissue-specific fluorescent labeling of endogenous proteins in zebrafish embryos. We further demonstrate that the split-mNG2 system can be used to control protein localization by anchoring the mNG2₁₋₁₀ fragment to specific cellular compartments.

Similar FP_{1-10/11} systems are now commonly used as endogenous protein labeling tools in cell lines (Cho et al., 2022; Feng et al., 2017; Kamiyama et al., 2016; Leonetti et al., 2016). The popularity of these systems is primarily due to the ease with which the short FP₁₁ sequences can be inserted into gene loci. The general utility of split-FP systems for protein labeling has also been demonstrated in multicellular organisms including zebrafish (Kesavan et al., 2021) and mouse embryos (O'Hagan et al., 2021), but in these studies the corresponding FP₁₋₁₀ fragment was delivered constitutively. Tissue specificity has been achieved in *C. elegans* (Goudeau et al., 2021; He et al., 2019; Hefel and Smolikove, 2019; Noma et al., 2017) and *Drosophila* (Kamiyama et al., 2021) and now in zebrafish (this study). The ability to spatially restrict fluorescent labeling is especially advantageous for

studying the tissue-specific function of an otherwise broadly expressed protein. In such cases, constitutive protein labeling would obscure the area under study due to competing signals coming from surrounding tissues, which cannot be easily removed without advanced microscopy or image processing methods. In contrast, our split-mNG2-based approach can achieve tissue-specific labeling using relatively straightforward and conventional techniques.

In this study, we demonstrated a novel application of the split-mNG2 system to control of protein localization via tethering Keratin 8 to mitochondria (Fig. 5). There are several potential applications for using the split-mNG2 system to experimentally manipulate protein localization. For example, mNG2₁₁-tagged proteins could be sequestered away from their normal site of function to achieve a loss-of-function effect. The same approach could also be used to achieve gain-of-function effects by constitutively anchoring a protein to its site of action or to an ectopic location. This approach could also be used to manipulate the properties of specific organelles or subcellular compartments through recruitment of mNG2₁₁-tagged enzymes. An advantage of the split-mNG2 approach is that successful (mis)localization can easily be confirmed because the reconstituted mNG2_{10/11} complexes retains their fluorescence. When combined with transgenic expression of mNG2₁₋₁₀, this approach can be applied to specific tissues of interest for even broadly expressed proteins.

Previous reports have suggested that not all proteins can be easily labeled with the split-mNG2 system (Cho et al., 2022; Leonetti et al., 2016; O'Hagan et al., 2021). Fluorescent labeling may fail because the target protein does not tolerate mNG2₁₁ tagging. Thus, mNG2₁₁ fusion proteins should be designed using the same considerations as with any epitope tag. Fluorescence brightness might also be a challenge. Split-FP systems are known to be dimmer than their intact counterparts; for example, split-mNG2 is about 60% as bright as intact mNeonGreen (Feng et al., 2017). Thus, even if tagging is tolerated, some proteins may not be expressed at high enough levels to produce a detectable fluorescent signal (Leonetti et al., 2016; O'Hagan et al., 2021). This challenge could be overcome by inserting multiple repeats of the mNG2₁₁ sequence to increase fluorescent signal as has been demonstrated for split-GFP (He et al., 2019; Hefel and Smolikove, 2019; Kamiyama et al., 2016, 2021; Noma et al., 2017). Additionally, a third generation split-mNG system was recently developed and reported to have improved spectral properties (Zhou et al., 2020), which may extend the use of split-FP labeling to low or moderately expressed proteins. A challenge specific to working with multicellular organisms is the difficulty of detecting fluorescence in very thick samples, such as late larval and older zebrafish stages. However, our demonstration that split-mNG2 fluorescence is preserved after paraformaldehyde fixation (Fig. S1) suggests that our method is compatible with tissue sectioning protocols.

In this study, we focused on the use of split-mNG2 as a protein labeling tool. However, the ability to control expression of these protein fragments independently, paired with their ability to self-assemble, could be leveraged for other applications. For example, they could be used as coincidence detectors to monitor cell states or signaling pathway activation. And because fluorescence is only produced when the two fragments bind, they could be used to

visualize interactions at multiple length scales, i.e., between proteins, organelles, cells, or adjacent tissues.

Conclusions

In summary, we have demonstrated that the split-mNG2 system can function in zebrafish to endogenously label proteins in a tissue-specific manner, with other potential applications that make it broadly useful to many areas of investigation.

Materials and Methods

Zebrafish strains

Adult *Danio rerio* zebrafish were maintained under standard laboratory conditions. Zebrafish in an outbred *AB*, *TL*, or *EKW* background were used as wild-type strains. Strains generated in this study are: *Tg(fezf2:mNG2₁₋₁₀)^{ucm120}*; *Tg(my17:mNG2₁₋₁₀)^{ucm121}*; *Tg(ubb:mNG2₁₋₁₀)^{ucm117}*; *krt8^{ucm132}*; and *tubb4b^{ucm132}*. This study was performed with the approval of the Institutional Animal Care and Use Committee (IACUC) of the University of California Merced (Protocol #2023-1144).

mRNA expression

All expression plasmids for *in vitro* mRNA synthesis were generated in a pCS2 backbone. To generate pCS2-GFP₁₋₁₀, GFP₁₋₁₀ was PCR amplified from pACUH-GFP₁₋₁₀ (Bo Huang, University of California San Francisco) and cloned into pCS2 by enzymatic assembly (Gibson et al., 2009). To generate pCS2-mNG2₁₋₁₀, mNG2₁₋₁₀ was PCR amplified from pSFFV-mNG2₁₋₁₀ (Bo Huang, University of California San Francisco) and cloned into pCS2 by enzymatic assembly. To generate pCS2-GFP₁₁-H2B and pCS2-mNG2₁₁-H2B, GFP₁₁ (5'-CGTGACCACATGGTCCTTCATGAGTATGTAATGCTGCTGGGATTACA-3') and mNG2₁₁ (5'-ACCGAGCTCAACTTCAAGGAGTGGCAAAGGCCTTTACCGATATGATG-3') were directly synthesized by Integrated DNA Technologies and H2B was PCR amplified from GFP-H2B (Hesselson et al., 2009); fragments were fused and cloned into pCS2 by enzymatic assembly. To generate pCS2-mito-mNG2₁₋₁₀, the outer mitochondrial membrane signal sequence was PCR amplified from pMSCV-FPPPP-mito (Bear et al., 2000) and cloned into pCS2-mNG2₁₋₁₀ by enzymatic assembly. Capped messenger RNA was synthesized using the mMESSAGE mMACHINE kit (Ambion), and 500 pg of each mRNA was injected at the one- or two-cell stage.

Generation of mNG2₁₋₁₀ transgenic lines

All transgene plasmids were generated in a p μ Tol2 backbone (LaBelle et al., 2021). mNG2₁₋₁₀ and promoter sequences for *fezf2* (Berberoglu et al., 2009), *my17* (Huang et al., 2003), or *ubb* (Mosimann et al., 2011) were PCR amplified then fused and cloned into p μ Tol2 by enzymatic assembly to generate p μ Tol2-fez:mNG2₁₋₁₀, p μ Tol2-my17:mNG2₁₋₁₀, and p μ Tol2-ubb:mNG2₁₋₁₀, respectively. The constructs were used to generate *Tg(fez:mNG2₁₋₁₀)^{ucm120}*; *Tg(my17:mNG2₁₋₁₀)^{ucm121}*; *Tg(ubb:mNG2₁₋₁₀)^{ucm117}* using standard transgenesis protocols (Clark et al., 2011; Kawakami, 2004).

CRISPR/Cas-directed insertion of mNG2₁₁

Guide RNAs (gRNAs) were designed using CRISPRscan (Moreno-Mateos et al., 2015) and synthesized as previously described (Varshney et al., 2016). The double-stranded DNA template for homology directed repair was assembled from two oligomers synthesized by Integrated DNA Technologies. Each oligomer contained the sequence for mNG2₁₁, a 10-amino acids-encoding linker sequence (5'-GGAGCTGGTGCAGGCGCTGGAGCCGGTGCC-3'), and a homology arm. Oligomers were hybridized to obtain a double-stranded template with single-stranded, 30 bp-long homology arms at each end (Liang et al., 2017). gRNAs, donor DNA, and Cas9 mRNA were injected at the one-cell stage as previously described (Gagnon et al., 2014).

To verify insertion, we pooled 40 injected embryos at 24 hpf, isolated genomic DNA, and performed PCR using two sets of primer pairs per gene covering the 5' and 3' insertion sites. The same primer sets were used for quantitative PCR (qPCR) to estimate knock-in efficiency. Each qPCR reaction contained 2X PerfeCTa[®] SYBR Green FastMix (Quantabio), five-fold diluted genomic DNA, and 325 nM of each primer. Reactions were carried out on a QuantStudio3 (Applied Biosystems) real-time PCR machine using the following program: initial activation at 95°C for 10 min, followed by 40 cycles of 30 s at 95°C, 30 s at 60°C, and 1 min at 72°C. Once the PCR was completed, a melt curve analysis was performed to determine reaction specificity. The gene *prox1a* was used as a reference. Primers used in this study (presented 5'–3'):

5' h2az2b-mNG211 forward: TTGTGTGTTTGTGCGTCCGC

5' h2az2b-mNG211 reverse: GCCACTCCTTGAAGTTGAGC

3' h2az2b-mNG211 forward: GCTCAACTTCAAGGAGTGCC

3' h2az2b-mNG211 reverse: ACGAAGCCCCGAAAGCACAC

5' mNG211-krt8 forward: ATACAGCGGCGGATACAGCG

5' mNG211-krt8 reverse: GCCACTCCTTGAAGTTGAGC

3' mNG211-krt8 forward: GCTCAACTTCAAGGAGTGCC

3' mNG211-krt8 reverse: AAGGCACGACAAGAGCGGTG

5' mNG211-tubb4b forward: CACATCTCGAATTACGACCTCA

5' mNG211-tubb4b reverse: GCCTTTTGCCACTCCTTGAAG

3' mNG211-tubb4b forward: GCTCAACTTCAAGGAGTGCC

3' mNG211-tubb4b reverse: AAAACAAGCAAGGATTAGCGTC

prox1a forward: TGTCATTTGCGCTCGCGCTG

prox1a reverse: ACCGCAACCCGAAGACAGTG

To verify germline transmission and establish stable lines, injected F₀ embryos were raised to adulthood then outcrossed to wild-type zebrafish. We pooled 40 of the resulting F₁ embryos at 24 hpf, isolated genomic DNA, and performed PCR using the same primer sets as above. PCR fragments were cloned into pGEM-T (Promega), and the inserts were sequenced by Sanger sequencing (University of California Berkeley DNA Sequencing Facility). Only clutches containing precise insertion of the mNG2₁₁ plus linker sequence were kept for propagation. At adulthood, individual F₁ zebrafish were genotyped by fin clipping using the same primer sets as described above. Only animals containing precise insertion of the mNG2₁₁ sequence were kept for line propagation.

Microscopy and image processing

Dechorionated embryos or larvae were embedded in 1.5% low-melting agarose (ISC BioExpress) containing 0.01% tricaine (Sigma-Aldrich) within glass-bottom Petri dishes (MatTek Corporation). For mitochondria labeling, embryos were incubated in 50 nM MitoTracker Red CMXRos (Invitrogen) for 30 minutes prior to agarose embedding. For paraformaldehyde fixation, embryos were incubated in 4% paraformaldehyde (Sigma-Aldrich) overnight at 4°C, washed three times for 10 min in 1X phosphate buffered saline with 0.1% Tween-20 (Sigma-Aldrich) at room temperature, then embedded in agarose for imaging. Identical image acquisition settings were used for all embryos from the same set of experiments.

Widefield fluorescence and brightfield images were acquired on an Olympus SZX16 stereomicroscope equipped with a DP23 monochrome camera and cellSens software (Evident). Brightfield images were acquired with transmitted light from an LED diascopic base (Evident). GFP or mNG2 fluorescence was excited with an LED light source (X-Cite) and 470/40 nm excitation filter (Chroma) and acquired with a 500 nm long-pass emission filter (Chroma).

Confocal images were acquired on an Olympus IX83 microscope (Evident) equipped with a spinning disk confocal unit (CSU-W1; Andor). Brightfield images were acquired using a transmitted LED light source. GFP or mNG2 fluorescence was excited with a 488 nm 150 mW solid state laser (Visitron Systems) and collected with a 525/50 nm emission filter. Images were acquired with a Prime 95b sCMOS camera (Teledyne Photometrics) controlled with MicroManager software (Edelstein et al., 2014). Z-stack optical sections were collected with a 10x/0.4NA objective lens (Evident) with a step-size of 5 µm or with a 30x/1.05 NA objective lens (Evident) with a step-size of 2 µm using a Piezo focus motor (ASI). For time-lapse experiments, z-stacks with a step-size of 4 µm were collected with a 30x/1.05 NA objective lens every 5 min, using an exposure time of 200 ms and 1×1 camera binning. All z-stacks are displayed as maximum z-projections.

Images were processed identically for each set of experiments using Fiji software (Schindelin et al., 2012) as follows: denoised using the Non-local Means Denoise plugin (Buades et al., 2005), brightness and contrast levels adjusted, converted to 8-bit depth, and cropped. In cases where the region of interest extended beyond the microscope's field of view, multiple images were stitched together using the pairwise stitching plugin. Brightfield

and fluorescence images were merged in Photoshop software (Adobe). Illustrations were created with BioRender (<https://biorender.com>).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Bo Huang (University of California San Francisco) for providing the split-GFP and split-mNG2 plasmids and advice, Anne Pipathsouk (University of California San Francisco) and Manuel Leonetti (Chan Zuckerberg Biohub) for helpful discussions, the Department of Animal Research Services (University of California Merced) for excellent fish care, and members of the Woo and Materna labs for helpful comments on the manuscript. This work was supported by grants from the Society for Developmental Biology, the National Institutes of Health (NIH R15HD102829), and the National Science Foundation (NSF-IOS-2238304) to S.W. J.L. was supported by NIH training grant T32GM141862. K.S. received support from the NSF-CREST: Center for Cellular and Biomolecular Machines at the University of California, Merced (NSF-HRD-1547848).

References

- Albadri S, Del Bene F, Revenu C. Genome editing using CRISPR/Cas9-based knock-in approaches in zebrafish. *Methods San Diego Calif* 2017;121–122:77–85.
- Auer TO, Del Bene F. CRISPR/Cas9 and TALEN-mediated knock-in approaches in zebrafish. *Methods San Diego Calif* 2014;69:142–50. [PubMed: 24704174]
- Bear JE, Loureiro JJ, Libova I, Fässler R, Wehland J, Gertler FB. Negative regulation of fibroblast motility by Ena/VASP proteins. *Cell* 2000;101:717–28. 10.1016/s0092-8674(00)80884-3. [PubMed: 10892743]
- Berberoglu MA, Dong Z, Mueller T, Guo S. fezf2 expression delineates cells with proliferative potential and expressing markers of neural stem cells in the adult zebrafish brain. *Gene Expr Patterns* 2009;9:411–22. 10.1016/j.gep.2009.06.002. [PubMed: 19524703]
- Buades A, Coll B, Morel J-M. A Non-Local Algorithm for Image Denoising. 2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. CVPR05, vol. 2, San Diego, CA, USA: IEEE; 2005, p. 60–5. 10.1109/CVPR.2005.38.
- Cho NH, Cheveralls KC, Brunner A-D, Kim K, Michaelis AC, Raghavan P, et al. OpenCell: Endogenous tagging for the cartography of human cellular organization. *Science* 2022;375:eabi6983. 10.1126/science.abi6983. [PubMed: 35271311]
- Clark KJ, Urban MD, Skuster KJ, Ekker SC. Transgenic zebrafish using transposable elements. *Methods Cell Biol* 2011;104:137–49. [PubMed: 21924161]
- Edelstein AD, Tsuchida MA, Amodaj N, Pinkard H, Vale RD, Stuurman N. Advanced methods of microscope control using µManager software. *J Biol Methods* 2014;1:e10. 10.14440/jbm.2014.36. [PubMed: 25606571]
- Feinberg EH, Vanhoven MK, Bendesky A, Wang G, Fetter RD, Shen K, et al. GFP Reconstitution Across Synaptic Partners (GRASP) defines cell contacts and synapses in living nervous systems. *Neuron* 2008;57:353–63. [PubMed: 18255029]
- Feng S, Sekine S, Pessino V, Li H, Leonetti MD, Huang B. Improved split fluorescent proteins for endogenous protein labeling. *2017;8:370*.
- Fischer B, Metzger M, Richardson R, Knyphausen P, Ramezani T, Franzen R, et al. p53 and TAp63 promote keratinocyte proliferation and differentiation in breeding tubercles of the zebrafish. *PLoS Genet* 2014;10:e1004048. 10.1371/journal.pgen.1004048. [PubMed: 24415949]
- Gagnon JA, Valen E, Thyme SB, Huang P, Akhmetova L, Pauli A, et al. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One* 2014;9:e98186. 10.1371/journal.pone.0098186. [PubMed: 24873830]

- Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 2009;6:343–5. [PubMed: 19363495]
- Goudeau J, Sharp CS, Paw J, Savy L, Leonetti MD, York AG, et al. Split-wrmScarlet and split-sfGFP: tools for faster, easier fluorescent labeling of endogenous proteins in *Caenorhabditis elegans*. *Genetics* 2021;217:iyab014. 10.1093/genetics/iyab014. [PubMed: 33693628]
- Harvey SA, Smith JC. Visualisation and quantification of morphogen gradient formation in the zebrafish. *PLoS Biol* 2009;7:e1000101. [PubMed: 19419239]
- He S, Cuentas-Condori A, Miller DM. NATF (Native and Tissue-Specific Fluorescence): A Strategy for Bright, Tissue-Specific GFP Labeling of Native Proteins in *Caenorhabditis elegans*. *Genetics* 2019;212:387–95. 10.1534/genetics.119.302063. [PubMed: 30952669]
- Hefel A, Smolikove S. Tissue-Specific Split sfGFP System for Streamlined Expression of GFP Tagged Proteins in the *Caenorhabditis elegans* Germline. *G3 Bethesda Md* 2019;9:1933–43. 10.1534/g3.119.400162. [PubMed: 30992318]
- Hesselson D, Anderson RM, Beinat M, Stainier DYR. Distinct populations of quiescent and proliferative pancreatic beta-cells identified by HOTcre mediated labeling. *Proc Natl Acad Sci U S A* 2009;106:14896–901. 10.1073/pnas.0906348106. [PubMed: 19706417]
- Huang C-J, Tu C-T, Hsiao C-D, Hsieh F-J, Tsai H-J. Germ-line transmission of a myocardium-specific GFP transgene reveals critical regulatory elements in the cardiac myosin light chain 2 promoter of zebrafish. *Dev Dyn Off Publ Am Assoc Anat* 2003;228:30–40.
- Jeong J-Y, Einhorn Z, Mercurio S, Lee S, Lau B, Mione M, et al. Neurogenin1 is a determinant of zebrafish basal forebrain dopaminergic neurons and is regulated by the conserved zinc finger protein Tof/Fezl. *Proc Natl Acad Sci U S A* 2006;103:5143–8. 10.1073/pnas.0600337103. [PubMed: 16549779]
- Kamiyama D, Sekine S, Barsi-Rhyne B, Hu J, Chen B, Gilbert LA, et al. Versatile protein tagging in cells with split fluorescent protein. *Nat Commun* 2016;7:11046. [PubMed: 26988139]
- Kamiyama R, Banzai K, Liu P, Marar A, Tamura R, Jiang F, et al. Cell-type-specific, multicolor labeling of endogenous proteins with split fluorescent protein tags in *Drosophila*. *Proc Natl Acad Sci U S A* 2021;118:e2024690118. 10.1073/pnas.2024690118. [PubMed: 34074768]
- Kawakami K. Transgenesis and gene trap methods in zebrafish by using the Tol2 transposable element. *Methods Cell Biol* 2004;77:201–22. [PubMed: 15602913]
- Kesavan G, Machate A, Brand M. CRISPR/Cas9-Based Split Fluorescent Protein Tagging. *Zebrafish* 2021;18:369–73. 10.1089/zeb.2021.0031. [PubMed: 34495758]
- Kimura Y, Hisano Y, Kawahara A, Higashijima S-I. Efficient generation of knock-in transgenic zebrafish carrying reporter/driver genes by CRISPR/Cas9-mediated genome engineering. *Sci Rep* 2014;4:6545. [PubMed: 25293390]
- LaBelle J, Ramos-Martinez A, Shen K, Motta-Mena LB, Gardner KH, Materna SC, et al. TAEL 2.0: An Improved Optogenetic Expression System for Zebrafish. *Zebrafish* 2021;18:20–8. [PubMed: 33555975]
- Leonetti MD, Sekine S, Kamiyama D, Weissman JS, Huang B. A scalable strategy for high-throughput GFP tagging of endogenous human proteins. *Proc Natl Acad Sci U S A* 2016;113:E3501–8. [PubMed: 27274053]
- Liang X, Potter J, Kumar S, Ravinder N, Chesnut JD. Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA. *J Biotechnol* 2017;241:136–46. [PubMed: 27845164]
- Moreno-Mateos MA, Vejnár CE, Beaudoin J-D, Fernandez JP, Mis EK, Khokha MK, et al. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods* 2015;12:982–8. [PubMed: 26322839]
- Mosimann C, Kaufman CK, Li P, Pugach EK, Tamplin OJ, Zon LI. Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish. *Dev Camb Engl* 2011;138:169–77.
- Noma K, Goncharov A, Ellisman MH, Jin Y. Microtubule-dependent ribosome localization in *C. elegans* neurons. *eLife* 2017;6:e26376. 10.7554/eLife.26376. [PubMed: 28767038]

- O'Hagan D, Kruger RE, Gu B, Ralston A. Efficient generation of endogenous protein reporters for mouse development. *Dev Camb Engl* 2021;148:dev197418. 10.1242/dev.197418.
- Rosen JN, Sweeney MF, Mably JD. Microinjection of zebrafish embryos to analyze gene function. *J Vis Exp JoVE* 2009;1115. 10.3791/1115.
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012;9:676–82. 10.1038/nmeth.2019. [PubMed: 22743772]
- Simiczyjew A, Mazur AJ, Popow-Wo niak A, Malicka-B\laskiewicz M, Nowak D. Effect of overexpression of β - and γ -actin isoforms on actin cytoskeleton organization and migration of human colon cancer cells. *Histochem Cell Biol* 2014;142:307–22. [PubMed: 24682235]
- Thisse C, Thisse B. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat Protoc* 2008;3:59–69. [PubMed: 18193022]
- Varshney GK, Carrington B, Pei W, Bishop K, Chen Z, Fan C, et al. A high-throughput functional genomics workflow based on CRISPR/Cas9-mediated targeted mutagenesis in zebrafish. *Nat Protoc* 2016;11:2357–75. 10.1038/nprot.2016.141. [PubMed: 27809318]
- Zhang Y, Marshall-Phelps K, De Almeida RG. Fast, precise and cloning-free knock-in of reporter sequences *in vivo* with high efficiency. *Development* 2023;150:dev201323. 10.1242/dev.201323. [PubMed: 37309812]
- Zhou S, Feng S, Brown D, Huang B. Improved yellow-green split fluorescent proteins for protein labeling and signal amplification. *PLoS One* 2020;15:e0242592. 10.1371/journal.pone.0242592. [PubMed: 33227014]
- Zhuo H-Q, Huang L, Huang H-Q, Cai Z. Effects of chronic tramadol exposure on the zebrafish brain: a proteomic study. *J Proteomics* 2012;75:3351–64. 10.1016/j.jprot.2012.03.038. [PubMed: 22507199]

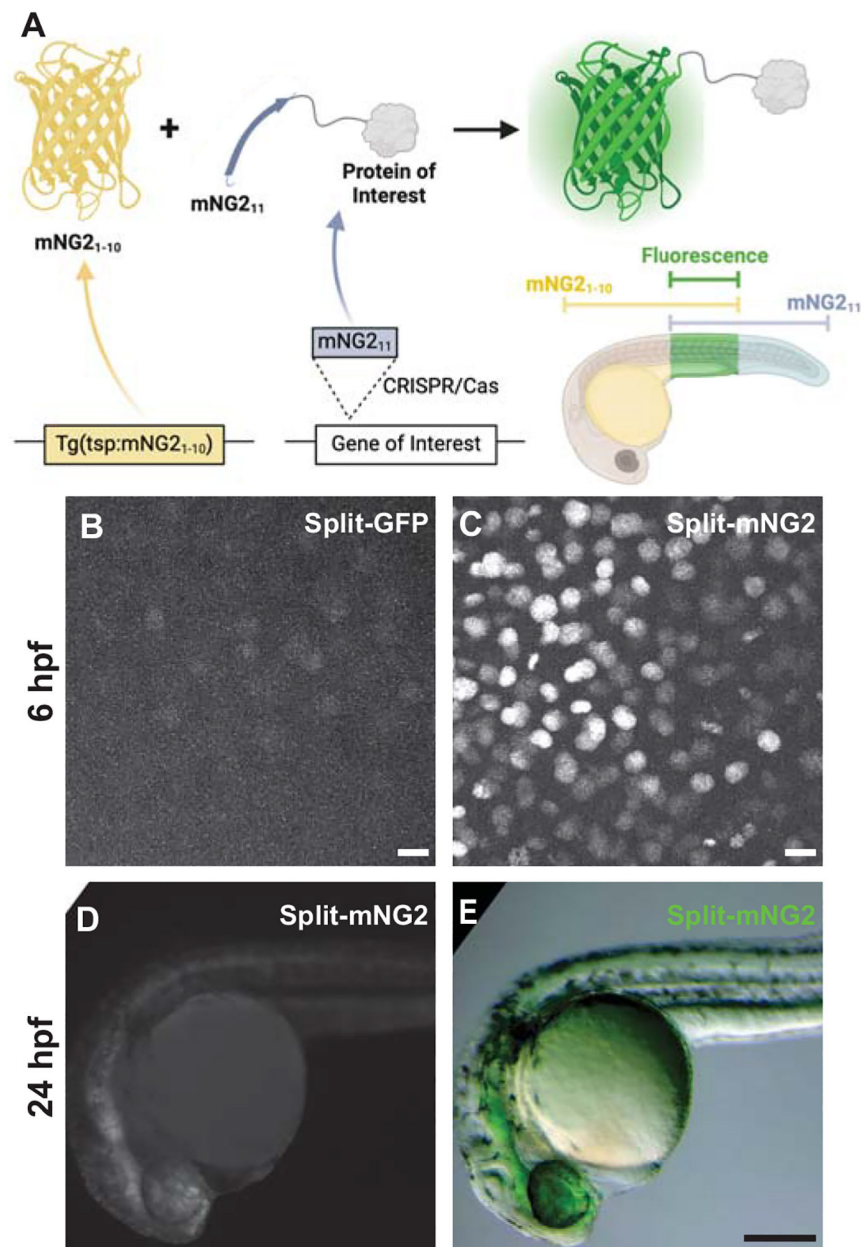


Figure 1. Split fluorescent protein fragments are functional in zebrafish embryos.
A. Schematic illustrating our protein labeling strategy using a split fluorescent protein. Transgenic (Tg) mNG2₁₋₁₀ is expressed under the control of a tissue-specific promoter (tsp) while mNG2₁₁ is inserted into protein-coding genes by CRISPR/Cas-directed gene editing. Fluorescence (green) is only generated in tissues co-expressing mNG2₁₋₁₀ and the mNG2₁₁-tagged protein of interest. **B–E.** Embryos were injected with GFP₁₋₁₀ and GFP11-H2B (split-GFP, B) or mNG2₁₋₁₀ and mNG211-H2B (split-mNG2, C–E) mRNAs then imaged at 6 hours post-fertilization (hpf) on a confocal microscope (B, C) or at 24 hpf on a fluorescence stereomicroscope (D, E). Confocal images are displayed as maximum z-projections. Scale bars in B and C, 50 μ m. Scale bar in E, 200 μ m.

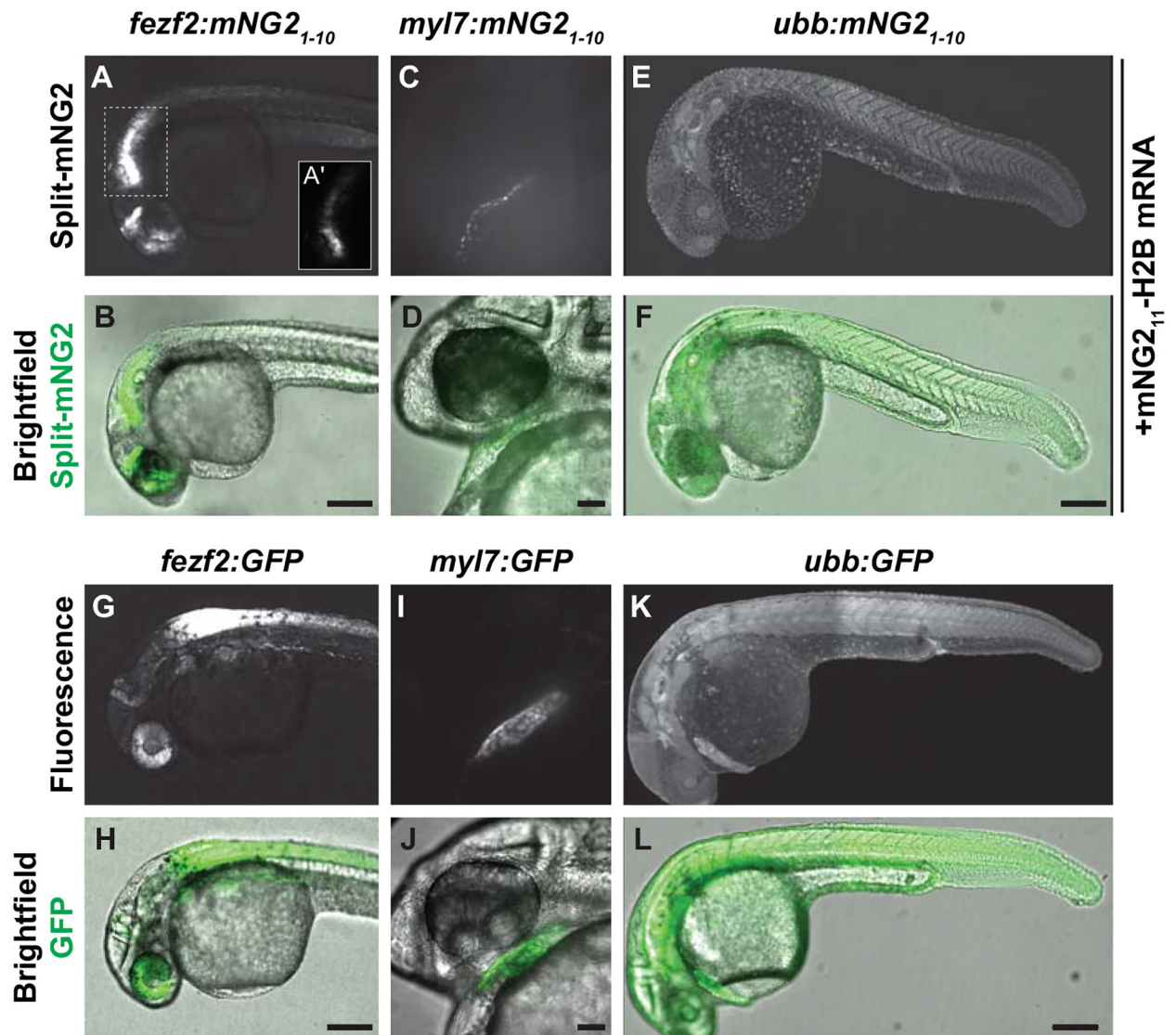


Figure 2. Split fluorescent protein labeling can be spatially restricted by transgenic expression of mNG2₁₋₁₀.

A–F. Transgenic embryos expressing mNG2₁₋₁₀ under control of the *fezf2* (A–B), *myl7* (C–D), or *ubb* (E–F) promoters and injected with mNG2₁₁-H2B mRNA. A' shows the boxed region in A with brightness rescaled to demonstrate fluorescence is localized to nuclei. G–L. Transgenic embryos expressing GFP under control of the *fez1* (G–H), *myl7* (I–J), or *ubb* (K–L) promoters. Images were acquired at 24 hours post-fertilization. Fluorescence images are maximum projections of confocal z-stacks. Scale bars in B, F, H, and L, 200 μm. Scale bars in D, J, 50 μm.

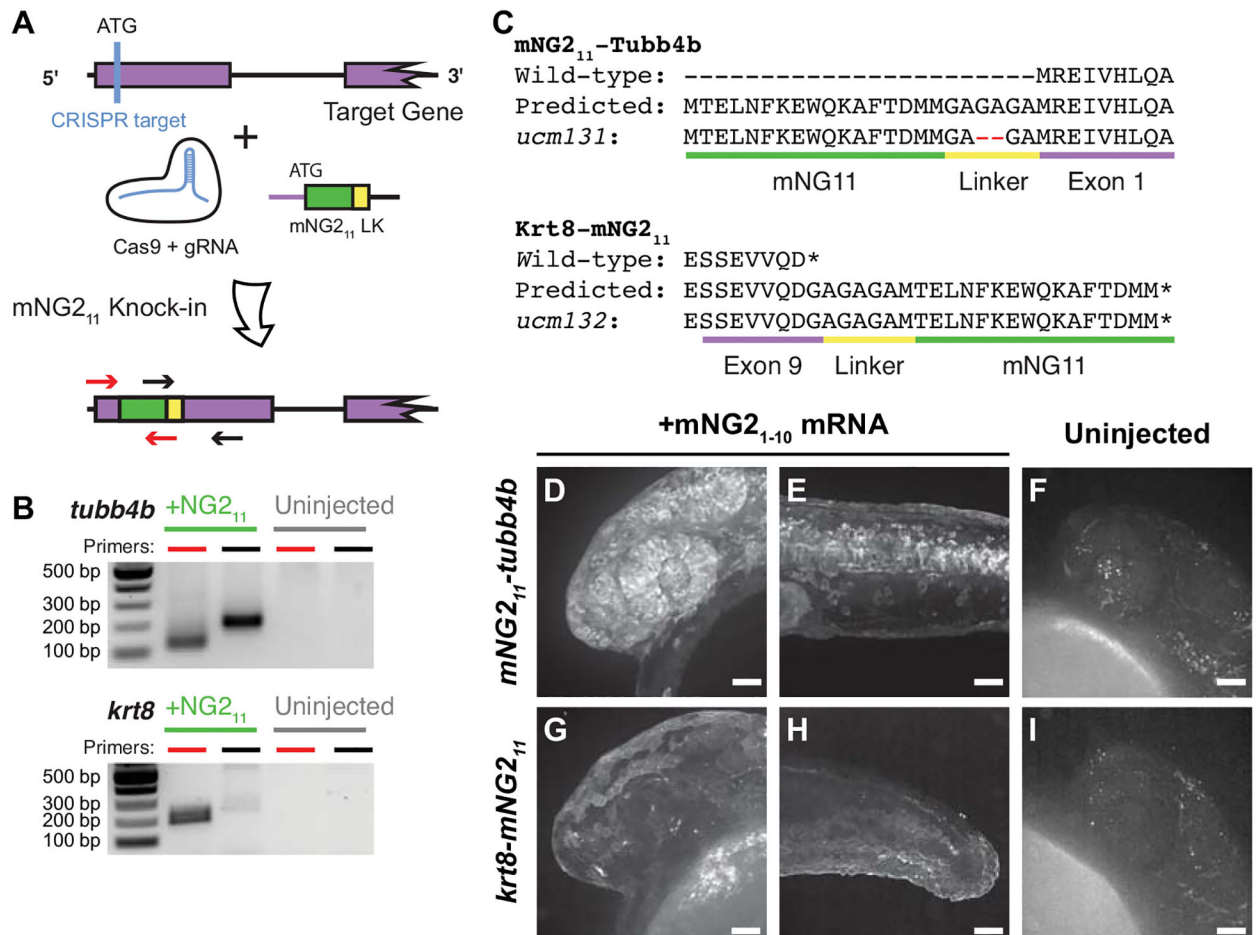


Figure 3. mNG2₁₁ tagging by CRISPR/Cas-directed gene editing.

A. Schematic of CRISPR/Cas-directed mNG2₁₁ insertion into target genes. Purple, endogenous exon sequence. Green, mNG2₁₁. Yellow, linker (LK). ATG, start codon. Arrows denote primers used in **B**. **B.** mNG2₁₁ insertion was assessed by PCR. The primers used correspond to the arrows shown in **A**. bp, base pairs. **C.** Amino acid sequences of wild-type, predicted mNG2₁₁ fusions, and recovered alleles for Tubb4b and Krt8. Mismatches between the predicted and recovered sequences are highlighted in red. Asterisks, stop codons. **D–I.** Representative images of *mNG2₁₁-tubb4b* (**D–F**) and *krt8-mNG2₁₁* (**G–I**) embryos injected with mNG2₁₋₁₀ mRNA (**D–E**, **G–H**) or uninjected (**F**, **I**). Maximum projections of confocal z-stacks. Images were acquired at 24 hours post-fertilization. Images in **F** and **I** have been overexposed to emphasize lack of fluorescence. Scale bars, 50 μm.

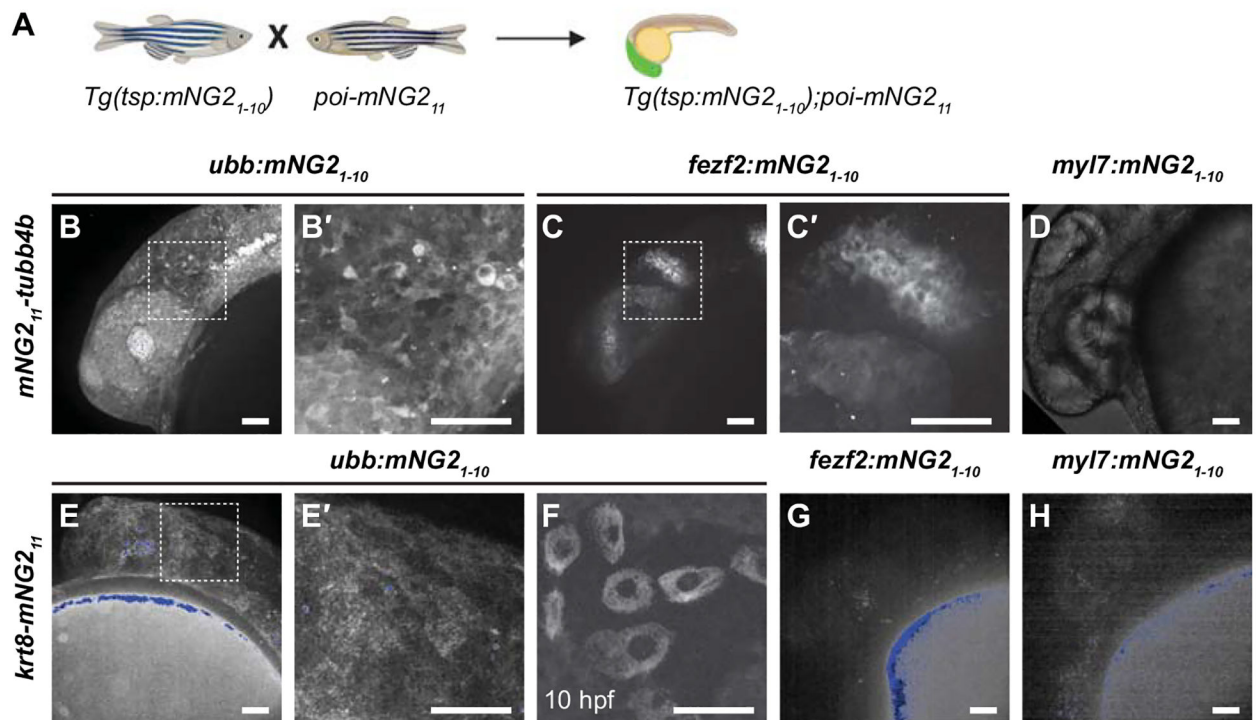


Figure 4. Combinatorial expression of tissue-specific $mNG2_{1-10}$ and $mNG2_{11}$ -tagged proteins.
A. Schematic of crossing strategy. tsp, tissue-specific promoter. poi, protein of interest.
B-D. Representative images of embryos obtained from crossing $mNG2_{11}-tubb4b$ and $ubb:mNG2_{1-10}$ (B-B'), $fezf2:mNG2_{1-10}$ (C-C'), or $myl7:mNG2_{1-10}$ (D). B' and C' show boxed regions in B and C, respectively. **E-H.** Representative images of embryos obtained by crossing $krt8-mNG2_{11}$ to $ubb:mNG2_{1-10}$ (E-F), $fezf2:mNG2_{1-10}$ (G), or $myl7:mNG2_{1-10}$ (H). E' shows boxed region in E. Maximum projections of confocal z-stacks. Images were acquired at 24 hours post-fertilization (hpf) unless otherwise noted. Image in D has been overexposed to emphasize lack of fluorescence. Autofluorescent speckles (yolk, pigment cells, and debris) are colored blue for display purposes. Scale bars, 50 μm .

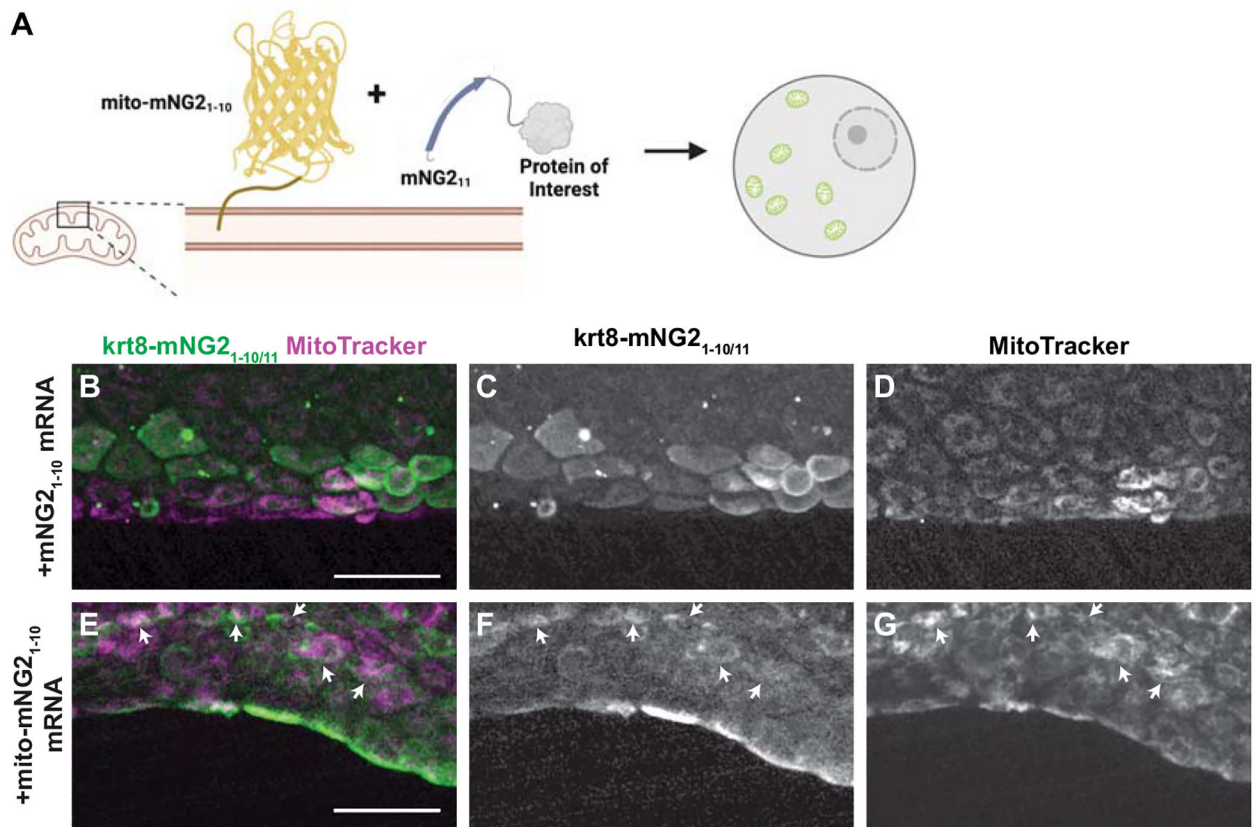


Figure 5. Directing protein localization with split-mNG2.

A. Schematic illustrating use of the split-mNG2 system to sequester proteins of interest on mitochondria. **B–G.** Representative images of *krt8-mNG2₁₁* embryos injected with *mNG2₁₋₁₀* (B–D) or *mito-mNG2₁₋₁₀* (E–G) mRNA and stained with MitoTracker dye to label mitochondria. Maximum projections of confocal z-stacks. Arrows indicate colocalization between split-mNG2 and MitoTracker fluorescence. Images were acquired from the tail fin epidermis at 48 hours post-fertilization (hpf). Scale bars, 50 μm .