Deterministic Barcoding of Neuron Identities through Multicolor Fluorescent Markers in *C. elegans*

by

Seung Hyeon Shim

Submitted to the Department of Biological Engineering
in partial fulfillment for the degree(s) of

Bachelors of Science in Biological Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2023

Authored by: Seung Hyeon Shim
Department of Biological Engineering
May 12th, 2023

Certified by: Edward S. Boyden
Professor of Biological Engineering
Thesis Supervisor

Accepted by: Forest White
Professor of Biological Engineering
UG Chair, Department of Biological Engineering

Deterministic Barcoding of Neuron Identities through Multicolor Fluorescent Markers in *C. elegans*

By

Seung Hyeon Shim

Submitted to the Department of Biological Engineering on May 12th, 2023.

Abstract

Mapping the connectivity of neurons, or the brain connectome, with a clear identification of neuron types and functions is a complicated problem that hinders the development of potential therapeutic interventions. Without understanding the identities of neurons that have fundamental roles in neurological or psychiatric disorders, uncovering the disease mechanism and pathology to observe, control, and even repair neural dynamics remains challenging. In this work, we achieve live animal imaging and cell identification by generating a multicolored fluorescent map of different neurons in *C. elegans*. We take advantage of neuron type-specific promoters to construct transgenes that label each neuron type with unique red, green, and blue fluorescent colors. We also tag subcellular markers of the nucleolus and nuclear membrane with fluorescent proteins to allow for a combinatorial barcode for neuron types, allowing the differentiation of neurons at a single-cell resolution. Through multicolor barcoding of individual neurons, we distinguish different cell types *in vivo*, helping reconstruct parts of the *C. elegans* connectome. With this deterministic strategy, we hope to enable cell typing throughout development and advance the understanding of both the connectome and the shifts in neural circuitry.

Thesis supervisor: Edward S. Boyden

Title: Professor of Biological Engineering

Acknowledgements

I thank Professor Edward S. Boyden and the Synthetic Neurobiology Group for providing me with space and resources to enable this work.

Moreover, I thank Yangning Lu, my direct research mentor, for his invaluable mentorship and feedback on experimental designs and my research progress. Yangning guided me through ideations to executing experiments and helped me to think critically and mature in my analytical reasoning. Without his support, this thesis would not have been possible.

I am also grateful to Anna Le for thesis writing advice, who has served as my communication mentor along with Yangning.

Finally, I would like to thank the Biological Engineering Undergraduate Department for this opportunity to present my research work as part of my thesis. This research endeavor would not have been possible without all the support from my mentors, lab, and the department.

Table of Contents

List of Figures

List of Tables

1. Introduction

Approximately one in six people, or greater than one billion worldwide, suffer from neurological or psychiatric disorders.[1,2] Despite the number of patients affected by brain-related diseases, most underlying causes of these conditions are still unknown.[3] With more than 85 billion neurons each making up to 15,000 synaptic connections, the complex structure of a human brain poses challenges for characterizing the unique identities of individual neurons that impact the disease pathology.[4–6] Therefore, comprehensive mapping of the pan-neuronal pathways, or the brain connectome, is imperative in understanding their involvement in disease progression and accelerating the development of potential effective therapeutic diagnostics and treatments.[7]

The billions of neurons largely communicate via endogenous chemical molecules known as neurotransmitters and neuropeptides. Neurotransmitters are released in vesicles at the synaptic junction, which is the site of electrical or chemical transmission between two neuronal processes.[8] There are various types of neurotransmitters, including dopamine, serotonin, acetylcholine, glutamine, serotonin, and gamma-aminobutyric acid (GABA), which uniquely determine each cell's function and alter the brain circuitry. For instance, glutaminergic neurons are excitatory and mediate neural plasticity for learning and memory, while GABAminergic neurons are inhibitory, regulating cell growth and differentiation.[9,10] Dopaminergic neurons also play an important role in essential brain functions like learning, reward, attention, motivation, and movement.[9,10] Similar to neurotransmitters, neuropeptides are diverse groups of small proteinaceous messengers released by neurons for cell communication. With over 100 different neuropeptides, they perform numerous functions, including direct regulation of neurotransmitters and modulation of both short and long range hormone pathways.[11,12] Altogether, these interneuronal variations add to the brain's complicated composition, making it difficult to resolve structural and functional connectivity.

Hence, to understand the fundamental mechanisms of neural dynamics, it is important to accurately identify cell types in live animals, resolve their functions in physiological conditions, and map the whole intact samples. *Caenorhabditis elegans* has been commonly used as a model organism to observe how the nervous system develops over time, due to its rapid life cycle, minimal maintenance requirements, self-fertilization, and transparency that enables efficient

imaging.[13,14] Moreover, they have already been studied in detail with known classification, lineage, morphology, and *in vivo* location of exactly 302 neurons, which facilitates identifying changes in neural connectivity.[14] As the brain composition alters over the course of development or at the onset of psychiatric or neurological diseases, having an approach to tracking the dynamic changes of the brain connectome and identifying the participating cell types would further elucidate the mechanisms of disease initiation and progression toward finding appropriate therapeutic solutions.

Common methods for cell type identification include immunostaining, RNA-fluorescence in situ hybridization (FISH), and single-cell sequencing. Immunostaining uses fluorescently labeled antibodies to selectively detect a target protein and measure protein expression, localization, and distribution patterns through the antigen-antibody complex.[15,16] Another technique, RNA FISH, induces binding between a specific mRNA transcript and DNA or RNA probes in a thermodynamically favorable condition to visualize gene expression, differentiating heterogeneity among individual cells.[17,18] Likewise, high-throughput single-cell sequencing extracts and amplifies the genome, epigenome, transcriptome, or proteins to classify cell types and determine their spatiotemporal roles, such as in development and disease pathology.[19,20] However, while robust at resolving complex cellular compositions, these methodologies require processed samples, and thus, only generate a snapshot of the cell type information, preventing cell identification in live animals in real time. In many cases, these snapshots will also bring issues in image registration.

Methods such as Brainbow and Bitbow address some of these challenges by expressing different combinations of fluorescent proteins in distinct cell populations *in vivo*, but at random.[21–23] This stochastic expression of fluorescent reporters prevents selective identification of cell characteristics in a transgenic animal. Previously, Yemeni *et al.* addressed these challenges by fluorescently labeling cell nuclei to map all neurons of nematode *C. elegans*, enabling deterministic whole-brain imaging in live animals.[6] Yet, without a multicolor palette that can also visualize neuron morphologies, it is challenging to elucidate the brain connectome and cell functions at different time points, such as during development, or task training. Having an additional description to invariantly identify neurons and their connections would thereby better inform the precise mapping of the brain.

Here, we encoded cell type information by deterministically color coding the cytosol of neurons, including neuronal processes, to visualize cell morphology and identify their functions *in vivo*. To differentiate neighboring cells and label individual neuron types in *C. elegans*, we developed a strategy for live labeling of seven neurotransmitters and two neuropeptides by driving the expression of fluorescent protein combinations under cell type-specific promoters. Because the color codes were limited to their encoding capacity, we specifically tagged selective fluorescent proteins to known subcellular structures with highly distinguishable morphologies, and utilized these subcellular structures to expand the number of encodable cell types. With this method, we report successful *in vivo* identification of neuron types in young adult *C. elegans*.

2. Methods

2.1 Generation of promoter-specific forward and reverse primers

All PCR and sequencing primers were designed using the Wormatlas (https://www.wormatlas.org/), UCSD genome browser (https://genome.ucsc.edu/), Wormbase (https://wormbase.org/#012-34-5), NCBI Primer Blast (https://www.ncbi.nlm.nih.gov/tools/primer-blast/), and SnapGene software. The forward and reverse primers to clone the promoter regions of specific cell types were designed to have 45-60% GC content, 55-70 ℃ melting temperature, and based on previously published regulatory sequence lengths and findings in Ortiz *et al.* for *gcy-13*,[39] Sze *et al.* for *tph-1*,[40] Suo *et al.* for *tbh-1*,[41] and Hobert lab database for *flp-7* and *flp-20*.[32] Flanking sequences were added to enable digestion with FseI (actgggccggcc) and AscI (tagaggcgcgcc) enzymes. Finally, the custom designed forward and reverse primers were ordered from Integrated DNA Technologies (IDT). Primer sequences used to clone the promoters are shown in Table 1.

2.2 Cloning neuron-specific promoters using PCR[42]

The neuron-specific promoters were acquired by running primer-specific PCR on the N2 genome. First, the NemaMetrix worm lysis protocol was used to extract the worm's genomic DNA. The ordered primers (IDT) were diluted to 100 µM with UltraPure water. Next, they were vortexed, dissolved at 60℃ for 30 minutes, and vortexed again prior to centrifugation at 12,000g for 1 min. To make a working solution of 5 µM, the primers were diluted again in a 1:20 ratio, where 10 µL of both forward and reverse primers were tandemly mixed into the 1.5 ml microfuge tube with 180 µL of water. 50 µL of PCR reactions with 2 µL of genome DNA per promoter were carried out according to the Q5 High-Fidelity 2X Master Mix protocol (New England Biolabs). The melting temperature for promoter regions of gcy-13, tph-1, tbh-1, flp-7, and flp-20 were 59℃, 65℃, 66℃, 66℃, and 63℃, respectively (calculated using the New England Biolabs TM calculator). The PCR products were run on the 1% agarose gel for 30 minutes. After the confirmation of appropriate bands, the gel was extracted and cleaned using the Monarch DNA Gel Extraction Kit (New England Biolabs). After the PCR clean up, the DNA concentration was determined with the Nanodrop (ThermoFisher).

2.3 Digestion and ligation

The plasmids composed of the neuron-specific promoter, fluorescent markers, and ampicillin resistance (Table 2) were generated through rounds of digestion and ligation (Figure 1). The backbone vectors and insert fragments were digested using 1 µL of FseI and AscI digestive enzymes, 5 µL of Instant Sticky-end Ligase Master Mix (New England Biolabs), 2 µL of vector, 10 to 13 µL of insert based on their concentrations, and UltraPure water up to 50 µL. The solution was incubated at 37℃ for 1 hr, loaded on a 1% agarose gel, and cleaned per the previous protocol. The concentration of digested inserts and vectors was calculated before the ligation. Following the Instant Sticky-end Ligase Master Mix Ligation Protocol (New England Biolabs), 3 µL of vectors, 2 µL of inserts, and 5 µL of the Instant Sticky-end Ligase Master Mix were mixed and set on ice for 1 hr if the concentration of the insert was less than 10 ng/µL.

2.4 *E. coli* transformation

The plasmids were transformed into *E. coli* for amplification per the New England Biolabs high efficiency transformation protocol. First, 2-5 µL plasmid DNA was mixed with 25 µL of competent New England Biolabs 5-alpha *E. coli* cells and incubated on ice for 30 min. Then, the DNA was introduced to the cells via heat shock for 30 s at 42°C. The tube was placed on ice for 5 min. After adding 950 µL of SOC media, cells recovered for 1 hr at 37°C. 2 µL and 20 µL of the transformation were plated onto an Luria-Bertani (LB) broth + carbenicillin agar plate and incubated at 37°C overnight for low and high density colony formations, respectively. Approximately 4 colonies per plasmid were picked and mixed with 30 µL of autoclaved water in a 1.5 mL centrifuge tube, of which 15 µL was transferred into the culture tube along with 4 µL of carbenicillin and 4 mL of LB growth media in a 1:1000 dilution. Then, the tubes were shaken overnight in a 37°C incubator at 225 rpm.

2.5 Clone purification and sequence analysis

Plasmid DNA was harvested from *E. coli* and their sequence was analyzed to confirm the proper generation of clones. The remaining 15 µL of each colony was sent for Sanger sequencing to IDT along with 10 µL of forward (5' - ACCGTTCGAGTAGATGTGGG - 3', 5' - CGTGACAGTCAAGATGGAAGT - 3', 5’ - CCGGAATCTTCCTTATGTATCTC - 3’, 5' - ACTGTTGCTTGGTCTTGTGC - 3', 5' - GTTGCAGTGAAGTTTGAGTTCCA - 3',) and reverse primers (5' - GCCACCGTCTTCGTATGTG - 3' for mTagBFP and 5' -

CTCCTGGCATTTGAACTGGTTT - 3' for mScarlet) for *gcy-13*-FIB-1-mTagBFP, *tph-1*-EMR-1-mScarlet, *tbh-1*-FIB-1-mTagBFP, *flp-7*-mTagBFP, and *flp-20*-mScarlet, respectively. After confirming the correct assembly, the plasmids were extracted using the Monarch Plasmid Miniprep Kit (New England Biolabs). SnapGene software was used to align the plasmid sequence with the sequencing results and whole plasmid sequencing (Primordium) was performed for the second confirmation of the correct plasmid. Then, another set of transformation and plasmid extraction was performed in an RNAse-free environment (NucleoSpin Plasmid purification kit, Macherey-Nagel) to prepare for an injection.

2.6 Digestion Mapping

In addition to sequencing, digestion mapping was performed to verify the correct plasmid generation by comparing the lengths of digested fragments with expected results. The digestion protocol was the same as previously mentioned with HincII, EcoRI, HincII, EcoRI, and NdeI as the digestive enzymes used for *gcy-13*, *tph-1*, *tbh-1*, *flp-7*, and *flp-20*, respectively. The digested products were run on 1% agarose gel for 30 minutes before imaging (Image Lab Touch Software, BioRad).

2.7 Microinjection

The constructs were microinjected into syncytial gonads of *C. elegans* for the generation of transgenic animals following the Wormbook protocol with appropriate modifications.[43,44] The injection solution was prepared by mixing a linearized pan neuronal GFP marker (Synthetic Neurobiology Group) and the constructs in a 1:10 dilution. Glass capillary-based needles (1.0mm, Kwik-Fil Borosilicate, World Precision Instruments, Inc.) were prepared with a needle puller with settings at heat:542, pull:45, vel:75, del: 90, and p500 (Sutter Instrument p-97). Then, approximately 0.5 μL of the injection solution was loaded into the unpulled ends of the capillary needles under the light microscope (ZEISS SteREO Discovery.V12) and set for 5 min. In the meantime, a needle breaker was made by pulling glass capillary needles over fire and placing the pulled portion on a glass cover slide with a drop of microinjection oil (Halocarbon oil 700, Sigma Aldrich). When the solution accumulated to the tip of the needle, the needle was mounted to a differential interference contrast (DIC) microscope (Nikon ECLIPSE Ts2R) at about 45°. Both the coarse and fine focus knobs were used to bring the needle tip close to the breaker. After

applying pressure supply at 1200 pi (Eppendorf™ FemtoJet™ 4i Microinjector), a gentle nudge of the scope table was used to break the needle, and application of the N_2 flow was discontinued.

After preparing the needle, the light microscope was used to pick 2-3 young N2 wild-type adults, which were swirled in oil to remove bacteria, immobilized to a dry agarose pad (3% agarose in M9 solution, dried on a glass slide) using a picker, and relocated to the DIC scope. Both the worm and the needle were focused using 10X and 40X DIC objectives with the distal arm of the gonad (clear chamber) facing the needle on the same plane. When their focus aligned, the worm was gently pierced by the needle and the injection solution was administered for about 4-5 s until the worms flooded. The swelling of the gonad indicated a successful injection, and the worms were released from the needle. The injected animals were picked and transferred to a newly *E. coli* seeded plate with a drop of M9 solution [3 g KH2PO4, 6 g Na2HPO4, 5 g NaCl, 1 ml 1 M MgSO4, H2O to 1L. Sterilized by autoclaving][45] under the light microscope for recovery. Finally, the fluorescing worms were selected and maintained until the F3 generation per Wormbook protocol.[45]

2.8 Sample Fixation and Gelation

Prior to imaging, the worms were fixed and gelled to conserve their configurations. First, a plate full of well-fed worms was washed in an M9 buffer using a glass pipette into a 15 ml tube, which was centrifuged at 50 g at 22 ℃ for 2 min. After removing the supernatant, the wash was repeated. Then, the worms were transferred into a 1.5 ml low-binding tube using a low-binding pipette with 1000 µL of M9 buffer and left on ice for 1 hr along with the fixative reagents. All the rest of the fixative steps were carried out on the ice. Within the last 30 minutes, 10 ml of fixative solution was prepared using 4X paraformaldehyde (PFA), 1X phosphate-buffered saline (PBS), and UltraPure water at pH 7.4. The worms were centrifuged at 400 g at 22 ℃ for 2 min and 1 ml of fixative was added to the tube after removing the supernatant. The sample was incubated at 4 ℃ for 4 hr and then in 1X PBS on ice for 30 min before washing with 1X PBS four times by centrifuging at 400 g at 22 ℃ for 2 min. The worm density was adjusted by adding 200 µL of PBS.

Subsequently, a gelatin chamber was made using a 25 mm x 75 mm glass slide as the base, two 22 mm x 22 mm coverslips stacked on each other on both sides, and finally, a 22 mm x 40 mm rectangular coverslip on the top. After closing the spaces between the coverslips using water droplets, the 50 µL of the fixed worms were gelled by adding 25 µL of 2x gel solution

[10% (w/v) acrylamide, 200 mM N,N'-diallyl-L-tartardiamide (DATD), 2x PBS], 15 µL of water, 5 µL of 10% (w/w) N,N,N',N'-tetramethyl-ethylenediamine (TEMED) stock solution, and 5 µL of 10% (w/w) ammonium persulfate (APS), in order. All these reagents were cooled on ice before mixing. Immediately, the worms in the gel solution were pipetted 10 times for thorough mixing and transferred into the gap between the bottom glass slide and the top coverslip. The gelatin chamber was filled with N2 for 5 min, incubated at humid 37 ℃ for 30 min, and removed from the gelation setup using a razor blade and paint brushes. 1-2 mm of the edges were removed and they were kept in PBS until imaging.

2.9 Characterizing fluorescence through confocal microscopy

The success of generating transgenic animals was confirmed by imaging their fluorescence under the confocal microscope (Nikon ECLIPSE Ti). If the worms weren't fixed and gelled, to immobilize the worms throughout imaging, a drop of 5% agarose in M9 solution was applied to the microscope slide, and another slide was used to thinly spread the solution.[46] After cooling, about 10 worms were picked and transferred into the droplet of M9 solution on the gel. Finally, a drop of NaN3 was added and a cover slip on the top was placed for imaging. If the worms were gelled, then the gel was transferred to a 6 well plate. Then, 2% of low gelling temperature agarose microwaved for 15 s was cooled and pipetted around the gel. Four channels (Brightfield, RFP channel (peak Excitation (Ex): 561 nm, Emission (Em): 582 nm), GFP channel (Ex: 488 nm, Em: 525 nm), and BFP channel (Ex: 405 nm, Em: 455 nm)) were used to identify neuron types and the image was collected using the NIS Element software. After collecting volumes of transgenic worms, the image was reconstructed using ImageJ software and neuron identities were characterized by comparing their color, morphologies, and locations to the established worm data.[27,35,47]

3. Results

The cellular composition is the structural basis for a tissue or an intact animal. By identifying neuron types and circuitry, their role in processes like development and disease onset and progression can be better understood, and therapeutic targeting can become more specific and localized. Nonetheless, monitoring cell types over time is challenging without a tool to deterministically characterize distinct neurons *in vivo*. Therefore, to distinguish different cells and label their functional identities, we chose to express fluorescent reporters under cell type-specific promoters. Briefly, we cloned promoters of different neuron types and color-coded them with fluorescent proteins (Figure 1A). Based on the sequencing results, the cloning strategy was devised for each construct, such as using different primers and melting temperatures to run PCR, to ensure the successful generation of plasmids. Then, these constructs were expressed in live worms to validate their expected fluorescent patterns and identify neuron types (Figure 1B).

3.1 Cloning neuron-specific promoters

First, we selected promoters that allow us to distinguish key neuronal cell types. The promoter region contains transcription regulatory sequences that are crucial for controlling gene expression.[24] Proteins like RNA polymerase and transcription factors can recognize and bind to the unique promoter sequence to initiate the transcription of a particular gene. With different neurons expressing different subsets of proteins, we inspected proteins specific to neuron types and obtained their promoter sequence. For instance, we chose the promoter region of the *tbh-1* gene to mark octopaminergic neurons because the *tbh-1* enzyme is exclusively responsible for the synthesis of octopamine from tyramine.[25,26] Accordingly, the promoters of tyraminergic, serotonergic (5-HT), octopaminergic, cholinergic, glutaminergic, dopaminergic, and GABAergic neurons were selected because they release major neurotransmitters and embody most of the 302 neurons in *C. elegans*.[27]

Next, the wildtype N2 genome was used to amplify promoter sequences using the forward and reverse primer sequences recorded in Table 1. These primer sequences include flanking sequences, which served as enzyme-cutting sites during the digestion and ligation (Table 1). Once the ten promoters were cloned, we confirmed the promoter band size by running gel electrophoresis (data not shown, whole construct confirmations are shown in Figures 2 and 3).

3.2 Color coding neuron types via fluorescent proteins and subcellular tags

In addition to neuron-specific promoters, the full plasmid construct included fluorescent and subcellular markers to differentiate neurons in live animals. Three fluorescent proteins, specifically mScarlet, enhanced green fluorescent protein (EGFP), and mTagBFP proteins, with unique excitation and red, green, and blue emission spectra, respectively, were chosen to color code the neurons. The mScarlet, EGFP, and mTagBFP fluorescent proteins have high brightness and non-overlapping emission peaks, allowing for clear separation of neuron types. Although more fluorescent colors could have been used to label cell identities, every additional probe increases the likelihood of overlapping emission spectra and may misrepresent neuron types during the image acquisition and analysis. Thus, whole neuron imaging is limited by the number of fluorescent reporters with minimal spectral cross-talk.

To overcome this challenge, we instead used subcellular markers that better increased the fluorescent labeling capacity. Beyond the three fluorescent colors that each tag three neurons, adding a single subcellular marker, such as for mitochondria, can multiply possible ways to characterize neurons by four times as before. The cytosol expression of mScarlet can be combined with red, green, blue, or negative mitochondria expression, creating four possible color combinations in total. In this work, we inserted FIB-1 and EMR-1 subcellular markers to *gcy-13*, *tph-1*, and *tbh-1* constructs to increase the likelihood of distinguishing neurons by color, location, and morphology. FIB-1 is a nucleolar marker that shows a bright, fluorescing dot, and EMR-1 is a nuclear membrane marker that makes a circular closure around the nucleus (refer to Table 2 for the specific color of these subcellular markers).[28,29]

Altogether, the fluorescent proteins and their color combinations were carefully assigned to neuron-specific promoters based on the neuron locations to visualize their identities as shown in Table 2. Cholinergic and glutaminergic neurons are most prevalent and in close proximity to each other in *C. elegans*, and thus, we separated them by labeling them with mScarlet and mTagBFP. GABAergic neurons are also well-numbered in the animal, but since they are largely nearer to the cholinergic neurons than glutaminergic neurons, we tagged them using mTagBFP. Similar approaches were taken to decide on fluorescent proteins for each neuron type. Subcellular markers provide an additional layer of separation for tyraminergic, serotonergic, and octopaminergic neurons as they are few in number and adjacent to multiple types of neurons.

Despite segregating cell identities by color, some regions have clusters of the same type of neurons that limit determining a neuron's specific function and location at the single-cell resolution. All 302 neurons in *C. elegans* have individual names that inform their respective roles beyond the neurotransmitters they release. For instance, PHBL and PHBR are a pair of chemosensory, cholinergic neurons located on the right and left of the worm's lumbar ganglia. Since PHB is near many other cholinergic neurons, having an additional marker that differentiates PHB would enable the specific identification of the cell and its functions. Moreover, characterizing PHB increases the probability of determining the identity of the neighboring neurons. Therefore, we made *flp-7* and *flp-20* neuropeptide constructs to help us distinguish neurons of the same neurotransmitter type and understand their discrete function (Table 2). The *flp-7* neuropeptides are critical for serotonin-dependent fat metabolism and are expressed in a myriad of neuron types, including GABAergic, glutaminergic, and cholinergic neurons, and especially eminent in ALA, AVG, PHB, RIC, PVW, and SAA neurons.[30–32] The *flp-20* neuropeptides are expressed in common touch receptor neurons (TRNs), including ALM, PVM, PLM, ALM, as well as LUA and PVR.[33,34] We coexpressed *flp-7* and *flp-20* neuropeptides to mark specific sets of neurons to increase the spatial resolution in distinguishing individual cells within a population of identically color-coded cholinergic and glutaminergic neurons.

3.3 Validating the successful generation of constructs for deterministic barcoding

Once all ten plasmid constructs were created (Table 2) via a series of digestion and ligation based on the proposed strategy (Figure 1A), we performed digestion mapping and sequencing to confirm the identity of the correct plasmids. The *unc-17*, *eat-4*, *dat-1*, *unc-25*, and *rgef-1* plasmids were not tested because they were formerly confirmed and provided by the lab. First, we simulated the digestion mapping pattern using the SnapGene software (Figure 2A). HincII, EcoRI, and NdeI proteins were selected as digestive enzymes because they had multiple plasmid-cutting sites and produced different band sizes (Figure 2A). After digesting the five newly-made constructs with their respective enzymes, we compared the gel's band patterns to the simulated gel and validated its correct design (Figure 2). Furthermore, both Sanger sequencing and whole genome sequencing affirmed the proper generation of plasmids, revealing clear chromatogram peaks and high-fidelity alignment of the sequence with minimal mutations (Figure 3).

3.4 Imaging transgenic animals to record the expression pattern of flp-20

After successfully cloning five additional plasmids, we first tested and confirmed the expression pattern of the *flp-20* construct *in vivo*. We microinjected the *flp-20*-mScarlet construct into the young hermaphrodite adults and visualized the F3 generation to validate their live expression in *C. elegans*. Young adults with few eggs were the best candidates for injection as they could most efficiently pass down extrachromosomal plasmids to their offspring. To establish a stable line of transgenic animals, we maintained the microinjected worms and waited until the F3 generation for imaging. When the F3 worms became adults, we picked and mounted them on agarose to minimize movement during the live-animal imaging.

As expected, not all neurons were expressed in every worm because we did not integrate the extrachromosomal array into the worm's genome, resulting in mosaicism. Nonetheless, the *flp-20* constructs were expressed in various neurons within one worm, including LUA, PLM, PVM, ALM, and ASE (Figure 4). In other worms, these neurons and other neurons like AVM could be found (Figure 4). The neuron processes were noticeably visible because of the cytoplasm labeling, which helped to confirm their identities. The two pairings of two neurons near the rear tail region of the worm suggested that these cell bodies were PLM and LUA left and right neurons (Figure 4). We also marked ALM neurons for their unique morphology extending from the head region to the midway within the anterior body (Figure 4). Similarly, by comparing the extent of neuron processes and the shapes to the known data, we identified and labeled further neurons. Overall, evaluating *flp-20* expression supported the potency of the deterministic color-coding mechanism. We then extended this approach to the rest of the plasmids by co-injecting all ten constructs into the worms.

3.5 Whole-brain neural circuitry and identity

The co-injection of all ten plasmids allowed us to produce a comprehensive map of the worm's brain connectome. Although live-animal imaging unquestionably provides us with quality images where neurons are easily segregatable, we decided to fully immobilize the worms to obtain the best co-injection data for analyzing and marking neuron identities. We thereby acquired higher-resolution images through fixation in addition to gelation, which eliminated the chance of worms altering their positions. All three emission channels exhibited fluorescence as expected and had their highest signals near the head due to the dense clusters of neurons (Figure

5A). Therefore, we focused our cell identification on the anterior to mid-body region, where the most diverse groups of neurons were located, instead of extensively representing all 302 neurons to confirm the proof-of-concept of this *in vivo* technique. The red fluorescence, largely from cholinergic neurons, was confirmed to be strongest near the center to the ventral side, while the blue fluorescence, mainly from GABAergic and glutamatergic neurons, was strongest from the center to the dorsal side (Figure 5A). The ventral nerve cord line was also present, displaying an extended signal of neuron processes and wide clumps of cell bodies (Figure 5A). Still, not all neurons were shown, even within the head, due to not only mosaicism but also unidentifiable key neurotransmitters in some neurons, the worm's orientation, and software light adjustment that caused some signals to be missed or lost in the image (Figure 5).

After validating the general expression pattern of fluorescent colors, we analyzed and mapped their unique identities using colors, morphologies, and relative locations in both projected and volume images of *C. elegans*. The directionality, whether it's unipolar or bipolar, as well as the shape of neuron processes like bending and circularity helped especially when distinguishing neurons of the same color (Figure 5B I). The subcellular structures also helped to identify tyraminergic, serotonergic (5-HT), and octopaminergic neurons in a pool of largely cholinergic, glutaminergic, and GABAergic neurons (Figure 5B II & III). Altogether, the neurons were analyzed using the volume and z stacks images to observe their relative locations and the extent of neuron processes in a three-dimensional setting.

Generally, characterizing neurons was easier and more reliable with a lower density of neurons at the very anterior. Some of these neurons were Il, I2, I3, MI, MCL, M3L, M4, and NSM (Figure 5C I & II). Il is a pharyngeal interneuron that releases acetylcholine at a spatial distance from other neurons. Since we marked cholinergic neurons with mScarlet, the single existence of a red-colored neuron was easily visible on the connectome (Figure 5C I). Consequently, the neighboring sets of neurons like I2, MCL, and NSM were identified using their relative location to Il and each other and color that indicated their neurotransmitter type (Figure 5C I). We applied a similar approach to reproduce the brain map of deterministically barcoded neurons in other *C. elegans* (Figure 5C II). As tagged with both mScarlet and mTagBFP, the CEPD and CEPV neurons were eminently purple (Figure 5C II). Moreover, by visualizing bipolar processes and a process that wraps around the worm body, ADE and DD

neurons could be identified, respectively (Figure 5C II). Through this analytical approach, we labeled neuron types and reproduced them across different *C. elegans*.

Taken together, we report the successful generation of transgenic worms that allows *in vivo* imaging of the whole brain and deterministic identification of neurons through multicolor fluorescent coding.

4. Discussion

Understanding the brain at a molecular, cellular, structural, and functional signaling level is crucially tied to resolving neural circuitry and synaptic communication. Without knowing the cellular compositions of the brain, it is challenging to understand the molecular networks that determine the function, development, and communication of the nervous system and their implications on disease progressions. In this work, we successfully generate cell-specific constructs to drive fluorescent expression of different neuron types in *C. elegans*, enabling live, single-cell identifications.

Amidst the brain's complexity, the deterministic approach to cell type labeling can allow deeper insight into the identity of each neuron *in vivo* and their anticipated roles within the brain. In the past, Brainbow and similar systems were developed as an *in-vivo* transgenic cell-labeling strategy that stochastically combined and expressed fluorescent proteins.[21,36] However, while useful for monitoring live tissues and animals at different time points, these preclude understanding cell types due to the random nature of Cre-loxP recombination for multicolor expression. Even in the work done by Yemini *et al.*, where they developed a multicolor landmark strain to label the whole nervous system of *C. elegans*, this technique was limited to labeling the nuclei due to the type of sensor they implemented.[6] Therefore, the labeling did not show the morphologies that are critical for identifying neurons. With the finite numbers of spectrally distinct fluorophores and high-density clusters of neuron bodies, whole-brain mapping in live animals remains challenging.

To meet this unmet need, we take advantage of neuron-specific promoters to construct transgenes that label each neuron and its processes with unique red, green, and blue fluorescent proteins. Importantly, we enable deterministic neuron identification using neurons' distinct morphologies and subcellular markers. With this multicolor palette approach, we thereby embrace an added layer of information about neuronal identities. Ultimately, we parsed out complex cellular structures, generating a multicolored fluorescent mapping of different neurons in *C. elegans* for live, whole-brain imaging and individual cell barcoding.

To further develop this technique, we anticipate rounds of optimizing the color-coding schemes. First, we plan to modify fluorescent proteins of *gcy-13*-FIB-1-mTagBFP, *tph-1*-EMR-1-mScarlet, and *tbh-1*-FIB-1-mTagBFP constructs. Currently, many of these

subcellular markers are co-expressed with fluorophores of the same color in the cytosol, making it difficult to recognize these structures. Therefore, by changing them into the opposite color, we would produce a sharp contrast between the cytosol and the subcellular tags that are easier to identify cells. Furthermore, the use of additional neuropeptidergic constructs with subcellular tags can help better mark neurons at a single-cell resolution. Consequently, it may be useful to explore diverse sets of fluorescent proteins to control the emission band, increase brightness, and reduce noise like unspecific expression of fluorophores. When the optimized plasmids are prepared for injection, we intend to integrate the plasmids into the stable germline and eliminate mosaicism. Beyond imaging in live *C. elegans*, this approach can be translated into other imaging setups, including samples like cultured cells and mice brain tissue. Similarly, although the process will no longer be in live animals, expansion microscopy can be applied to conduct high resolution, fixed animal imaging, and thus, mapping brain connectome with even greater fidelity.[37,38]

Ultimately, through this method of deterministic barcoding of neuron identities, we aim to aid cell typing throughout development and advance the understanding of both the connectome and the changes in neural circuitry over time. As previously mentioned, we expect this technique to be much more practical and compatible with other imaging setups, thus, opening doors to new experimental findings under normal and diseased physiological conditions.

5. Conclusion

The multicolor, deterministic barcoding to label neuron identities in vivo is a potent tool to image live animals over time. Through monitoring the changes in neuronal activity and connectivity, how the brain connectome alters during the development and even disease pathology can be better understood. This genetically targeting approach is particularly compatible when investigating neuron identities as their unique morphology of neuronal processes can be visualized. Beyond that, the deterministic toolbox enables studying cellular compositions in diverse cell populations like tissues and multicellular cultures. With neuron-specific promoters tagged with fluorescent proteins and subcellular markers helping to differentiate same-colored neuron clusters, accurately discriminating and identifying neurons in C. elegans became more robust. This technique can also be applied to study the cell compositions of other model organisms like mice, Drosophila, and zebrafish. We believe this advanced form of multicolor cell typing will be a valuable tool in investigating future genetics and development studies.

6. References

1. Mental Illness. National Institute of Mental Health (NIMH). Accessed April 29, 2023. https://www.nimh.nih.gov/health/statistics/mental-illness

2. Nearly 1 in 6 of world's population suffer from neurological disorders – UN report | UN News. Published February 27, 2007. Accessed April 29, 2023. https://news.un.org/en/story/2007/02/210312

3. Overview | Brain Initiative. Accessed April 29, 2023. https://braininitiative.nih.gov/about/overview

4. Herculano-Houzel S. The Human Brain in Numbers: A Linearly Scaled-up Primate Brain. *Front Hum Neurosci*. 2009;3:31. doi:10.3389/neuro.09.031.2009

5. Nguyen T. Total Number of Synapses in the Adult Human Neocortex. *Undergrad J Math Model One Two*. 2010;3(1). doi:http://dx.doi.org/10.5038/2326-3652.3.1.26

6. Yemini E, Lin A, Nejatbakhsh A, et al. NeuroPAL: A Multicolor Atlas for Whole-Brain Neuronal Identification in C. elegans. *Cell*. 2021;184(1):272-288.e11. doi:10.1016/j.cell.2020.12.012

7. Leergaard T, Hilgetag C, Sporns O. Mapping the Connectome: Multi-Level Analysis of Brain Connectivity. *Front Neuroinformatics*. 2012;6. Accessed April 29, 2023. https://www.frontiersin.org/articles/10.3389/fninf.2012.00014

8. Wong MY, Kaeser PS. Active Zone☆. In: *Reference Module in Biomedical Sciences*. Elsevier; 2014. doi:10.1016/B978-0-12-801238-3.04486-X

9. Valenzuela CF, Puglia MP, Zucca S. Focus On: Neurotransmitter Systems. *Alcohol Res Health*. 2011;34(1):106-120.

10. Sheffler ZM, Reddy V, Pillarisetty LS. Physiology, Neurotransmitters. In: *StatPearls*. StatPearls Publishing; 2023. Accessed April 30, 2023. http://www.ncbi.nlm.nih.gov/books/NBK539894/

11. Russo AF. Overview of neuropeptides: awakening the senses? *Headache*. 2017;57(Suppl 2):37-46. doi:10.1111/head.13084

12. Burbach JPH. What Are Neuropeptides? In: Merighi A, ed. *Neuropeptides: Methods and Protocols*. Methods in Molecular Biology. Humana Press; 2011:1-36. doi:10.1007/978-1-61779-310-3_1

13. Working with Worms: Caenorhabditis elegans as a Model Organism - Meneely - 2019 - Current Protocols Essential Laboratory Techniques - Wiley Online Library. Accessed April 30, 2023. https://currentprotocols.onlinelibrary.wiley.com/doi/full/10.1002/cpet.35

14. Hobert O. Neurogenesis in the nematode Caenorhabditis elegans. In: *WormBook: The Online Review of C. Elegans Biology [Internet]*. WormBook; 2018. Accessed April 30, 2023. https://www.ncbi.nlm.nih.gov/books/NBK116086/

15. Maity B, Sheff D, Fisher RA. Chapter 5 - Immunostaining: Detection of Signaling Protein Location in Tissues, Cells and Subcellular Compartments. In: Conn PM, ed. *Methods in Cell Biology*. Vol 113. Laboratory Methods in Cell Biology. Academic Press; 2013:81-105. doi:10.1016/B978-0-12-407239-8.00005-7

16. Duraiyan J, Govindarajan R, Kaliyappan K, Palanisamy M. Applications of immunohistochemistry. *J Pharm Bioallied Sci*. 2012;4(Suppl 2):S307-S309. doi:10.4103/0975-7406.100281

17. Young AP, Jackson DJ, Wyeth RC. A technical review and guide to RNA fluorescence in situ hybridization. *PeerJ*. 2020;8:e8806. doi:10.7717/peerj.8806

18. Zenklusen D, Singer RH. Chapter 26 - Analyzing mRNA Expression Using Single mRNA Resolution Fluorescent In Situ Hybridization. In: *Methods in Enzymology*. Vol 470. Guide to Yeast Genetics: Functional Genomics, Proteomics, and Other Systems Analysis. Academic Press; 2010:641-659. doi:10.1016/S0076-6879(10)70026-4

19. Tang X, Huang Y, Lei J, Luo H, Zhu X. The single-cell sequencing: new developments and medical applications. *Cell Biosci*. 2019;9(1):53. doi:10.1186/s13578-019-0314-y

20. Kukurba KR, Montgomery SB. RNA Sequencing and Analysis. *Cold Spring Harb Protoc*. 2015;2015(11):951-969. doi:10.1101/pdb.top084970

21. Weissman TA, Pan YA. Brainbow: New Resources and Emerging Biological Applications for Multicolor Genetic Labeling and Analysis. *Genetics*. 2015;199(2):293-306. doi:10.1534/genetics.114.172510

22. Richier B, Salecker I. Versatile genetic paintbrushes: Brainbow technologies. *WIREs Dev Biol*. 2015;4(2):161-180. doi:10.1002/wdev.166

23. Li Y, Walker LA, Zhao Y, et al. Bitbow Enables Highly Efficient Neuronal Lineage Tracing and Morphology Reconstruction in Single Drosophila Brains. *Front Neural Circuits*. 2021;15. Accessed April 30, 2023. https://www.frontiersin.org/articles/10.3389/fncir.2021.732183

24. Promoter. Genome.gov. Published September 14, 2022. Accessed April 30, 2023. https://www.genome.gov/genetics-glossary/Promoter

25. Özbey NP, Imanikia S, Krueger C, et al. Tyramine Acts Downstream of Neuronal XBP-1s to Coordinate Inter-tissue UPRER Activation and Behavior in C. elegans. *Dev Cell*. 2020;55(6):754-770.e6. doi:10.1016/j.devcel.2020.10.024

26. Yu J, Vogt MC, Fox BW, et al. Parallel pathways for serotonin biosynthesis and metabolism in C. elegans. *Nat Chem Biol*. 2023;19(2):141-150. doi:10.1038/s41589-022-01148-7

27. Gendrel M, Atlas EG, Hobert O. A cellular and regulatory map of the GABAergic nervous system of C. elegans. Shen K, ed. *eLife*. 2016;5:e17686. doi:10.7554/eLife.17686

28. The Nucleolus of Caenorhabditis elegans - PMC. Accessed April 30, 2023. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3345250/

29. Age-related changes of nuclear architecture in Caenorhabditis elegans. doi:10.1073/pnas.0506955102

30. Palamiuc L, Noble T, Witham E, Ratanpal H, Vaughan M, Srinivasan S. A tachykinin-like neuroendocrine signalling axis couples central serotonin action and nutrient sensing with peripheral lipid metabolism. *Nat Commun*. 2017;8(1):14237. doi:10.1038/ncomms14237

31. Li C, Kim K. Neuropeptides. In: *WormBook: The Online Review of C. Elegans Biology [Internet]*. WormBook; 2018. Accessed April 30, 2023. https://www.ncbi.nlm.nih.gov/books/NBK116087/

32. NeuroPAL ·. Accessed April 30, 2023. https://www.hobertlab.org/neuropal/

33. An Afferent Neuropeptide System Transmits Mechanosensory Signals Triggering Sensitization and Arousal in C. elegans - ScienceDirect. Accessed April 30, 2023. https://www.sciencedirect.com/science/article/pii/S0896627318306767

34. Kang C, Avery L. The FMRFamide Neuropeptide FLP-20 Acts as a Systemic Signal for Starvation Responses in Caenorhabditis elegans. *Mol Cells*. 2021;44(7):529-537. doi:10.14348/molcells.2021.0051

35. WormAtlas, Altun, Z.F., Herndon, L.A., Wolkow, C.A., Crocker, C., Lints, R. and Hall, D.H. (ed.s) 2002-2023. http://www.wormatlas.org

36. Cai D, Cohen KB, Luo T, Lichtman JW, Sanes JR. Improved tools for the Brainbow toolbox. *Nat Methods*. 2013;10(6):540-547. doi:10.1038/nmeth.2450

37. Chen F, Tillberg PW, Boyden ES. Expansion microscopy. *Science*. 2015;347(6221):543-548. doi:10.1126/science.1260088

38. Wassie AT, Zhao Y, Boyden ES. Expansion microscopy: principles and uses in biological research. *Nat Methods*. 2019;16(1):33-41. doi:10.1038/s41592-018-0219-4

39. Searching for Neuronal Left/Right Asymmetry: Genomewide Analysis of Nematode Receptor-Type Guanylyl Cyclases | Genetics | Oxford Academic. Accessed April 30, 2023. https://academic.oup.com/genetics/article/173/1/131/6061598

40. Food and metabolic signalling defects in a Caenorhabditis elegans serotonin-synthesis mutant | Nature. Accessed April 30, 2023. https://www.nature.com/articles/35000609

41. Suo S, Kimura Y, Van Tol HHM. Starvation Induces cAMP Response Element-Binding Protein-Dependent Gene Expression through Octopamine–Gq Signaling in Caenorhabditis elegans. *J Neurosci*. 2006;26(40):10082-10090. doi:10.1523/JNEUROSCI.0819-06.2006

42. Hobert O. PCR fusion-based approach to create reporter gene constructs for expression

analysis in transgenic C. elegans. *BioTechniques*. 2002;32(4):728-730. doi:10.2144/02324bm01

43. Evans TC. Transformation and microinjection. In: *WormBook: The Online Review of C. Elegans Biology [Internet]*. WormBook; 2006. Accessed April 30, 2023. https://www.ncbi.nlm.nih.gov/books/NBK19648/

44. Mello C, Fire A. DNA transformation. *Methods Cell Biol*. 1995;48:451-482.

45. Stiernagle T. Maintenance of C. elegans. In: *WormBook: The Online Review of C. Elegans Biology [Internet]*. WormBook; 2006. Accessed April 30, 2023. https://www.ncbi.nlm.nih.gov/books/NBK19649/

46. Yu CC (Jay), Barry NC, Wassie AT, et al. Expansion microscopy of C. elegans. Dernburg AF, Ron D, eds. *eLife*. 2020;9:e46249. doi:10.7554/eLife.46249

47. NIH Image to ImageJ: 25 years of image analysis | Nature Methods. Accessed April 30, 2023. https://www.nature.com/articles/nmeth.2089

Appendix A. Figures

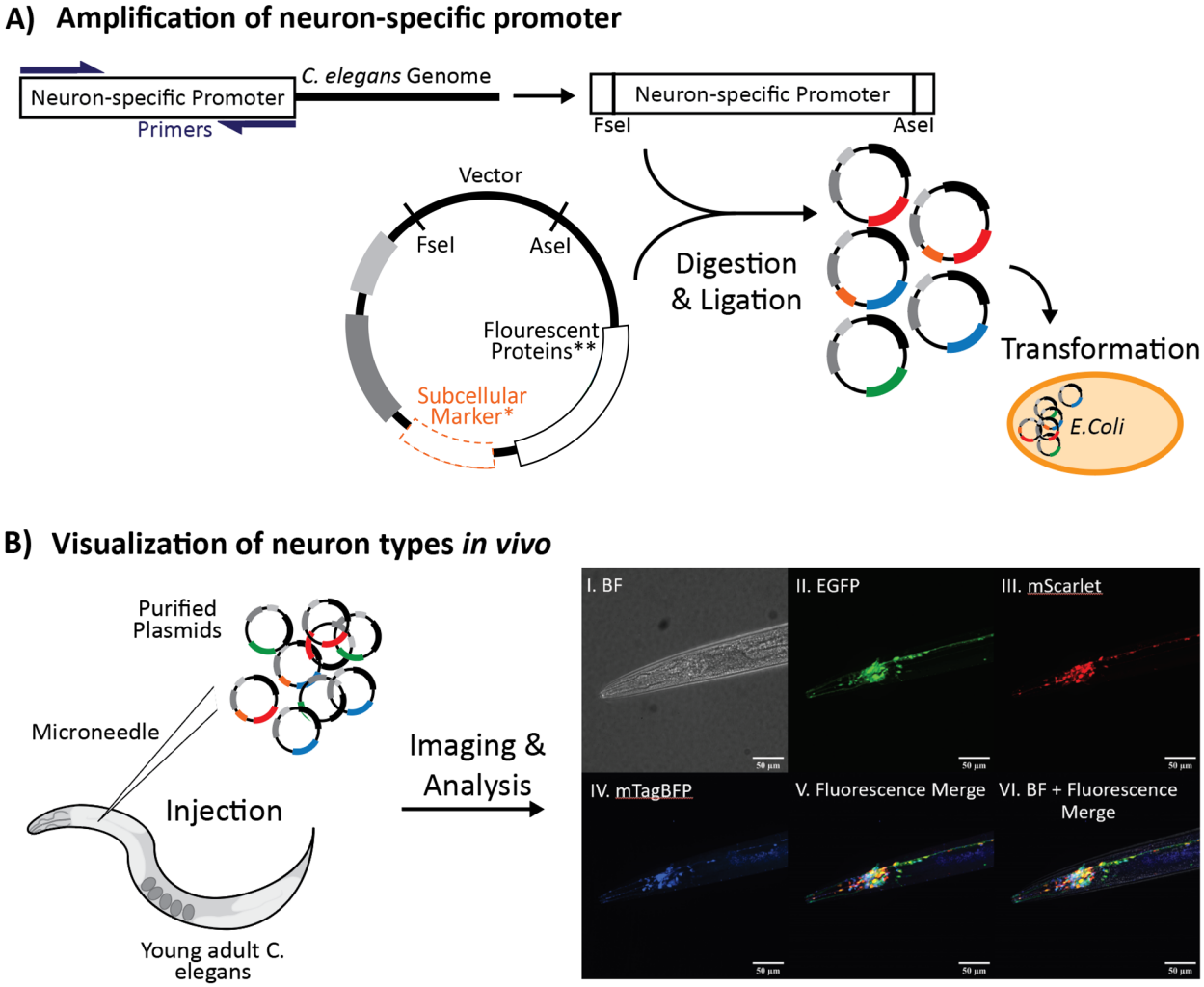
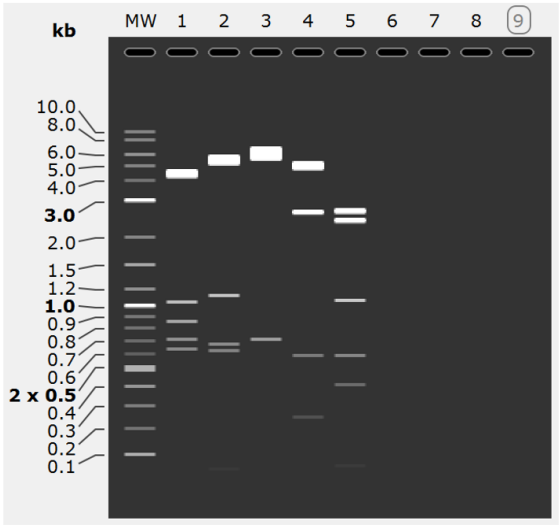


Figure 1. Overview schematic of deterministic color coding. A) Nine neuron-specific promoters and one pan-neuronal promoter were cloned from the wild-type (strain N2) genome to generate color-coded constructs. They were then amplified using *E. coli* cells. *Not all constructs contained subcellular markers. **Fluorescent proteins were either red, green, or blue. B) Purified plasmids were microinjected into the gonad of *C. elegans* to create a transgenic animal. Subsequent imaging was performed under the confocal microscope to confirm the cell identities. The worms were imaged in I) brightfield (BF, grayscale), II) emission channel (Em) of 582 nm (mScarlet, red, driven by *tph-1*, *flp-20*, *unc-17*, and *dat-1* promoters), III) Em of 525 nm (GFP, green, driven by pan-neuronal markers), and IV) Em of 455 nm (mTagBFP, blue, driven by *gcy-13*, *tbh-1*, *flp-7*, *eat-4*, *dat-1*, and *unc-25* promoters). The merged images (V and VI) combined three fluorescent color images into one for comparing the neurons' relative locations.

A)



Lane	Promoter Name	Digestive Enzyme	Expected band size (bp)
1	gcy-13	HincII	4347, 1050, 860, 708, 635
2	tph-1	EcoRI	5376, 1123, 677, 621, 48
3	tbh-1	HincII	5964, 708
4	flp-7	EcoRI	4902, 2657, 591, 247
5	flp-20	NdeI	2712, 2427, 1066, 597, 413, 61

B)

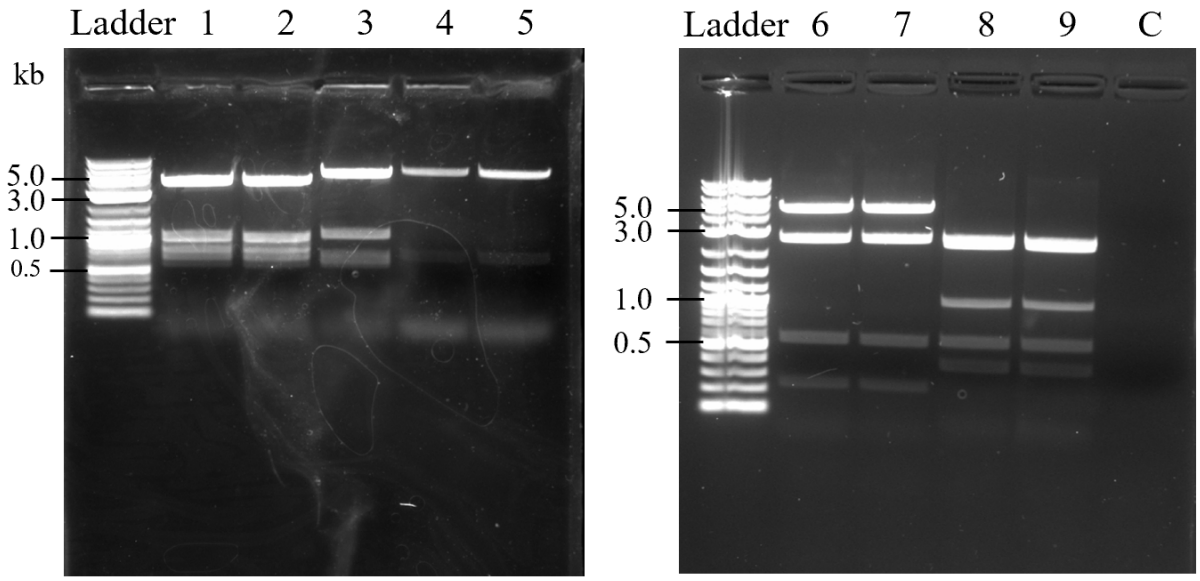
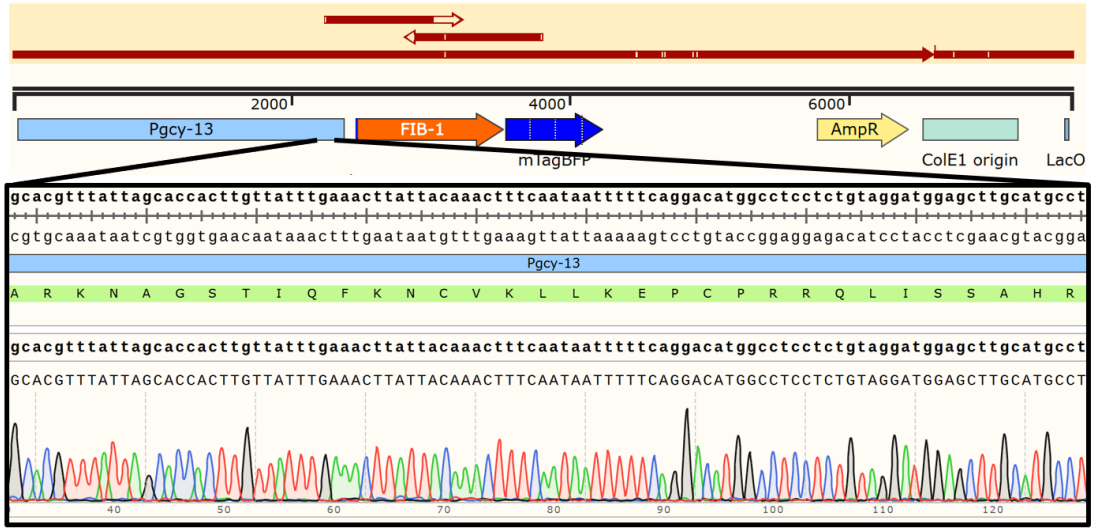
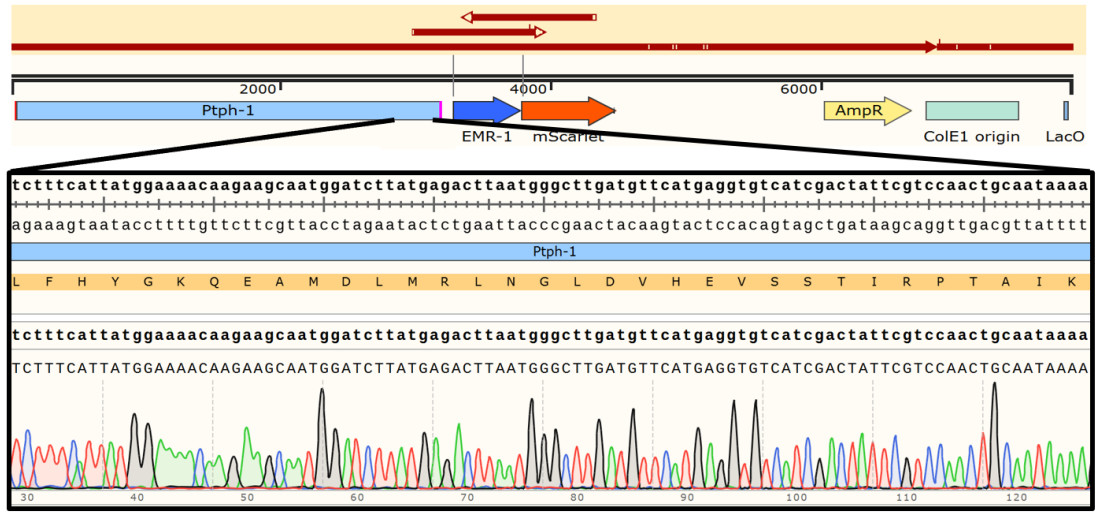


Figure 2. Digestion mapping indicates anticipated length of each cloned construct. 1% agarose gel was used to visualize the digestion pattern of constructed plasmids for *gcy-13*, *tph-1*, *tbh-1*, *flp-7*, and *flp-20* using 1 kb Plus DNA Ladder. A) Expected bands length from the given restriction digest enzymes for each construct are shown. Generated by SnapGene. B) Image of the digested plasmids. Lanes 1-2 corresponds to plasmids containing *gcy-13*, lane 3 to *tph-1*, lanes 4-5 to *tbh-1*, lanes 6-7 to *flp-7*, and lanes 8-9 to *flp-20*. The final lane (C) is the control of 50 µL water and buffer.

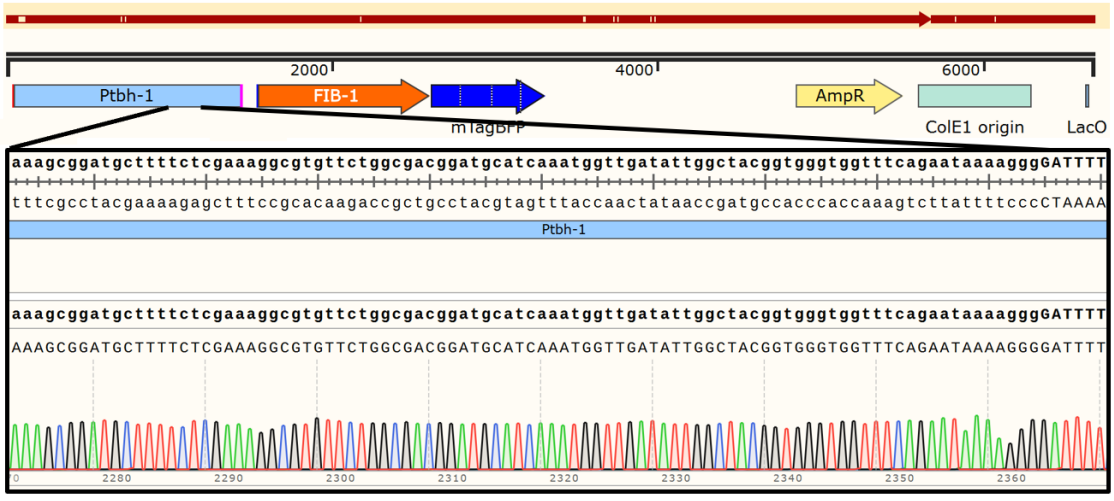
A) *Pgcy-13*-FIB-1-mTagBFP



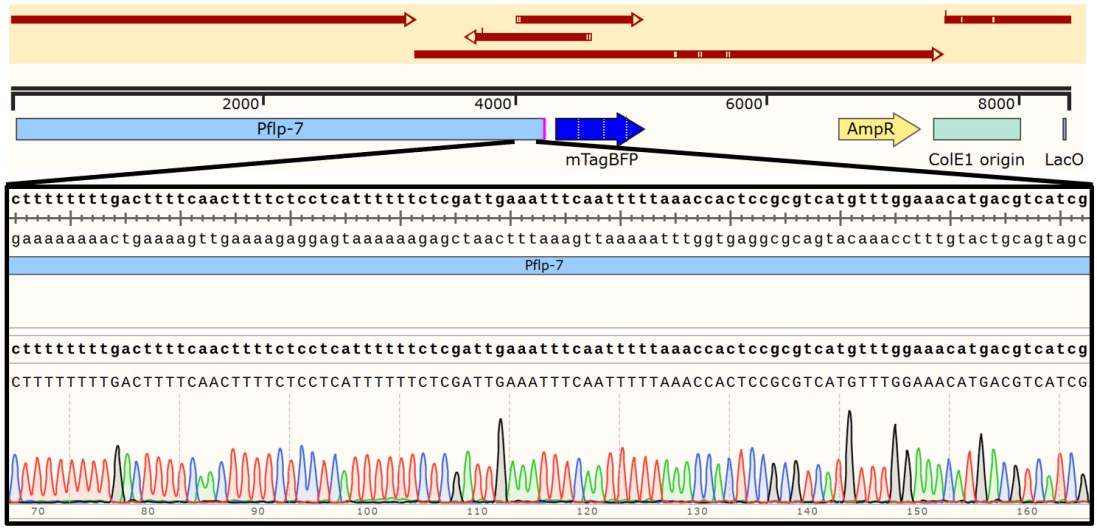
B) *Ptph-1*-EMR-1-mScarlet



C) *Ptbh-1*-FIB-1-mTagBFP



D) *Pflp-7*-mTagBFP



E) *Pflp-20*-mScarlet

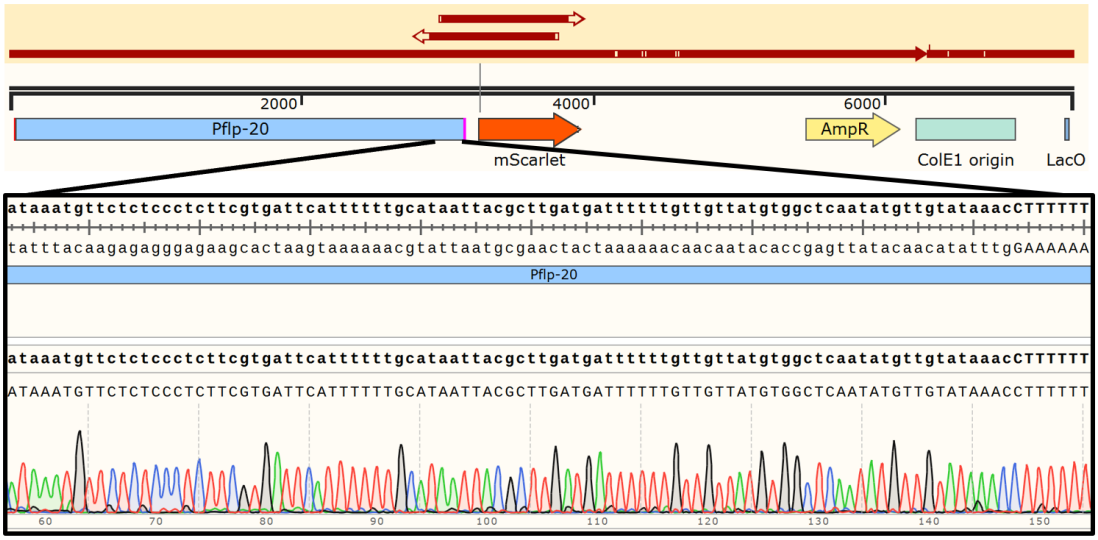


Figure 3. SnapGene sequence alignment reveals successful cloning of neuron-specific constructs. Both the Sanger sequencing and whole genome sequencing results are shown (except for *tbh-1* in which only whole genome sequencing is shown). For each plasmids (A through E), chromatogram exhibits clear peaks.

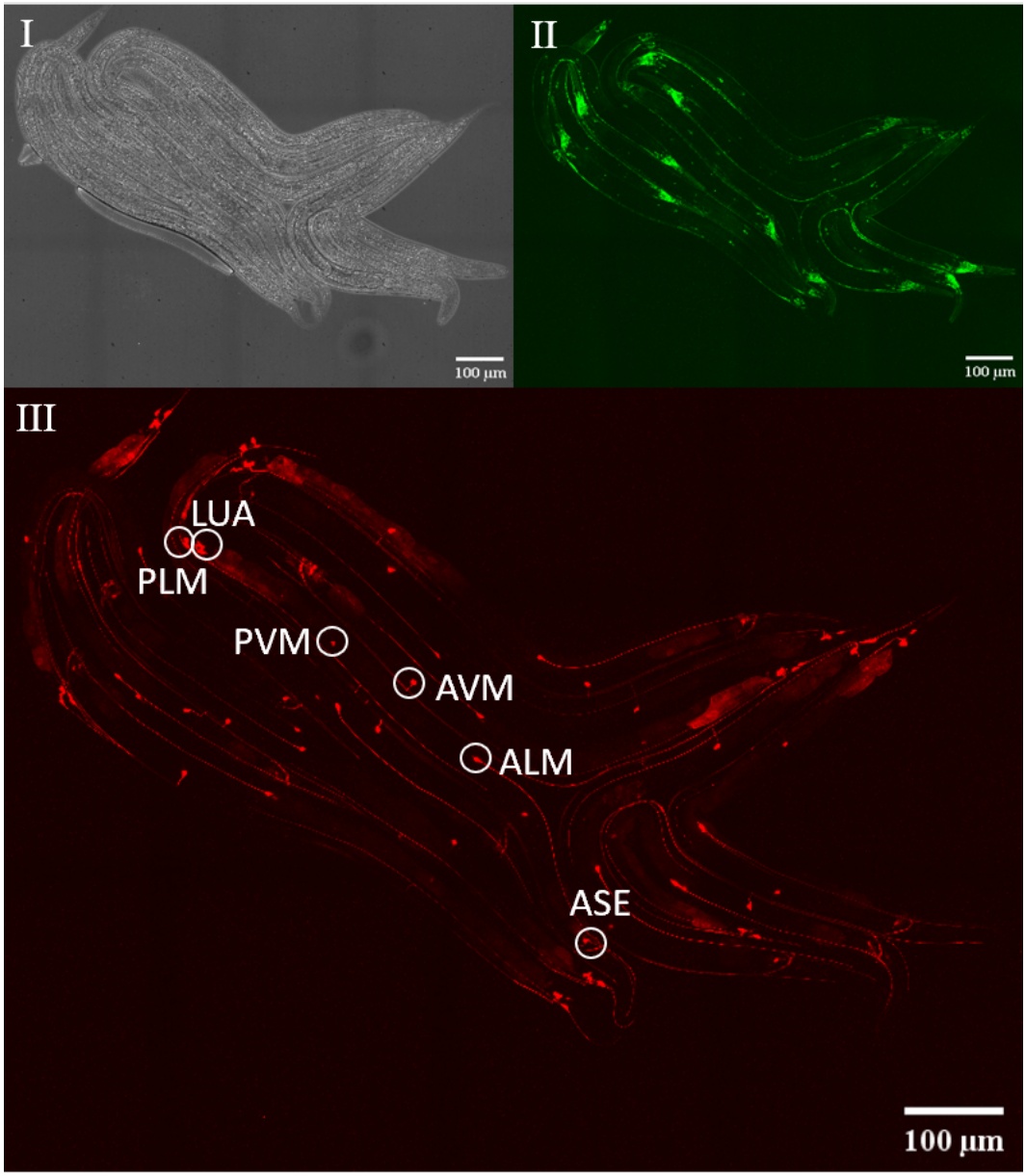
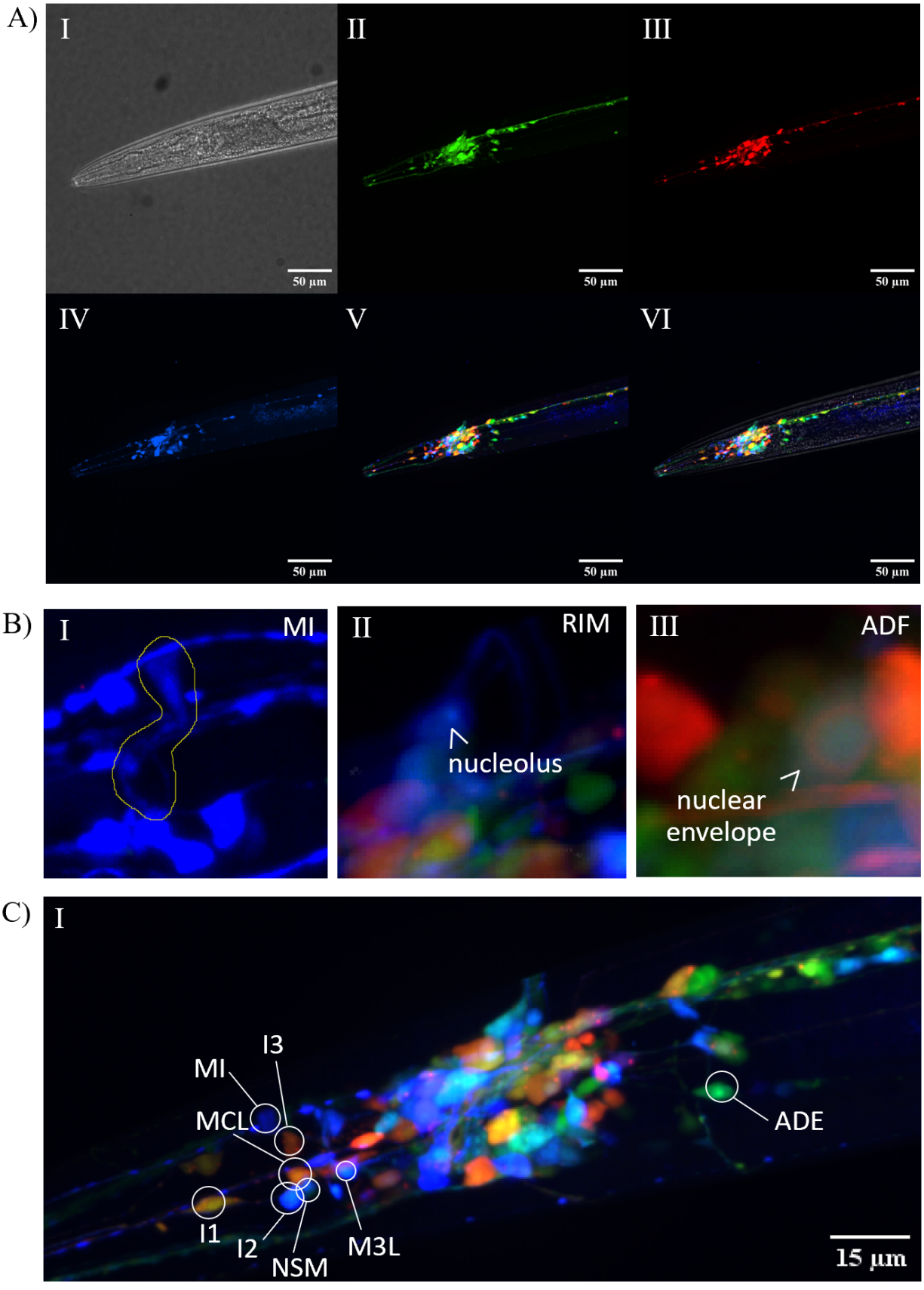


Figure 4. Expression pattern of flp-20-mScarlet construct incorporated into live *C. elegans.* The I) brightfield (BF, grayscale), II) GFP (green), and III) RFP (red) images of the dense clusters of worms are shown. GFP is used as the pan neuronal marker and RFP is expressed in mechanosensory neurons that express *flp-20* in their bodies, including LUA, PLM, PVM, AVM, ALM, and ASE. The scale bar indicates 100 µm, and neuron type labels were based on the mapping data from WormAtlas.[35]


A)
I
II
III
IV
V
VI
50 µm
B)
I
MI
II
RIM
nucleolus
III
ADF
nuclear
envelope
C)
I
I3
MI
MCL
I1
I2
NSM
M3L
ADE
15 µm

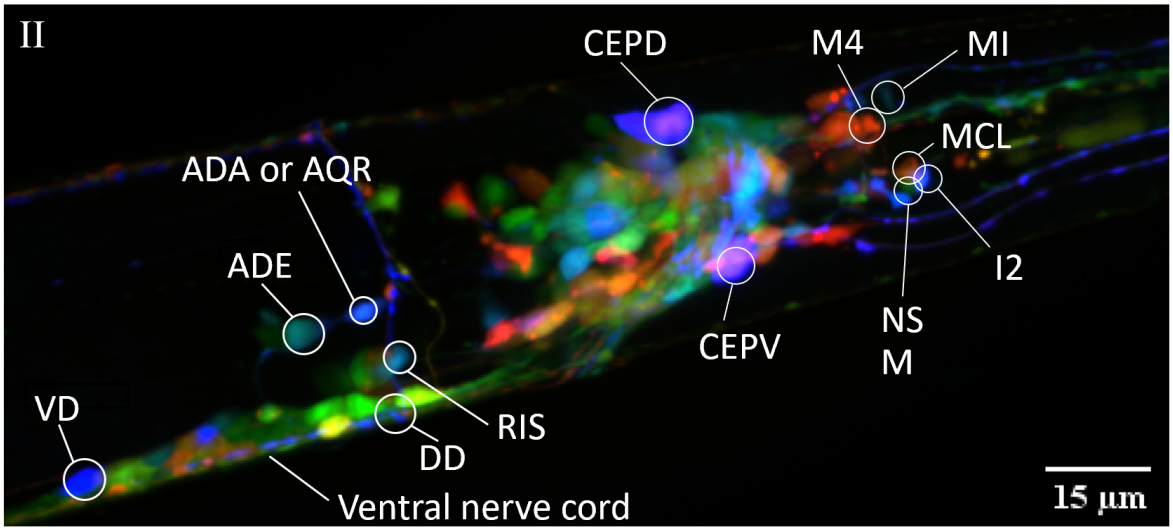


Figure 5. Confocal microscopic images present the unique expression patterns of neuron-specific constructs. Ten plasmids (*gcy-13*, *tph-1*, *tbh-1*, *flp-7*, *flp-20*, *unc-17*, *eat-4*, *dat-1*, *unc-25*, and *rgef-1*) were microinjected into the worms, and their expression patterns are shown above. A) The worms were visualized per the previous imaging setting with each panel representing images under the channels for I) BF, II) GFP, III) mScarlet, IV) mTagBFP, V) Merged image of II-IV, and VI) Merged image of I-IV. Scale bar, 50 µm. B) Unique neuron morphology and nucleolus and nuclear membrane subcellular structures guided tagging neuron identities. For instance, (I) MI neuron has distinct unipolar, circular processes; (II) RIM, a tyraminergic neuron, exhibits a nucleolus marker as encoded by the *gcy-13*-FIB-1-mTagBFP construct; (III) ADF neuron that releases both serotonin and acetylcholine is fluorescently tagged around the nuclear membrane with mScarlet. C) (I) In a zoomed-in version of panel V from part A, subsets of neurons are labeled with their unique identities again using the same protocol per Figure 4. (II) Marking unique expression patterns of neuron-specific constructs is reproduced in another set of *C. elegans*. Scale bar, 15 µm.

Appendix B. Tables

Neuron Type	Promoter	Forward Primer	Reverse Primer	Promoter Length
Tyraminergic	*gcy-13*	actgggccggccGCCACAATGACAATTATCTC	tagaggcgcgccAGTCGACCTGCAGGCATGCAAGCTCCATCCTACAGAGGAGGC	2323bp
Serotonergic (5-HT)	*tph-1*	actgggccggccGATGATCATAGTCAGTCGAGCC	tagaggcgcgccCCGGTTGTCGCATCTGAAAAAGG	3129bp
Octopaminergic	*tbh-1*	actgggccggccCCGGAATCTTCCTTATGTATCTC	tagaggcgcgccGAAACAGGAGATGCTCCGAAAATC	1348bp
Neuropeptidergic	*flp-7*	ActgggccggccTTCTCGGCCAAATCTATGTATTTCG	tagaggcgcgccGAAGCGGGATCCAAGCATTTC	5152bp
Neuropeptidergic	*flp-20*	actgggccggccGAAACATTGGTCGGGAGATG	tagaggcgcgccTTGCTCACTAAAGTTTCTGAAAAAAAG	3067bp
*Cholinergic	*unc-17*	actgggccggccTACACCAATCATTTCTCCCCTCCGAAA	tagaggcgcgccCTGAAAATTAAATATTTTAGTGTAAAACTTTTGGGGGCA	3180bp
*Glutaminergic	*eat-4*	actgggccggccTTGAAGTAGCTCACTGATGGATCGTAATTT	tagaggcgcgccGGTTTCTGAAAATGATGATGATGATGATGGAGT	5624bp
*Dopaminergic	*dat-1*	actgggccggccGCCTATTCCAGTATGACC	tagaggcgcgccCTGAAAACACATGAATCTAGTATAG	400bp
*GABAergic	*unc-25*	actgggccggccGAAATATCGATTTTTTCCGAGATATCTAAAATTGAAAGAAAT	tagaggcgcgccTTTTGGCGGTGAACTGAGCTTTTC	1277bp

*All neurons	*rgef-1*	Obtained by cutting plasmid pJH3971 (a gift from Mei Zhen lab at University of Toronto; Addgene #177736) with NheI and SpeI, and ligate with EGFP fragment using the same restriction enzyme sites	3444bp (length of rgef-1)

Table 1. Primer Sequences designed to amplify neuron-specific promoters. The sequence "actgggccggcc" and "tagaggcgcgcc" are flanking sequences, used as cutting sites for the FseI and AscI digestion enzymes. The primers are approximately 2300 bp to 5200 bp long.
* The following primer sequences are donated by the Neuro Synthetic Group, produced by Yangning Lu (Thesis mentor).

Neuron Type	Promoter	Gene Expressed
Tyraminergic	*gcy-13*	FIB-1-mTagBFP
Serotonergic (5-HT)	*tph-1*	EMR-1-mScarlet
Octopaminergic	*tbh-1*	FIB-1-mTagBFP
Neuropeptidergic	*flp-7*	mTagBFP
Neuropeptidergic	*flp-20*	mScarlet
Cholinergic	*unc-17*	mScarlet
Glutaminergic	*eat-4*	mTagBFP
Dopaminergic	*dat-1*	mScarlet, mTagBFP
GABAergic	*unc-25*	mTagBFP
All neurons	*rgef-1*	EGFP

Table 2. Full plasmid constructs expression.