A pilot study for whole proteome tagging in *C. elegans*

Matthew Eroglu and Oliver Hobert

Columbia University, Department of Biological Sciences, Howard Hughes Medical Institute, New York, NY 10027
Correspondence: me2839@columbia.edu & or38@columbia.edu

Abstract

Tagging all proteins encoded by an animal genome with a fluorescent tag would open many windows to the discovery of unexpected patterns of protein expression and localization. To scale such an approach, it would be beneficial to introduce multiple, spectrally distinct fluorophore tags in parallel. As a first step in this direction, we undertook a pilot study in the nematode *C. elegans,* in which we set out to tag 30 different genetic loci with three different fluorophores, with 3 tags being introduced at a time. By choosing essential genes, predicted based on transcriptomics to cover a range of expression levels, we explore issues relating to disrupting gene function and visibility of tagged proteins. We demonstrate that such a tagging approach is highly efficient and indeed reveals unanticipated patterns of cellular sites of expression, as well as subcellular protein localization. We hope that this pilot study will motivate attempts to scale this tagging approach to more loci and, ultimately, the whole genome.

Introduction

Gene transcript-based expression atlases have provided invaluable insights into animal biology. The next frontier lies in generating whole animal protein expression atlases with single cell resolution that capture the cellular and subcellular specificity of protein expression and localization and visualize the dynamics of these parameters during development and during changes in internal and external states. The establishment of protein expression atlases is particularly urgent in light of the generally appreciated discordance between transcript detection – often used to infer cellular functions – and protein expression, a reflection of multifarious layers of posttranscriptional gene regulation (Liu et al., 2016; Schwanhäusser et al., 2011; Vogel and Marcotte, 2012).

One means to acquire protein expression and localization on a single cell level with whole proteome coverage lies in tagging all genes in the genome with a fluorescent reporter. This approach has been pioneered in yeast (Huh et al., 2003), where chromosomally GFP-tagged strains provided the first cellular map covering ~75% of the yeast proteome, assigning localization for 70% of previously unlocalized proteins while uncovering the compartmental logic underlying protein function and interactions. Tagging proteins in multicellular organisms provides a host of new insights since the differentiation of cells into distinct types and their assembly into tissues deploys many more proteins and encompasses the establishment and dynamic control of more complex protein expression and localization patterns.

We envision the tagging of the entire proteome of a multicellular organism to be most feasible in the nematode *C. elegans*. CRISPR/Cas9 genome engineering is now routinely used to tag individual genetic loci in *C. elegans* with a fluorophore. So far, around 1,500 genes have been tagged already by individual labs (Leyhr et al., 2025). With recent procedural advances (Eroglu et al., 2023; Ghanta and Mello, 2020; Paix et al., 2015), such engineering is highly effective and rapid, given the short generation time of *C. elegans*. The transparency of the organism facilitates imaging of fluorophores in live animals. Advances in light-sheet fluorescence microscopy, such as dual-view inverted selective plane illumination microscopy (diSPIM), have enabled high-resolution, 3-dimensional live imaging of dynamic fluorophore localization patterns throughout all stages of development (Kumar et al., 2014; Wu et al., 2017). Super-resolution imaging approaches such as image scanning microscopy (ISM) (Huff, 2015) and stimulated emission depletion (STED) microscopy (Rankin et al., 2011) enable increasingly finer localization of proteins in live worms. Further resolution can also be achieved in fixed worms with expansion microscopy, without the requirement for specialized microscopy equipment (Yu et al., 2022). Extensive anatomical atlases are either already available (Altun et al., 2002) or in the process of being generated (Cook et al., 2019; Santella et al., 2022; Witvliet et al., 2021), allowing the mapping of fluorophore localization patterns onto anatomical features at electron microscopical scales (Vergara et al., 2021).

There is also ample precedent in the *C. elegans* literature that reporter tagging results in unanticipated insights. To cite two recent examples from our own lab, we found that Golgi-resident proteins displayed cell type-specific expression patterns not anticipated by transcriptomic analysis (Leyva-Díaz et al., 2025). As another example, the subcellular localization pattern of a DEG/ENaC channel in enteric neurons indicated unexpected functional insights into the protein (Bayer et al., 2025). It is also notable that the widely appreciated discordance of mRNA and protein expression levels is evident in *C. elegans* (Grün et al., 2014; Reilly et al., 2022, 2020), attesting to the need to extend currently existing transcriptome atlases (Ghaddar et al., 2023; Packer et al., 2019; Roux et al., 2023; Taylor et al., 2025, 2021) with protein expression atlases.

As a proof of concept for scaling such a CRISPR/Cas9-engineered project to the whole genome, we report here the results of a pilot study in which we aimed to CRISPR/Cas9-engineer reporter tags into 30 genomic loci, generated through simultaneous tagging of three genes at a time. A subset of these 30 genes reveals unexpected patterns of expression and localization, illustrating how protein tagging approaches provide a route for new discoveries. We discuss here several aspects of such tagging approaches, highlighting promises, as well as limitations.

Results

Pipeline for pooled high throughput tagging by CRISPR/Cas9

Recent methods for CRISPR/Cas9 mediated genomic insertions have pushed efficiencies to sufficient levels to simultaneously insert multiple fluorophores (mNeonGreen and mScarlet) as well as a co-CRISPR marker (*dpy-10*) at three independent loci in a single injection (Eroglu et al., 2023). Thus, in principle current methods could allow tagging of at least 3 independent loci in one injection, particularly if

a co-CRISPR marker is omitted. Endogenously tagged genes can be detected by fluorescence stereomicroscopy if the fluorophores are expressed at sufficiently high levels and hence are readily visible as heterozygotes in F1 by visual screening without co-CRISPR markers. We reasoned that ubiquitously and highly expressed genes could be used as insertion markers and pooled with low or modestly expressed genes, with a first round of visual screening of the highly expressed genes followed by targeted PCR or higher power imaging of the lower expressed genes **(Fig. 1A)**.

To facilitate screening by fluorescence stereomicroscopy and co-imaging, we chose the three currently brightest fluorescent proteins in the blue, green and red emission spectra: mTagBFP2 (Subach et al., 2011), mStayGold(J) (Ando et al., 2024) and mScarlet3 (Gadella et al., 2023)**,** respectively (**Fig. 1B**). These tags are significantly brighter than first generation fluorophores, thereby allowing the detection of even very lowly expressed proteins. Autofluorescence in *C. elegans* is highest in blue and green wavelengths, and lowest in red. Furthermore, we noted that nematode growth medium (NGM) plates seeded with OP50 also contribute to background fluorescence with conventional blue and green filters on a fluorescent stereomicroscope **(Fig. 1C)**. In contrast, we observed minimal background with yellow and red emission filters from worms as well as bacteria. We also noted from existing tagged strains that mStayGold could be observed brightly with YFP filters, which, combined with the minimal background, yielded a greater contrast with untagged worms enabling easier screening of mStayGold positive worms compared to conventional GFP filters **(Fig 1C)**. However, given the suboptimal excitation and emission of visualizing mStayGold with a YFP filter, mScarlet3 in practice yielded the greatest signal-to-noise ratio for visual screening. Thus, to balance fluorescence signal-to-noise ratios with predicted gene expression, we assigned the highest expressed genes to be tagged with mTagBFP2, modestly expressed genes with mStayGold, and lowest expressed genes with mScarlet3.

In the pilot experiment, we sought to establish three major principles: (1) the endogenous expression ranges at which fluorophores could be utilized for visual screening; (2) success rate of triple pooled tagging; (3) functionality of tagged proteins. We thus selected 30 genes across a variety of bulk transcript expression ranges which are predicted to be ubiquitously expressed based on molecular function or single cell RNA sequencing (scRNA-seq) data **(Table 1, Fig. 2A, B)** (Taylor et al., 2021). The proteins are generally predicted to be broadly expressed and cover a very wide range of predicted cellular functions and predicted subcellular localization patterns (nuclear, cytosolic, transmembrane). 26 of 30 genes are annotated to display lethal, sterile, or other phenotypes by RNAi, enabling the assessment of tagged protein functionality.

The *C. elegans* genome encodes around 20,000 genes. Approximately 6,700 marker genes would need to be sufficiently highly expressed to be used as visible markers to group with the entire genome in sets of three. In bulk RNA-seq of N2 worms at L4 stage, 6,804 genes are expressed above 20 fragments per kilobase of transcript per million mapped reads (FPKM) (Eroglu et al., 2024; Gerstein et al., 2010; Harris et al., 2020). Based on our prior experience comparing bulk transcript FPKM with visibility, we divided these genes into three subsets to assess the detection limit of the three fluorophores: ≥500 RPKM for mTagBFP2; 100-500 RPKM for mStayGold; and 18-100 RPKM for mScarlet3.

We tagged genes either at their N- or C-terminus, primarily aiming to capture as many predicted isoforms as possible while retaining a potential guide RNA target sequence within 10 nucleotides of the start or stop codon. Where both termini satisfied these conditions, we selected the C-terminus over the N-terminus, as previous studies have suggested that C-terminal tags are less likely to disrupt protein localization and expression (Davidi et al., 2019; Palmer and Freeman, 2004) and that genome wide tagging attempts in yeast have successfully used C-termini (Huh et al., 2003) with high concordance to prior annotated data. However, we acknowledge that subclasses of proteins such as peroxisomal proteins, C-tail- or glycophosphatidylinositol-anchored proteins, lipidated proteins and proteins with retrieval motifs, may only show correct localization when tagged at the N-terminus (Yofe et al., 2016).

Efficient tagging across loci, expression levels and target pools

We co-injected the 30 selected targets as 10 pools (**Table 1**), with each pool being injected into 5 worms. Overall, we successfully isolated 8/10 mTagBFP2 tags, 9/10 mStayGold tags, and 7/10 mScarlet3 tags **(Fig. 3A, B)**. All but one tag (*cox-6B*::mTagBFP2) was visible in the F1 generation of injected P0 animals, and these were subsequently isolated among F2 worms positive for the other tags in the pool. In 4 pools, we isolated all 3 tags within the same worm, and in an additional experiment all 3 tags but in pairs. The remaining tags were isolated as pairs or singles. Even when one or two tags in the pool failed, the remaining tags in the same pool could be independently isolated. Overall, 24 of 30 tags were successfully isolated, all of which were visible with fluorescence stereomicroscopy.

During our screening, we made several observations which enabled us to assess whether failures were due to technical or biological reasons. Our first attempt with pools 4, 6 and 10 yielded no positive progeny for any of the 3 targets in the respective pools despite the injected worms showing a healthy number of viable offspring **(Fig. 3C)**. We reasoned that, given overall efficiency for individual tags, this was due to technical failures rather than the impossibility of functionally tagging all three genes within the pool. Indeed, upon reconstitution of the injection mixes with additional quality control steps (i.e., ensuring that repair templates were of sufficient purity as assessed by proper A260/230 and A260/280 ratios; see Methods), we successfully obtained 6/9 of the originally failed tags (**Fig. 3C**). In contrast, failures in obtaining one or two of the tags in the pool were generally interpreted as loss of function tags, and in these cases, we frequently observed fluorescent but dead embryos on the plates. While the combination of high brood size with 0/3 successful tags was a good indicator to reattempt individual pools, overall brood size across successful injections did not correlate with efficiency of insertion (**Fig. 3D**).

Unanticipated cell type specific enrichment of targets

We assessed the expression patterns of the protein fusions by confocal microscopy in adult worms and compared them to expected tissue type specific abundances predicted by stage-matched scRNA-seq datasets (Ghaddar et al., 2023; Taylor et al., 2021). Based on molecular functions in fundamental biological processes,

all genes were expected to be ubiquitously expressed. Likewise, scRNA-seq indicated that transcript abundance was ubiquitous and without strong tissue-specific enrichment with few exceptions.

In contrast, we found notable examples of mismatch between expected and observed expression patterns. The translation elongation factor EEF-1A.1 showed strong germline enrichment whereas it was predicted to be enriched in various cell types including germ cells, epidermis, intestinal and excretory cells by scRNA seq (**Fig. 4**). The acyl-CoA dehydrogenase ACDH-10 was somatically broadly and evenly expressed but strongly depleted in the germline, while we expected enrichment in intestinal and epidermal tissues given its role in fatty acid metabolism (**Fig. 4**). Strikingly, the glucokinase HXK-1 showed strong enrichment in the gonadal sheath, whereas we expected ubiquitous expression with potential enrichment in glycolysis upregulated tissues such as muscle and neurons (**Fig. 5**). The lysyl-tRNA synthetase KARS-1 showed germline enrichment, while it was expected to be evenly distributed across tissues given its role in translation, but was correctly predicted by scRNA seq to be germline enriched (**Fig. 6A-C**). The nucleosome assembly protein NAP-1 was highest expressed in the germline, whereas it was predicted to be highest in the distal tip cell and pharynx (**Fig. 7A**). In contrast, the phosphoribosyl transferase HPRT-1 and ATP citrate lyase ACLY-1 were enriched in subsets of glia and neurons, respectively, consistent with scRNA-seq datasets (**Fig. 7A, B**). The cytochrome c oxidase subunit 6B COX-6B was highest enriched in the distal gonad and intestine (**Fig. 8A**), while scRNA-seq predicted highest expression in the pharynx. The tryptophan 2,3-dioxygenase TDO-2 showed highly enriched or specific expression in epidermis, whereas its expression predicted by scRNA seq was ubiquitous with modest epidermal enrichment (**Fig. 8A, B**). The hydroxyacyl-CoA dehydrogenase F54C8.1, which we independently isolated from the same pool as COX-6B and TDO-2, showed high enrichment in sperm (**Fig. 8B, C**) whereas it was predicted to be enriched in the intestine based on function but correctly suggested by scRNA-seq as highly enriched in or specific to sperm. We note that the gut granule signal in F54C8.1::mScarlet3 tagged worms was comparable to that observed in untagged N2 worms, indicating that it is background autofluorescence **(Fig. 8D)**. The heat shock protein HSP-1 was enriched in the germline while predicted to be ubiquitous and evenly expressed by scRNA-seq (**Fig. 9**). Similarly, the triosephosphate isomerase TPI-1 was observed to be enriched in neurons, which was not predicted from scRNA-seq data (**Fig. 9**). Overall, numerous genes with ubiquitous and evenly distributed transcript expression showed tissue-specific enrichment not predicted by transcript levels with a general trend towards underestimation of relative protein levels in the germline.

Unanticipated subcellular localization patterns of proteins

Given that our tags were designed to be translationally fused to the proteins of interest, we assessed their subcellular localizations and compared them to localizations annotated experimentally or predicted by homology (Harris et al., 2020). Most notably, we found two proteins, GDH-1 and KARS-1, detected in mitochondria by mass spectrometry (Guo et al., 2025; Li et al., 2009), and co-tagged in the same worm, to display mitochondrial-like localizations; however, they did not overlap (**Fig. 6B, C**). This finding suggests that subsets of mitochondria in the gonad may have varying markers and

metabolic specializations. Indeed, a study published around the time of our pilot experiment described metabolically distinct subpopulations of mitochondria depending on cellular energy demand (Ryu et al., 2024).

Furthermore, we found that the V-ATPase VHA-2, involved in proton transport and predicted to be localized to cell and vesicular membranes and detected in membrane fractions by mass spectrometry (Mawuenyega et al., 2003), displayed perinuclear localization in the gonad specifically (**Fig. 10**). This indicates that cell type-specific nuclear envelopes or endoplasmic reticula may have unique demands for pH control which are yet unexplored, to our knowledge.

In general, subcellular localizations were predicted more accurately than cell type-specific abundance, though we observed additional dynamic localization patterns unanticipated by molecular function. For instance, the DNA-binding nuclear phosphoprotein p8 homolog ZK632.9 was predicted to be nuclear but observed diffusely with modest nuclear enrichment (**Fig. 4**). We note that this may be due to its small size of 62 amino acids, which could be influenced by the larger 226 amino acid mStayGold tag. However, the tagged worm does not show any annotated ZK632.9 depletion phenotypes such as sterility (Simmer et al., 2003), indicating that it is to some extent functional. Additionally, the histone acetyltransferase 1 HAT-1 showed broadly nuclear localization as predicted but was cytoplasmic in early embryos (**Fig. 10**), indicating dynamic subcellular localization. Finally, the tryptophan 2,3-dioxygenase TDO-2 with previously unknown localization appeared cytoplasmic (**Fig. 8A, B**).

Discussion

We have provided proof of concept for the ease and efficiency of tagging at least three proteins at once during a single round of injection. We also find that modern-day fluorophores allow the visualization of proteins at a wide range of expression levels, even at the lowest expression levels tested. Extrapolating our results to annotated transcript levels of *C. elegans* genes, we predict that at least 324 proteins could be directly visualized by mTagBFP2, 1578 with mStayGold, and 7249 with mScarlet3 for high throughput screening using low powered fluorescence stereomicroscopy. At minimum, this could enable annotation of one third of *C. elegans* proteins and at best provide sufficient markers to co-tag the entire genome in groups of 3 genes. In the latter scenario, brighter genes would be used as coCRISPR markers grouped with dimmer genes which would be isolated among bright gene positive worms through alternative screening methods such as PCR or higher magnification, longer exposure fluorescence microscopy (**Fig. 1A**). Given many genes are not broadly expressed, their bulk transcript levels in whole worms are likely to underestimate expression within subsets of cells, which may aid in their detectability and isolation. Similarly, nuclear localized proteins will likely be easier to detect even at low expression levels, given the concentration of signal in small subcellular compartments. Even with the relatively small number of targets tested, we have observed numerous unanticipated patterns of tissue-specific enrichment and subcellular localization.

By tagging essential genes and in most cases observing no overt effect on viability of the animal, we have also shown that many N- or C-terminal tags do not disrupt protein

function in an obvious way. Given most genes we targeted were essential, we could infer the effect of tagging through the viability of tagged worms. We interpret failed tagging attempts as disruptions to protein function, which would necessarily prevent isolation of viable worms, which occurred in 20% of attempted tags. In our experience, inclusion of an 18 amino acid linker reduces the chances of disrupting protein function, and for all our tags we have included a universal 18 amino acid linker. However, we do fully recognize that N- or C-terminal tagging will not always produce properly localized and/or functional protein. We have shown a few examples of functional disruptions here, and it is well appreciated in the field that many types of proteins are difficult to tag without either disruption of function or localization, or both. We also recognize that in some cases where worms appear superficially wild type, tags may have more subtle effects on protein function which may impair fitness. One strategy to predict whether a fused protein may lose function is by comparing the predicted structure of the fusion to the original protein of interest with Alphafold (Jumper et al., 2021). However, even if the fused protein is predicted to fold normally, steric interactions, multimeric complex formation and other intrinsic properties of the protein such as solubility may still be affected and are difficult to predict based solely on structure. Smaller tags, including epitope tags like 3xFLAG, V5, and SPOT, can be expected to minimize the risk of such disruptions. However, these present their own problems. Classic epitope tags require fixation of samples, thereby preventing the ability for live imaging and introducing caveats of tissue deformation during fixation and permeabilization. The Halotag system (Los et al., 2008) has the advantage of also being small and allowing the conjugation of fluorescent dyes, but the need for even dye permeability across cell types and tissues introduces an important limitation. The Alfa tag system that operates through stabilizing fluorescent protein-fused nanobodies (Götzke et al., 2019) may, for steric reasons, carry the same risk of disruption of endogenous protein function or subcellular localization as covalent tagging. For genes with no obvious mutant phenotypes, assessing the functionality of tagged alleles will be difficult. For such genes, if an interaction-based phenotype such as synthetic lethality is annotated, the tagged gene can be crossed to the interactor to assess synthetic phenotypes. We envision that any novel or unexpected subcellular localizations observed will require validation by independent methods, as in every high-throughput experiment. Overall, the balance between throughput and functionality is, in our view, best struck with translational fusions of fluorophores to proteins of interest, which allows direct live visualization while carrying well understood limitations and artefacts.

Proteins expressed exclusively in a particular cell type and restricted to very specific subcellular sites, such as a handful of synapses in the nerve ring, pose significant challenges for accurate cell identification. Likewise, the cellular origin of secreted proteins may be difficult to trace based solely on translational reporter fusions. In such cases, the labeling of the entire cytoplasm of cells (or, alternatively, its membrane) may be useful. In theory this could relatively easily be achieved by using a double fluorophore tag, where the second fluorophore is separated by a splice leader or T2A sequence from the protein-tagging, first fluorophore. However, in practical terms, such a cassette is much larger than 1 kilobases, and in our experience larger inserts lead to a significant drop in the efficiency of successful insertion, therefore negatively affecting throughput. Where such cases are observed, a small, self-cleaving T2A peptide sequence, which frees up the tag from the

fused protein, can be easily inserted in a second round of CRISPR/Cas9 to validate the cell of origin.

Fluorescent landmarks can facilitate delineation of cell boundaries, nuclei, cytoskeleton, or other organelles like mitochondria. Use of visible fluorophores such as mTagBFP2, mStayGold and mScarlet3 leaves the far-red spectrum open for the inclusion of markers such as membrane-targeted far-red fluorescent proteins like miRFP680 (Zhang et al., 2023) or cell-permeant near-infrared dyes. In an ideal case, pools of target-specific mTagBFP2, mStayGold and mScarlet3 would be directly injected into integrated marker strains expressing miRFP680 targeted to landmarks of interest. Another option is to group genes strategically to include one ubiquitously expressed predicted cytoplasmic gene. With confocal or super-resolution microscopy, such impromptu landmarks may provide some indication of subcellular localization for particularly difficult proteins as a starting point.

Illustrating the discovery aspects of protein tagging, we found that two predicted mitochondrial proteins, KARS-1 and GDH-1, localized to distinct subsets of mitochondria. This is consistent with parallel reports of functionally differentiated mitochondria in cells with varying metabolic demands (Ryu et al., 2024). To increase the chance of such discoveries, it may be informative to group multiple proteins predicted to localize to the same subcellular compartments together. With sufficient coverage, this can yield information on the differentiation or uniformity of organelles within the same cell.

We routinely inject 30 animals per day, with 5 worms per injection mix. In theory, this can enable injection of 6 pools targeting 18 genes per day, or 90 per week. A single lab can thus expect 1x coverage of the genome in approximately 290 weeks or 5.6 years, working at a modest pace, assuming that 30% of injections will need to be repeated. With simple quality control steps and careful template preparation, such as ensuring that purity of repair templates is sufficient, we project that fewer injections will need to be repeated. Recent advances in automated injection systems (Pan et al., 2024) combined with fluorescent worm sorters may further accelerate coverage. The market cost of ordering crRNAs and annealing with tracrRNA is currently $76 USD; the cost of each target-specific primer pair for this study was $23 USD. Approximately $8 USD of commercially purified Cas9 is used per injection mix; $1.53 USD for spin columns. Without the one-time small cost of universal fluorophore template synthesis, the material cost of tagging each gene is $109 USD and would be $2.2 million USD for the entire genome. While the time, material and personnel costs would be daunting for an individual lab, it would be feasible for a group of committed labs within a few years. A centrally coordinated, but distributed effort would avoid wasteful duplications in gene tagging efforts that are already documented in the literature, such as PGL-1, DAF-16 and several Argonaute genes which each have been tagged at least 4 independent times with GFP (Leyhr et al., 2025). The available Caenorhabditis Genetics Center (CGC) strain repository funded by NIH Office of Research Infrastructure Programs (P40 OD010440) as well as emerging database infrastructure dedicated to annotating endogenously tagged strains (Leyhr et al., 2025) will facilitate strain as well as reagent (e.g., guide RNA sequence) cataloging and dissemination.

Materials and methods

Nematode maintenance

All *C. elegans* strains were maintained at 20 °C on nematode growth medium (NGM) agar plates with OP50 Escherichia coli as a food source (Brenner, 1974). All CRISPR/Cas9 insertion strains were generated in the Bristol N2 background.

CRISPR/Cas9-mediated genomic insertions

Sequences of oligonucleotides, fluorophore templates and CRISPR RNAs (crRNAs) are provided in **Supplementary Table 1**. Cas9 protein was ordered from Integrated DNA Technologies (Park Coralville, IA, USA) (Alt-R® S.p. Cas9 Nuclease V3, 500 µg, Cat: 1081059), aliquoted at 5 mg/mL concentration, and stored at -80°C until use. tracrRNA was ordered from IDT (Alt-R® CRISPR-Cas9 tracrRNA, 20 nmol, Cat: 1072532), resuspended in the provided nuclease-free duplex buffer to 0.4 µg/µl (18 µM), aliquoted, and stored at -80°C until use (Ghanta and Mello, 2020).

All guide RNAs were designed using IDT's guide RNA design tool. The highest 'on-target' scoring guide within 10 bp of the desired insert site was selected and ordered at 2 nmol scale as a custom Alt-R™ crRNA from IDT, resuspended in 20 µL nuclease-free duplex buffer (provided by IDT with tracrRNA order), and stored at 100 µM at -20°C.

Single-stranded DNA repair templates were generated as described **(Eroglu et al., 2023)** with minor modifications. Double-stranded linear DNA templates encoding *C. elegans* codon-optimized mTagBFP2, mStayGold(J), and mScarlet3 were ordered as 'FragmentGENEs' from GENEWIZ (South Plainfield, NJ, USA) and used as universal templates. Primers encoding 45 bp gene-specific homology followed by universal priming sequences (Eurofins Genomics, 10 nmol scale) were used to introduce homology arms to universal fluorophore templates. One primer was phosphorylated at the 5' end to enable strand-specific digestion by lambda exonuclease. Templates were amplified using Q5 polymerase from New England Biolabs (NEB; M0491) with 0.5 ng of template per reaction at an annealing temperature of 70°C and extension time of 30 seconds with 40 cycles in total. For each gene, a single 50 µL reaction was run. After amplification, individual 50 µL reactions were pooled in sets of three as indicated in **Table 1** before proceeding to column purification (PureLink™ PCR Purification Kit, Invitrogen, Cat: K310001). Purified pools of double stranded DNA were digested with lambda exonuclease (NEB, cat: M0262) for 20-30 minutes at 37°C followed by micro-column purification (Monarch® Spin PCR & DNA Cleanup Kit (5 µg)), NEB, Cat: T1130) using a modified protocol described previously (Eroglu et al., 2023). For re-runs of pools 4, 6 and 10 which failed initially, we regenerated the repair templates and ensured that after each column purification, the A260/230 ratio of the purified DNA was ≥2.2 and A280/260 was 1.8 ± 0.05 when measured with a Nanodrop spectrophotometer.

Pooled injection mixes were set up by first annealing the crRNAs with tracrRNA as follows: 3.8 µL of tracrRNA and 0.4 µL each of 3 target-specific crRNAs were mixed, then incubated at 95 °C for 5 min followed by 10 °C for 5 min. The Cas9 RNP complex (Paix et al., 2015) was formed by incubating 2.55 µL of tracr-crRNA annealed guide pools with

0.5 µL of Cas9 for 5 min at room temperature. Following Cas9 RNP formation, 4 µL of pooled single stranded DNA template (≥400 ng/µL) was added. Constituted injection mixes were injected into *C. elegans* gonads using standard practice. Progeny of injected worms were screened for fluorescence under a Leica M205 FCA fluorescence stereomicroscope with a Planapo 1.6x M-series objective, pE 300 white MB light source and the following filters: Filter set ET DAPI BP - M205FA/M165FC for mTagBFP2; Filter set ET YFP - M205FA/M165FC for mStayGold; and Filter set ET mCherry - M205FA/M165FC for mScarlet3. Worms positive for fluorescence were isolated on individual plates to establish strains.

Confocal microscopy

Worms at the adult stage were immobilized with 50mM sodium azide on 5% agarose pads and imaged on a Zeiss LSM980 laser scanning confocal microscope (C-Apochromat 40x/1.20 water objective) with the 405, 488 and 561nm lasers at variable power settings. Detection was on spectral channels of the 32-channel GaAsP-PMT detector with gain at 800 V for all strains except for group 2 (Fig. 4) which was at 725V. For mTagBFP2, detection wavelengths were set to 411-481 nm; for mStayGold, detection was at 490-552 nm; for mScarlet3, detection was at 570-694 nm. Multitracking with the LSM980 acousto-optic tunable filter (AOTF) was used to excite and detect mTagBFP2 and mScarlet3 simultaneously on the first light track (405 nm and 561 nm) while mStayGold was excited and detected on a second track (488 nm) with switching between tracks line by line. Scanning was bidirectional with 8x line averaging; pinhole was set to 31 µM (1.00 AU for mScarlet3). Worms were imaged at 1x sampling (213.13 µm x 213.13 µm at 2576 x 2576 pixels) then stitched together on ImageJ (FIJI) (Preibisch et al., 2009; Schindelin et al., 2012) to reconstitute the whole worm. Worms were linearized using the straighten function on ImageJ.

Software and statistical analysis

Genome plots showing gene loci were adapted from the University of California Santa Cruz (UCSC) genome browser (*http://genome.ucsc.edu*) (Casper et al., 2026) and represent ce11 version of the *Caenorhabditis elegans* genome. Plots were generated and statistical analyses performed on GraphPad Prism version 10.6.1.

Data availability

Reported strains are available from the authors upon request. Source data for Figs. 3 and 8D are provided in **Supplementary Table 2**.

Acknowledgements

We thank Chi Chen for performing all injections, members of the Hobert lab, as well as Stephen Nurrish, Emily Bayer, Luisa Cochella, Geraldine Seydoux, Jean-Louis Bessereau, Hari Shroff, David Sherwood, Jake Leyhr, Brent Derry and Piali Sengupta for

comments on the manuscript. Matthew Eroglu is supported by the Leon Levy Foundation and the New York Academy of Sciences through a Leon Levy Neuroscience Scholarship. This work was supported by the Howard Hughes Medical Institute.

References

Altun ZF, Herndon LA, Wolkow CA, Crocker C, Lints R, Hall DH. 2002. WormAtlas.

Ando R, Shimozono S, Ago H, Takagi M, Sugiyama M, Kurokawa H, Hirano M, Niino Y, Ueno G, Ishidate F, Fujiwara T, Okada Y, Yamamoto M, Miyawaki A. 2024. StayGold variants for molecular fusion and membrane-targeting applications. Nature Methods **21**:648–656. DOI: https://doi.org/10.1038/s41592-023-02085-6

Bayer EA, Mango SE, Hobert O, Schier AF. 2025. The DEG/ENaC channel DEGT-1 is a proprioceptor of C. elegans foregut movement. Current biology: CB **35**:5332-5343.e5. DOI: https://doi.org/10.1016/j.cub.2025.09.065, PMID: 41135517

Brenner S. 1974. The genetics of Caenorhabditis elegans. Genetics **77**:71–94. DOI: https://doi.org/10.1093/genetics/77.1.71, PMID: 4366476

Casper J, Speir ML, Raney BJ, Perez G, Nassar LR, Lee CM, Hinrichs AS, Gonzalez JN, Fischer C, Diekhans M, Clawson H, Benet-Pages A, Barber GP, Vaske CJ, van Baren MJ, Wang K, Rodriguez YJP, Jenkins-Kiefer JA, Chalamala M, Haussler D, Kent WJ, Haeussler M. 2026. The UCSC Genome Browser database: 2026 update. Nucleic Acids Research **54**:D1331–D1335. DOI: https://doi.org/10.1093/nar/gkaf1250, PMID: 41251146

Cook SJ, Jarrell TA, Brittin CA, Wang Y, Bloniarz AE, Yakovlev MA, Nguyen KCQ, Tang LT-H, Bayer EA, Duerr JS, Bülow HE, Hobert O, Hall DH, Emmons SW. 2019. Whole-animal connectomes of both Caenorhabditis elegans sexes. Nature **571**:63–71. DOI: https://doi.org/10.1038/s41586-019-1352-7

Davidi D, Weill U, Krieger G, Avihou Z, Milo R, Schuldiner M. 2019. Assessment of GFP tag position on protein localization and growth fitness in yeast. Journal of molecular biology **431**:636–641. DOI: https://doi.org/10.1016/j.jmb.2018.12.004, PMID: 30550779

Eroglu M, Yu B, Derry WB. 2023. Efficient CRISPR/Cas9 mediated large insertions using long single-stranded oligonucleotide donors in C. elegans. The FEBS journal. DOI: https://doi.org/10.1111/febs.16876, PMID: 37254814

Eroglu M, Zocher T, McAuley J, Webster R, Xiao MZX, Yu B, Mok C, Derry WB. 2024. Noncanonical inheritance of phenotypic information by protein amyloids. Nature Cell Biology 1–13. DOI: https://doi.org/10.1038/s41556-024-01494-9

Gadella TWJ, van Weeren L, Stouthamer J, Hink MA, Wolters AHG, Giepmans BNG, Aumonier S, Dupuy J, Royant A. 2023. mScarlet3: a brilliant and fast-maturing red fluorescent protein. Nature Methods **20**:541–545. DOI: https://doi.org/10.1038/s41592-023-01809-y

Gerstein MB, Lu ZJ, Van Nostrand EL, Cheng C, Arshinoff BI, Liu T, Yip KY, Robilotto R, Rechtsteiner A, Ikegami K, Alves P, Chateigner A, Perry M, Morris M, Auerbach RK, Feng X, Leng J, Vielle A, Niu W, Rhrissorrakrai K, Agarwal A, Alexander RP, Barber G, Brdlik CM, Brennan J, Brouillet JJ, Carr A, Cheung M-S, Clawson H, Contrino S, Dannenberg LO, Dernburg AF, Desai A, Dick L, Dosé AC, Du J, Egelhofer T, Ercan S, Euskirchen G, Ewing B, Feingold EA, Gassmann R, Good

PJ, Green P, Gullier F, Gutwein M, Guyer MS, Habegger L, Han T, Henikoff JG, Henz SR, Hinrichs A, Holster H, Hyman T, Iniguez AL, Janette J, Jensen M, Kato M, Kent WJ, Kephart E, Khivansara V, Khurana E, Kim JK, Kolasinska-Zwierz P, Lai EC, Latorre I, Leahey A, Lewis S, Lloyd P, Lochovsky L, Lowdon RF, Lubling Y, Lyne R, MacCoss M, Mackowiak SD, Mangone M, McKay S, Mecenas D, Merrihew G, Miller DM, Muroyama A, Murray JI, Ooi S-L, Pham H, Phippen T, Preston EA, Rajewsky N, Rätsch G, Rosenbaum H, Rozowsky J, Rutherford K, Ruzanov P, Sarov M, Sasidharan R, Sboner A, Scheid P, Segal E, Shin H, Shou C, Slack FJ, Slightam C, Smith R, Spencer WC, Stinson EO, Taing S, Takasaki T, Vafeados D, Voronina K, Wang G, Washington NL, Whittle CM, Wu B, Yan K-K, Zeller G, Zha Z, Zhong M, Zhou X, modENCODE Consortium, Ahringer J, Strome S, Gunsalus KC, Micklem G, Liu XS, Reinke V, Kim SK, Hillier LW, Henikoff S, Piano F, Snyder M, Stein L, Lieb JD, Waterston RH. 2010. Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project. Science (New York, N.Y.) **330**:1775–1787. DOI: https://doi.org/10.1126/science.1196914, PMID: 21177976

Ghaddar A, Armingol E, Huynh C, Gevirtzman L, Lewis NE, Waterston R, O'Rourke EJ. 2023. Whole-body gene expression atlas of an adult metazoan. Science Advances **9**:eadg0506. DOI: https://doi.org/10.1126/sciadv.adg0506, PMID: 37352352

Ghanta KS, Mello CC. 2020. Melting dsDNA Donor Molecules Greatly Improves Precision Genome Editing in Caenorhabditis elegans. Genetics **216**:643–650. DOI: https://doi.org/10.1534/genetics.120.303564, PMID: 32963112

Götzke H, Kilisch M, Martínez-Carranza M, Sograte-Idrissi S, Rajavel A, Schlichthaerle T, Engels N, Jungmann R, Stenmark P, Opazo F, Frey S. 2019. The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications. Nature Communications **10**:4403. DOI: https://doi.org/10.1038/s41467-019-12301-7

Grün D, Kirchner M, Thierfelder N, Stoeckius M, Selbach M, Rajewsky N. 2014. Conservation of mRNA and protein expression during development of C. elegans. Cell Reports **6**:565–577. DOI: https://doi.org/10.1016/j.celrep.2014.01.001, PMID: 24462290

Guo X, Lu J, Miao L, Shen E. 2025. Mitochondrial Proteome Reveals Metabolic Tuning by Restricted Insulin Signaling to Promote Longevity in Caenorhabditis elegans. Biology **14**:279. DOI: https://doi.org/10.3390/biology14030279

Harris TW, Arnaboldi V, Cain S, Chan J, Chen WJ, Cho J, Davis P, Gao S, Grove CA, Kishore R, Lee RYN, Muller H-M, Nakamura C, Nuin P, Paulini M, Raciti D, Rodgers FH, Russell M, Schindelman G, Auken KV, Wang Q, Williams G, Wright AJ, Yook K, Howe KL, Schedl T, Stein L, Sternberg PW. 2020. WormBase: a modern Model Organism Information Resource. Nucleic Acids Research **48**:D762–D767. DOI: https://doi.org/10.1093/nar/gkz920, PMID: 31642470

Huff J. 2015. The Airyscan detector from ZEISS: confocal imaging with improved signal-to-noise ratio and super-resolution. Nature Methods **12**:i–ii. DOI: https://doi.org/10.1038/nmeth.f.388

Huh W-K, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK. 2003. Global analysis of protein localization in budding yeast. Nature **425**:686–691. DOI: https://doi.org/10.1038/nature02026

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. Nature **596**:583–589. DOI: https://doi.org/10.1038/s41586-021-03819-2

Kumar A, Wu Y, Christensen R, Chandris P, Gandler W, McCreedy E, Bokinsky A, Colón-Ramos DA, Bao Z, McAuliffe M, Rondeau G, Shroff H. 2014. Dual-view plane illumination microscopy for rapid and spatially isotropic imaging. Nature Protocols **9**:2555–2573. DOI: https://doi.org/10.1038/nprot.2014.172, PMID: 25299154

Lambert TJ. 2019. FPbase: a community-editable fluorescent protein database. Nature Methods **16**:277–278. DOI: https://doi.org/10.1038/s41592-019-0352-8

Leyhr J, Chi Q, Zeng L, Shen E-Z, Zou W, Sherwood D. 2025. WormTagDB: A Systematic Survey of Endogenously Tagged Proteins in C. elegans and Roadmap Towards the Tagged Proteome. DOI: https://doi.org/10.64898/2025.12.02.691955

Leyva-Díaz E, Cesar M, Pe K, Jordá-Llorens JI, Valdivia J, Hobert O. 2025. Alternative splicing controls pan-neuronal homeobox gene expression. Genes & Development **39**:209–220. DOI: https://doi.org/10.1101/gad.352184.124, PMID: 39730199

Li J, Cai T, Wu P, Cui Z, Chen X, Hou J, Xie Z, Xue P, Shi L, Liu P, Yates JR, Yang F. 2009. Proteomic analysis of mitochondria from Caenorhabditis elegans. Proteomics **9**:4539–4553. DOI: https://doi.org/10.1002/pmic.200900101, PMID: 19670372

Liu Y, Beyer A, Aebersold R. 2016. On the Dependency of Cellular Protein Levels on mRNA Abundance. Cell **165**:535–550. DOI: https://doi.org/10.1016/j.cell.2016.03.014

Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, Wood MG, Learish R, Ohana RF, Urh M, Simpson D, Mendez J, Zimmerman K, Otto P, Vidugiris G, Zhu J, Darzins A, Klaubert DH, Bulleit RF, Wood KV. 2008. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. ACS chemical biology **3**:373–382. DOI: https://doi.org/10.1021/cb800025k, PMID: 18533659

Mawuenyega KG, Kaji H, Yamuchi Y, Shinkawa T, Saito H, Taoka M, Takahashi N, Isobe T. 2003. Large-scale identification of Caenorhabditis elegans proteins by multidimensional liquid chromatography-tandem mass spectrometry. Journal of Proteome Research **2**:23–35. DOI: https://doi.org/10.1021/pr025551y, PMID: 12643540

Packer JS, Zhu Q, Huynh C, Sivaramakrishnan P, Preston E, Dueck H, Stefanik D, Tan K, Trapnell C, Kim J, Waterston RH, Murray JI. 2019. A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell resolution. Science (New York, N.Y.) **365**:eaax1971. DOI: https://doi.org/10.1126/science.aax1971, PMID: 31488706

Paix A, Folkmann A, Rasoloson D, Seydoux G. 2015. High Efficiency, Homology-Directed Genome Editing in Caenorhabditis elegans Using CRISPR-Cas9 Ribonucleoprotein Complexes. Genetics **201**:47–54. DOI: https://doi.org/10.1534/genetics.115.179382, PMID: 26187122

Palmer E, Freeman T. 2004. Investigation Into the use of C- and N-terminal GFP Fusion Proteins for Subcellular Localization Studies Using Reverse Transfection Microarrays. Comparative and Functional Genomics **5**:342–353. DOI: https://doi.org/10.1002/cfg.405, PMID: 18629169

Pan P, Zoberman M, Zhang P, Premachandran S, Bhatnagar S, Pilaka-Akella PP, Sun W, Li C, Martin C, Xu P, Zhang Z, Li R, Hung W, Tang H, MacGillivray K, Yu B, Zuo R, Pe K, Qin Z, Wang S, Li A, Derry WB, Zhen M, Saltzman AL, Calarco JA, Liu X. 2024. Robotic microinjection enables large-scale transgenic studies of Caenorhabditis elegans. Nature Communications **15**:8848. DOI: https://doi.org/10.1038/s41467-024-53108-5

Preibisch S, Saalfeld S, Tomancak P. 2009. Globally optimal stitching of tiled 3D microscopic image acquisitions. Bioinformatics **25**:1463–1465. DOI: https://doi.org/10.1093/bioinformatics/btp184

Rankin BR, Moneron G, Wurm CA, Nelson JC, Walter A, Schwarzer D, Schroeder J, Colón-Ramos DA, Hell SW. 2011. Nanoscopy in a Living Multicellular Organism Expressing GFP. Biophysical Journal **100**:L63–L65. DOI: https://doi.org/10.1016/j.bpj.2011.05.020, PMID: 21689517

Reilly MB, Cros C, Varol E, Yemini E, Hobert O. 2020. Unique homeobox codes delineate all the neuron classes of C. elegans. Nature **584**:595–601. DOI: https://doi.org/10.1038/s41586-020-2618-9, PMID: 32814896

Reilly MB, Tekieli T, Cros C, Aguilar GR, Lao J, Toker IA, Vidal B, Leyva-Díaz E, Bhattacharya A, Cook SJ, Smith JJ, Kovacevic I, Gulez B, Fernandez RW, Bradford EF, Ramadan YH, Kratsios P, Bao Z, Hobert O. 2022. Widespread employment of conserved C. elegans homeobox genes in neuronal identity specification. PLoS genetics **18**:e1010372. DOI: https://doi.org/10.1371/journal.pgen.1010372, PMID: 36178933

Roux AE, Yuan H, Podshivalova K, Hendrickson D, Kerr R, Kenyon C, Kelley D. 2023. Individual cell types in C. elegans age differently and activate distinct cell-protective responses. Cell Reports **42**:112902. DOI: https://doi.org/10.1016/j.celrep.2023.112902, PMID: 37531250

Ryu KW, Fung TS, Baker DC, Saoi M, Park J, Febres-Aldana CA, Aly RG, Cui R, Sharma A, Fu Y, Jones OL, Cai X, Pasolli HA, Cross JR, Rudin CM, Thompson CB. 2024. Cellular ATP demand creates metabolically distinct subpopulations of mitochondria. Nature **635**:746–754. DOI: https://doi.org/10.1038/s41586-024-08146-w

Santella A, Kolotuev I, Kizilyaprak C, Bao Z. 2022. Cross-modality synthesis of EM time series and live fluorescence imaging. eLife **11**:e77918. DOI: https://doi.org/10.7554/eLife.77918

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: an open-source platform for biological-image analysis. Nature Methods **9**:676–682. DOI: https://doi.org/10.1038/nmeth.2019

Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. 2011. Global quantification of mammalian gene expression control. Nature **473**:337–342. DOI: https://doi.org/10.1038/nature10098

Simmer F, Moorman C, van der Linden AM, Kuijk E, van den Berghe PVE, Kamath RS, Fraser AG, Ahringer J, Plasterk RHA. 2003. Genome-wide RNAi of C. elegans using the hypersensitive rrf-3 strain reveals novel gene functions. PLoS biology **1**:E12. DOI: https://doi.org/10.1371/journal.pbio.0000012, PMID: 14551910

Subach OM, Cranfill PJ, Davidson MW, Verkhusha VV. 2011. An Enhanced Monomeric Blue Fluorescent Protein with the High Chemical Stability of the Chromophore. PLOS ONE **6**:e28674. DOI: https://doi.org/10.1371/journal.pone.0028674

Taylor SR, Olson C, Ripoll-Sanchez L, Valperga G, McWhirter R, Barney ST, Atkinson A, Goel S, Weinreb A, Hardin A, Rolfson A, Pattee J, Aguilar GR, Merritt DM, Eroglu M, Majeed M, Grundvig E, Child E, Beets I, Vértes PE, Schafer WR, Varol E, Hammarlund M, Hobert O, Miller DM. 2025. A gene expression atlas of a juvenile nervous system. bioRxiv: The Preprint Server for Biology 2025.11.21.689793. DOI: https://doi.org/10.1101/2025.11.21.689793, PMID: 41332732

Taylor SR, Santpere G, Weinreb A, Barrett A, Reilly MB, Xu C, Varol E, Oikonomou P, Glenwinkel L, McWhirter R, Poff A, Basavaraju M, Rafi I, Yemini E, Cook SJ, Abrams A, Vidal B, Cros C, Tavazoie S, Sestan N, Hammarlund M, Hobert O, Miller DM. 2021. Molecular topography of an entire nervous system. Cell **184**:4329-4347.e23. DOI: https://doi.org/10.1016/j.cell.2021.06.023

Vergara HM, Pape C, Meechan KI, Zinchenko V, Genoud C, Wanner AA, Mutemi KN, Titze B, Templin RM, Bertucci PY, Simakov O, Dürichen W, Machado P, Savage EL, Schermelleh L, Schwab Y, Friedrich RW, Kreshuk A, Tischer C, Arendt D. 2021. Whole-body integration of gene expression and single-cell morphology. Cell **184**:4819-4837.e22. DOI: https://doi.org/10.1016/j.cell.2021.07.017, PMID: 34380046

Vogel C, Marcotte EM. 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nature Reviews Genetics **13**:227–232. DOI: https://doi.org/10.1038/nrg3185

Witvliet D, Mulcahy B, Mitchell JK, Meirovitch Y, Berger DR, Wu Y, Liu Y, Koh WX, Parvathala R, Holmyard D, Schalek RL, Shavit N, Chisholm AD, Lichtman JW, Samuel ADT, Zhen M. 2021. Connectomes across development reveal principles of brain maturation. Nature **596**:257–261. DOI: https://doi.org/10.1038/s41586-021-03778-8

Wu Y, Kumar A, Smith C, Ardiel E, Chandris P, Christensen R, Rey-Suarez I, Guo M, Vishwasrao HD, Chen J, Tang J, Upadhyaya A, La Riviere PJ, Shroff H. 2017. Reflective imaging improves spatiotemporal resolution and collection efficiency in light sheet microscopy. Nature Communications **8**:1452. DOI: https://doi.org/10.1038/s41467-017-01250-8

Yofe I, Weill U, Meurer M, Chuartzman S, Zalckvar E, Goldman O, Ben-Dor S, Schütze C, Wiedemann N, Knop M, Khmelinskii A, Schuldiner M. 2016. One library to make them all: streamlining the creation of yeast libraries via a SWAp-Tag strategy. Nature Methods **13**:371–378. DOI: https://doi.org/10.1038/nmeth.3795

Yu C-C (Jay), Orozco Cosio DM, Boyden ES. 2022. ExCel: Super-resolution imaging of C. elegans with expansion microscopy. Methods in molecular biology (Clifton,

N.J.) **2468**:141–203. DOI: https://doi.org/10.1007/978-1-0716-2181-3_9, PMID: 35320565

Zhang H, Papadaki S, Sun X, Wang X, Drobizhev M, Yao L, Rehbock M, Köster RW, Wu L, Namikawa K, Piatkevich KD. 2023. Quantitative assessment of near-infrared fluorescent proteins. Nature methods **20**:1605–1616. DOI: https://doi.org/10.1038/s41592-023-01975-z, PMID: 37666982

Figure captions

Figure 1: Pipeline for triple pooled CRISPR-Cas9 mediated gene tagging.
(A) Schematic of injection and screening strategies. FP, fluorescent protein.
(B) Excitation and emission spectra of selected fluorescent proteins used in the pilot study. Theoretical brightness of the fluorophores, calculated as a product of the extinction coefficient and quantum yield, is annotated as numbers adjacent to the respective emission spectra. The spectra and brightness values were adapted from FPbase (Lambert, 2019).
(C) Schematic of the practical visibility of fluorophores observed under a fluorescent stereomicroscope with different filter sets in live worms on NGM plates.

Figure 2: Selection of genes and predicted transcript expressions across tissues.
(A) Location of selected genes on *C. elegans* chromosomes. Numbers in parentheses indicate the injection pool the gene was assigned to. Color of text indicates the fluorophore assigned to the gene: blue, mTagBFP2; green, mStayGold; red, mScarlet3.
(B) Transcript expression levels of the selected genes across major tissues of *C. elegans* as measured by single cell RNA sequencing of adult hermaphrodites. Data was adapted from the CeNGEN adult hermaphrodite set (Taylor et al., 2021).

Figure 3: Successful tagging across loci, fluorophores and gene pools.
(A) Total number of targets successfully tagged. White bar denotes all 3 genes within the pool were successfully tagged in any combination. In four pools, all 3 targets were isolated in a single, triple-tagged individual worm. In one pool, we obtained all 3 targets but in doubles.
(B) Total insertion efficiencies for the indicated targets across injection pools. Dots, fraction of F1s positive for indicated fluorophore per injected P0. Five worms were injected per pool except for pools 4, 10 and 6 which were repeated after initial failures as shown in panel C. Data shown in panel B only represents the repeat runs and omits the original failed injections. Bars, mean ± SD.
(C) Viable brood size of injected P0 worms represented as individual points. Lines, mean ± SD. Numbers above plots, number of successfully isolated tags out of 3 within the pool. 4R, 6R and 10R are re-injections of groups 4, 6 and 10.
(D) Correlation of viable brood of P0s with fraction of F1 worms positive for any fluorophore tags. Goodness of fit is represented as an R^2 value from a simple linear regression. Slope was not significantly greater than zero (P = 0.2497, F-test on Prism).

Figure 4: Expression and localization of genes in pool 2.
Expressions and localizations of translation elongation factor 1 alpha 1 EEF-1A.1::mTagBFP2, nuclear protein 1 homolog ZK632.9::mStayGold and acyl-CoA dehydrogenase ACDH-10::mScarlet3 proteins in a live adult hermaphrodite worm by confocal microscopy. All three tags were isolated within the same individual then homozygosed at the F2 generation before imaging. Germline and intestine of the animal are delineated in magenta and white dashed lines, respectively. The worm was imaged

in separate overlapping frames with high magnification then stitched to reconstruct the whole animal. A single plane near the mid-plane of the animal was imaged.

Figure 5: Expression and localization of genes in pool 3.
Expressions and localizations of 14-3-3 protein PAR-5::mTagBFP2, heat shock protein family E Y22D7AL.10::mStayGold and the glucokinase HXK-1::mScarlet3 proteins in a live adult hermaphrodite worm by confocal microscopy. All three tags were isolated within the same individual then homozygosed at the F2 generation before imaging. Magenta and white dashed lines, germline and intestine respectively. Image reconstructed from multiple overlapping high magnification images. Single plane near the mid-plane of the animal.

Figure 6: Expression and localization of genes in pool 9.
(A) Expressions and localizations of glutamate dehydrogenase GDH-1::mTagBFP2, lysyl-tRNA synthetase KARS-1::mStayGold and the CREB binding protein CBP-1::mScarlet3 proteins in a live adult hermaphrodite worm by confocal microscopy. All three tags were isolated within the same individual then homozygosed at the F2 generation before imaging. Magenta and white dashed lines, germline and intestine respectively. Image reconstructed from multiple overlapping high magnification images. Single plane near the mid-plane of the animal.
(B) Germline insets showing non-overlapping localizations of GDH-1 and KARS-1, predicted mitochondrial proteins.
(C) Left, maximum entropy thresholds of GDH-1 and KARS-1 signals. Right, correlation plot of signal intensities of GDH-1 and KARS-1.

Figure 7: Expression and localization of genes in pool 10.
(A) Expressions and localizations of nucleosome assembly protein NAP-1::mTagBFP2, phosphoribosyl transferase HPRT-1::mStayGold, and ATP citrate lyase ACLY-1::mScarlet3 proteins in a live adult hermaphrodite worm by confocal microscopy. Magenta and white dashed lines, germline and intestine respectively. Image reconstructed from multiple overlapping high magnification images. Single plane near the mid-plane of the animal.
(B) Expressions and localizations of nucleosome assembly protein NAP-1::mTagBFP2, phosphoribosyl transferase HPRT-1::mStayGold, and ATP citrate lyase ACLY-1::mScarlet3 proteins in the head of a live adult hermaphrodite worm by confocal microscopy. Image shows a maximum intensity projection of a Z-stack optically sectioned at 1 µm per stack.

Figure 8: Expression and localization of genes in pool 5.
(A) Expressions and localizations of cytochrome C oxidase subunit 6B COX-6B::mTagBFP2 and TDO-2::mStayGold proteins in a live adult hermaphrodite worm by confocal microscopy. The two tags were isolated within the same individual then homozygosed at the F2 generation before imaging. Magenta and white dashed lines, germline and intestine respectively. Image reconstructed from multiple overlapping high magnification images. Single plane near the mid-plane of the animal.

(B) Expressions and localizations of tryptophan 2,3-dioxygenase TDO-2::mStayGold and hydroxyacyl-CoA dehydrogenase F54C8.1::mScarlet3 proteins in a live adult hermaphrodite worm by confocal microscopy. The two tags were isolated within the same individual then homozygosed at the F2 generation before imaging. Magenta and white dashed lines, germline and intestine respectively. Image reconstructed from multiple overlapping high magnification images. Single plane near the mid-plane of the animal.
(C) F54C8.1::mScarlet3 signal in sperm compared to untagged N2 worms imaged under the same settings.
(D) F54C8.1::mScarlet3 signal in gut granules compared to untagged N2 worms imaged under the same settings. N = 20 F54C8.1::mScarlet3 granules and 24 N2 granules quantified from 2 worms for each genotype. Lines, mean ± SD.

Figure 9: Expression and localization of genes in pool 7.
Expressions and localizations of heat shock protein family A HSP-1::mTagBFP2 and triosephosphate isomerase TPI-1::mStayGold proteins in a live adult hermaphrodite worm by confocal microscopy. The two tags were isolated within the same individual then homozygosed at the F2 generation before imaging. Magenta and white dashed lines, germline and intestine respectively. Image reconstructed from multiple overlapping high magnification images. Single plane near the mid-plane of the animal.

Figure 10: Expression and localization of genes in pool 8.
(A) Expressions and localizations of ATPase H+ transporting V0 subunit C VHA-2::mTagBFP2 and histone acetyltransferase HAT-1::mScarlet3 proteins in a live adult hermaphrodite worm by confocal microscopy. The two tags were isolated within the same individual then homozygosed at the F2 generation before imaging. Magenta and white dashed lines, germline and intestine respectively. Image reconstructed from multiple overlapping high magnification images. Single plane near the mid-plane of the animal.
(B) Inset showing germ cell nuclei. The proton pump VHA-2 localizes to nuclear membranes in germ cells specifically, indicated by white arrows.

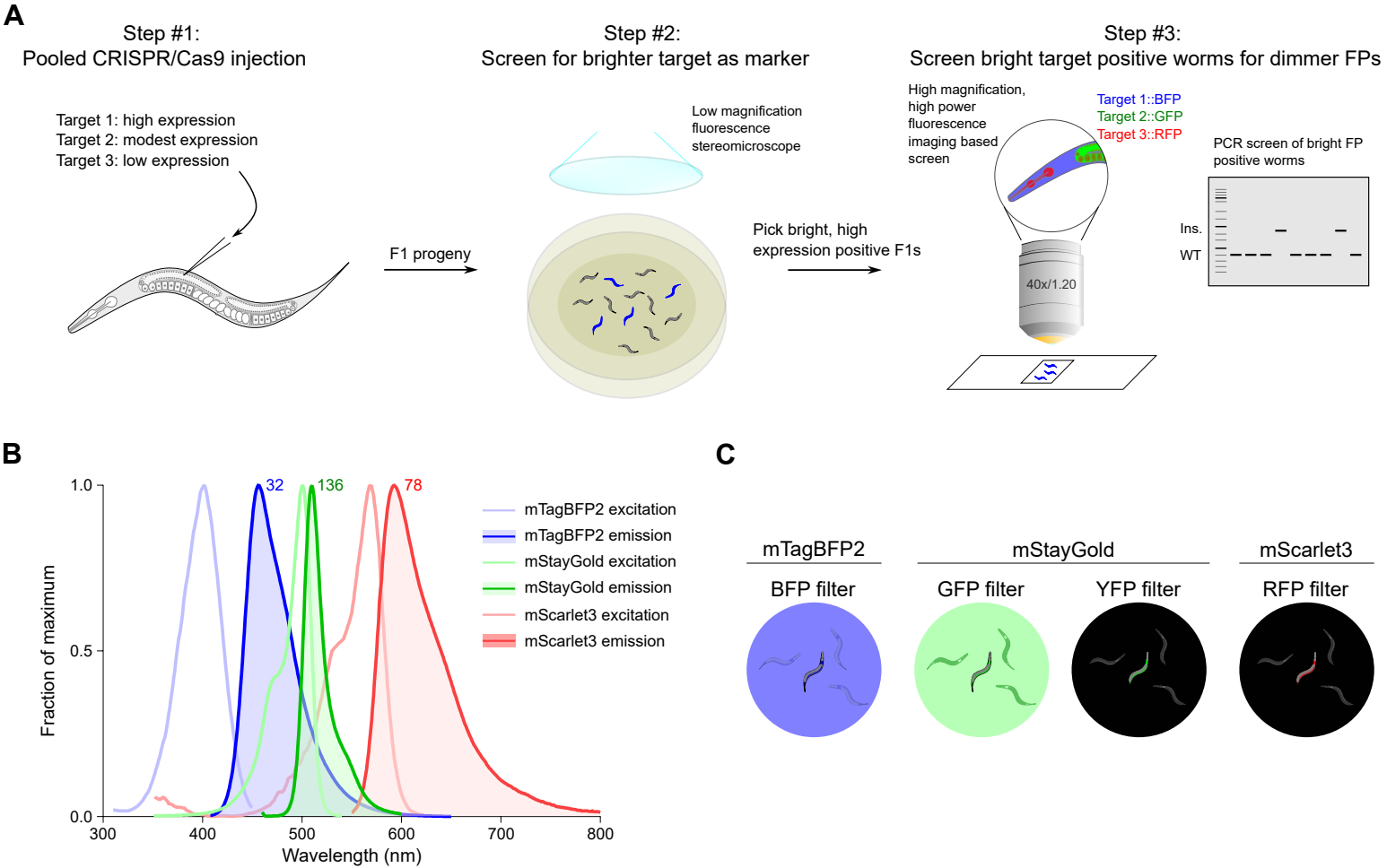

A
Step #1:
Pooled CRISPR/Cas9 injection
Target 1: high expression
Target 2: modest expression
Target 3: low expression
F1 progeny
Step #2:
Screen for brighter target as marker
Low magnification
fluorescence
stereomicroscope
Pick bright, high
expression positive F1s
Step #3:
Screen bright target positive worms for dimmer FPs
High magnification,
high power
fluorescence
imaging based
screen
Target 1::BFP
Target 2::GFP
Target 3::RFP
40x/1.20
PCR screen of bright FP
positive worms
Ins.
WT
B
Fraction of maximum
Wavelength (nm)
32
136
78
mTagBFP2 excitation
mTagBFP2 emission
mStayGold excitation
mStayGold emission
mScarlet3 excitation
mScarlet3 emission
C
mTagBFP2
mStayGold
mScarlet3
BFP filter
GFP filter
YFP filter
RFP filter

Figure 1

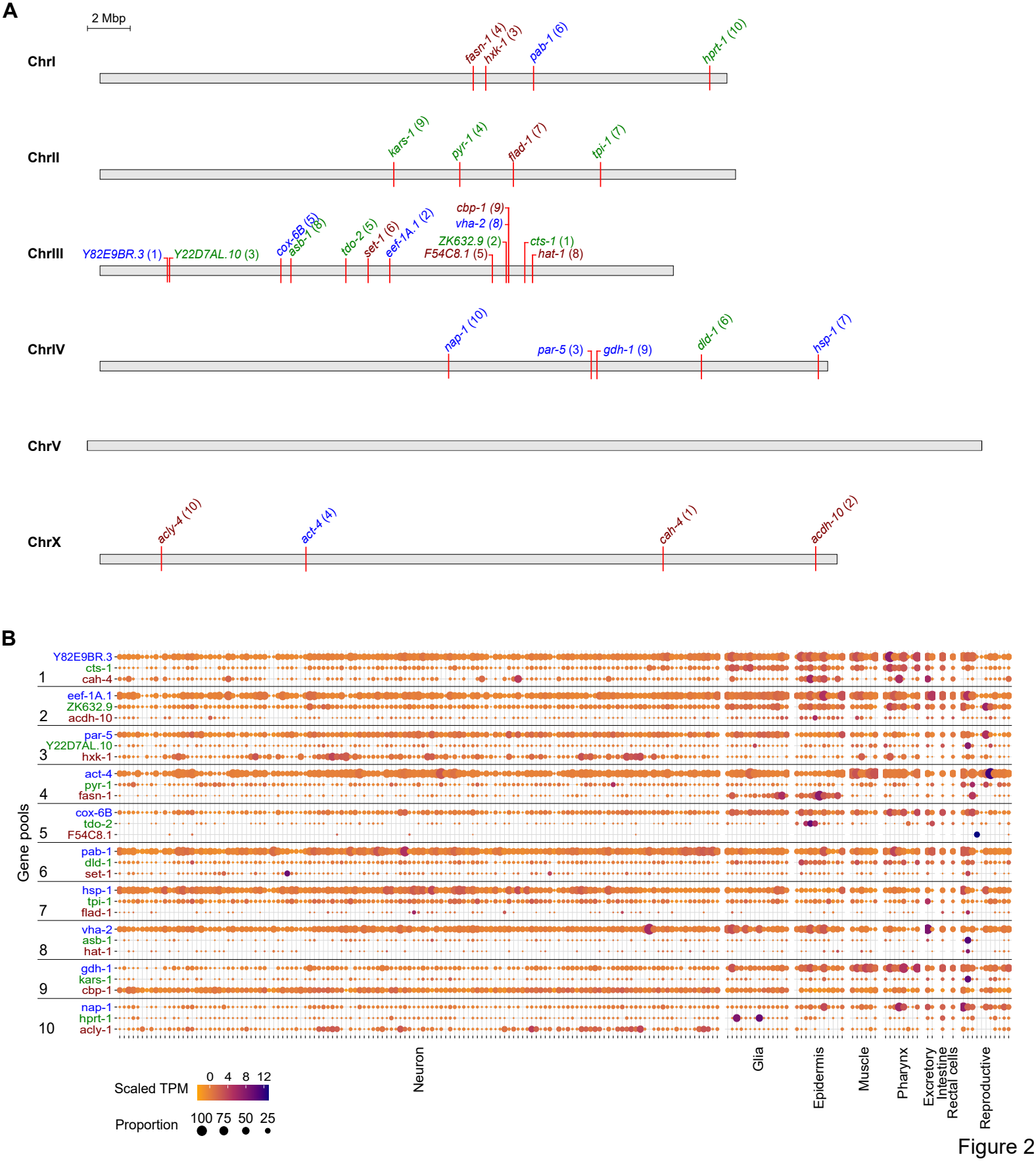

A
2 Mbp
ChrI
fasn-1 (4)
hxk-1 (3)
pab-1 (6)
hprt-1 (10)
ChrII
kars-1 (9)
pyr-1 (4)
flad-1 (7)
tpi-1 (7)
ChrIII
Y82E9BR.3 (1)
Y22D7AL.10 (3)
cox-6B (5)
asb-1 (8)
tdo-2 (5)
set-1 (6)
eef-1A.1 (2)
cbp-1 (9)
vha-2 (8)
ZK632.9 (2)
F54C8.1 (5)
cts-1 (1)
hat-1 (8)
ChrIV
nap-1 (10)
par-5 (3)
gdh-1 (9)
dld-1 (6)
hsp-1 (7)
ChrV
ChrX
acly-4 (10)
act-4 (4)
cah-4 (1)
acdh-10 (2)
B
Gene pools
1
Y82E9BR.3
cts-1
cah-4
2
eef-1A.1
ZK632.9
acdh-10
3
par-5
Y22D7AL.10
hxk-1
4
act-4
pyr-1
fasn-1
5
cox-6B
tdo-2
F54C8.1
6
pab-1
dld-1
set-1
7
hsp-1
tpi-1
flad-1
8
vha-2
asb-1
hat-1
9
gdh-1
kars-1
cbp-1
10
nap-1
hprt-1
acly-1
Neuron
Glia
Epidermis
Muscle
Pharynx
Excretory
Intestine
Rectal cells
Reproductive
Scaled TPM
0 4 8 12
Proportion
100 75 50 25

Figure 2

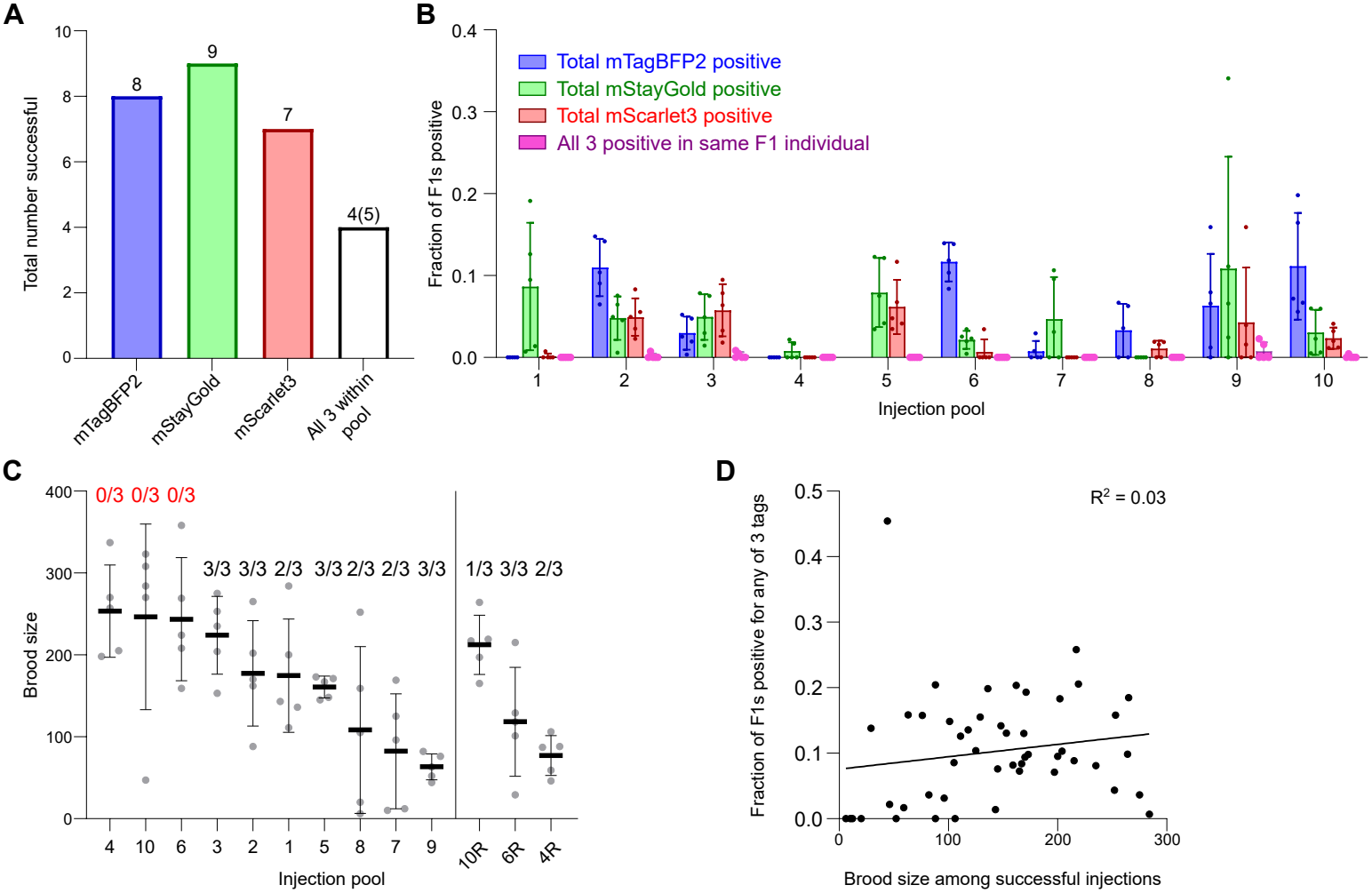

A
Total number successful
8
9
7
4(5)
mTagBFP2
mStayGold
mScarlet3
All 3 within pool
B
Fraction of F1s positive
Total mTagBFP2 positive
Total mStayGold positive
Total mScarlet3 positive
All 3 positive in same F1 individual
Injection pool
C
Brood size
0/3 0/3 0/3
3/3 3/3 2/3 3/3 2/3 2/3 3/3
1/3 3/3 2/3
Injection pool
D
Fraction of F1s positive for any of 3 tags
R² = 0.03
Brood size among successful injections

Figure 3

Group 2: *eef-1A.1*::mTagBFP-2 mStayGold::*ZK632.9* (III); *acdh-10*::mScarlet3 (X)

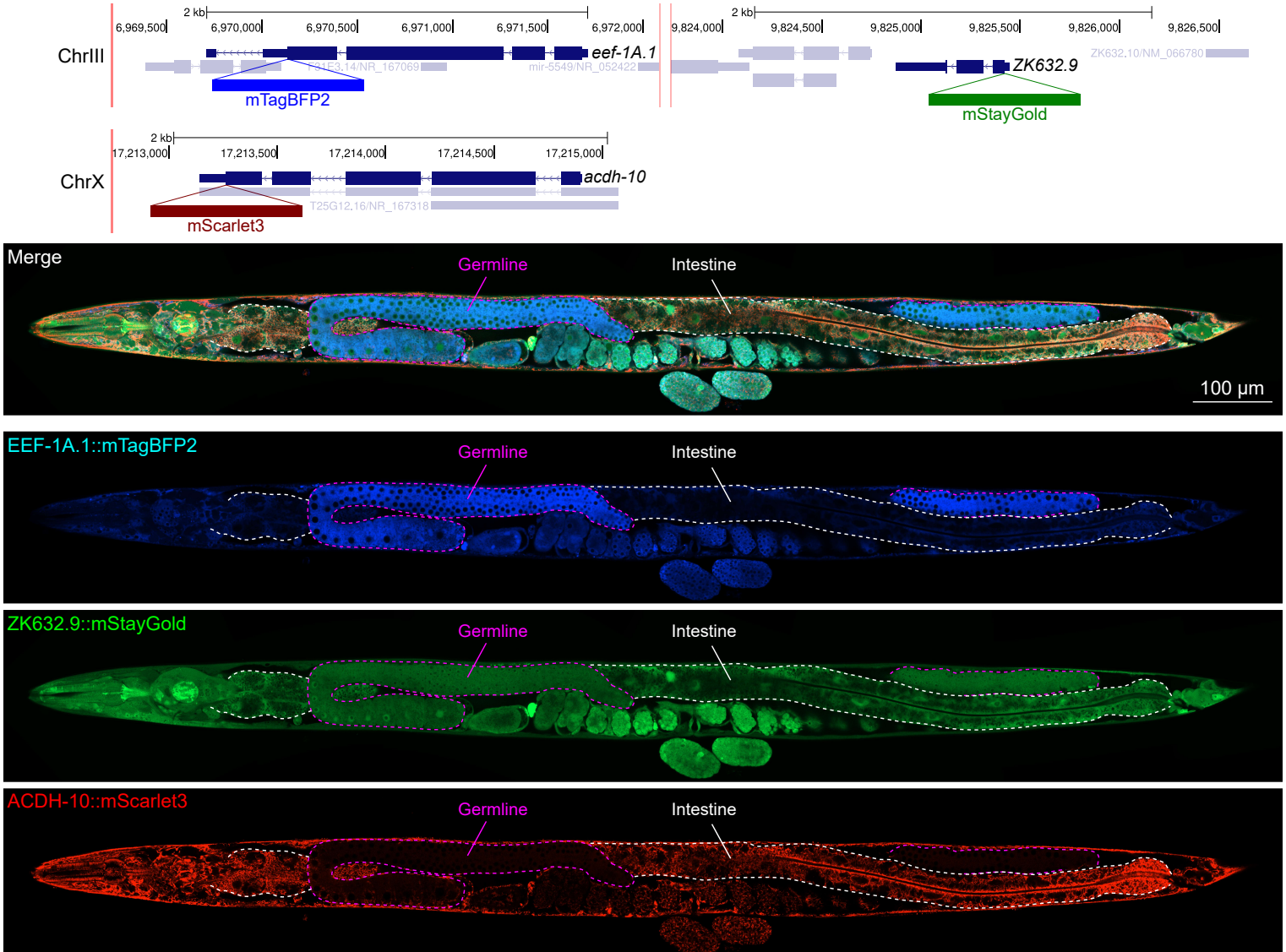


Figure 4

Group 3: *hxk-1*::mScarlet3 (I); *Y22D7AL.10*::mStayGold (III); *par-5*::mTagBFP2 (IV)

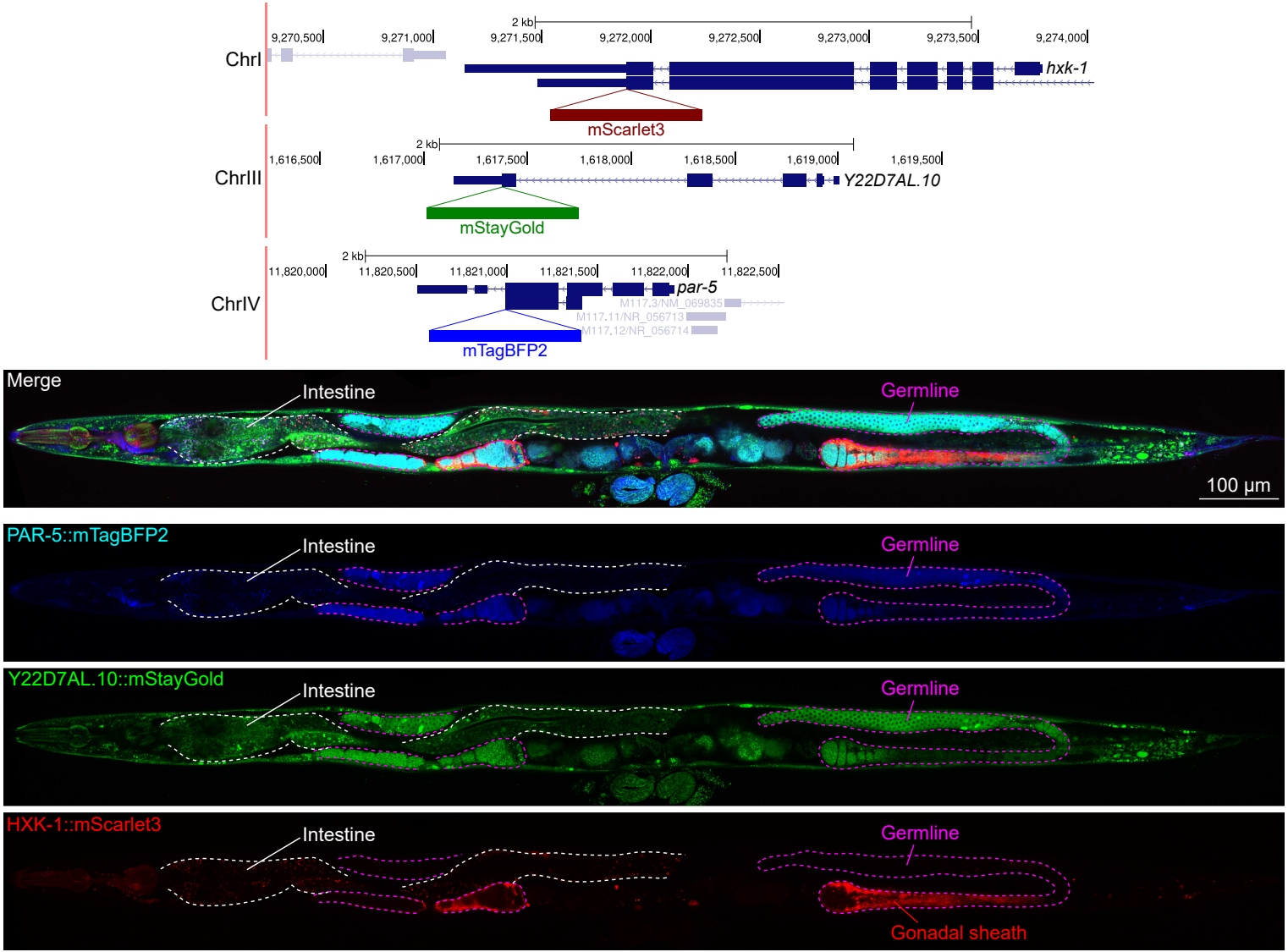
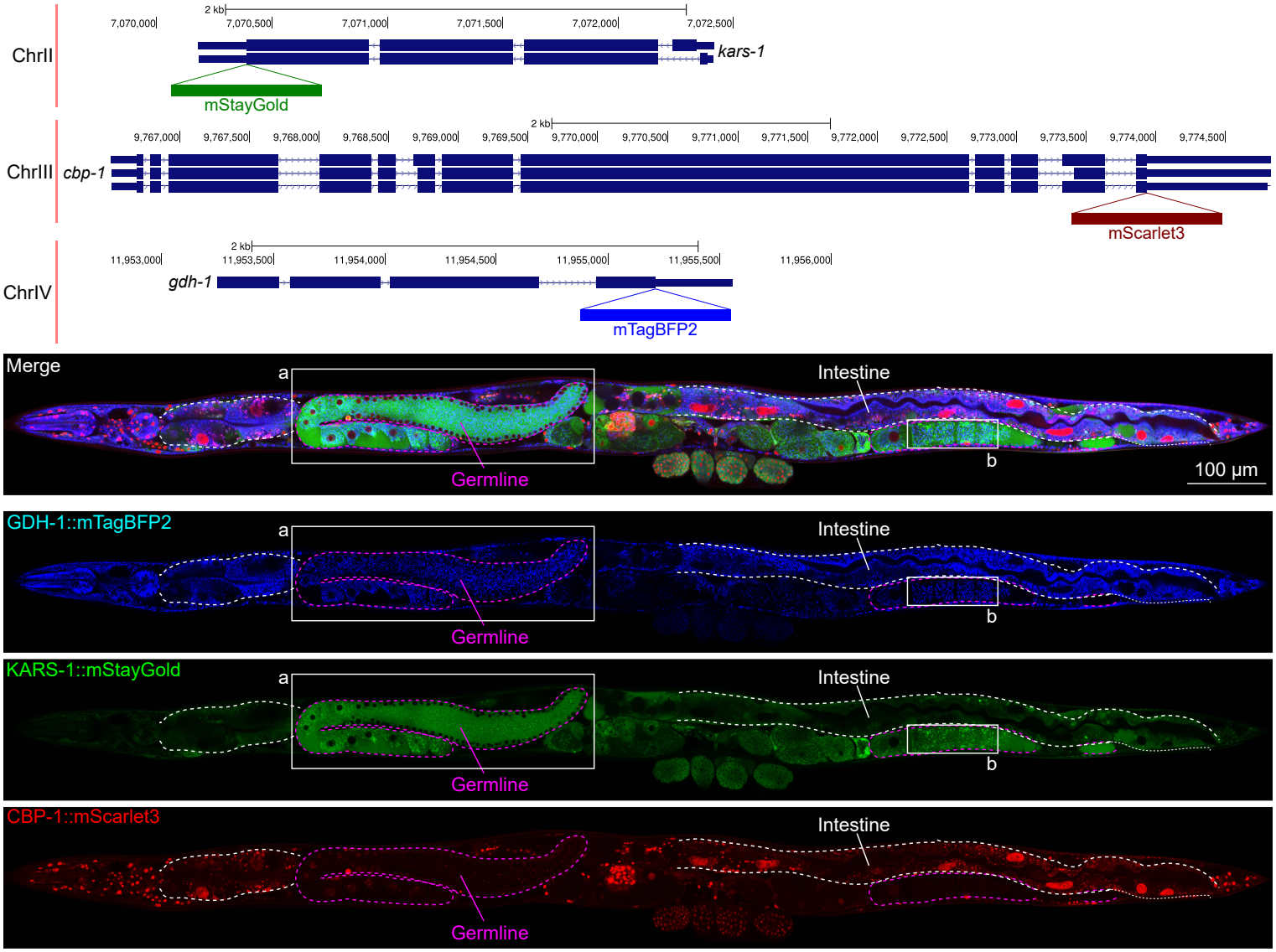


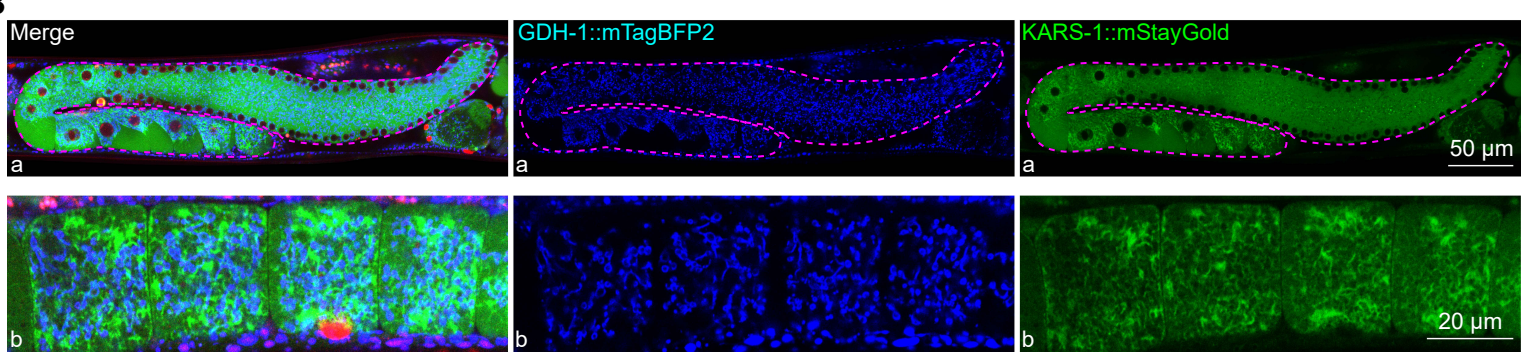
Figure 5

A

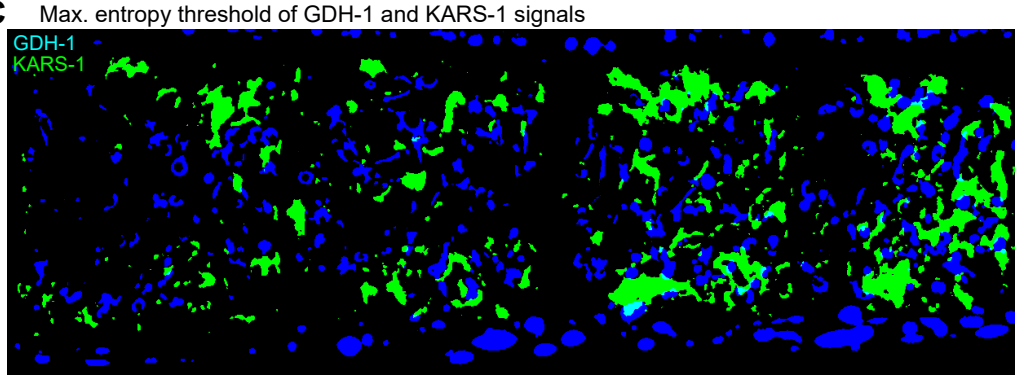
Group 9: *kars-1*::mStayGold (II); *cbp-1*::mScarlet3 (III); *gdh-1*::mTagBFP2 (IV)



B



C Max. entropy threshold of GDH-1 and KARS-1 signals



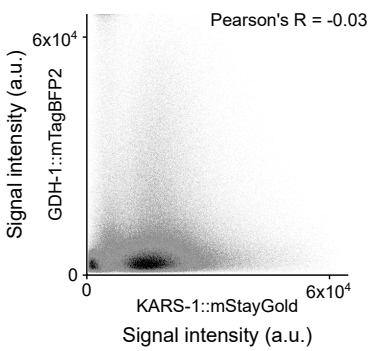
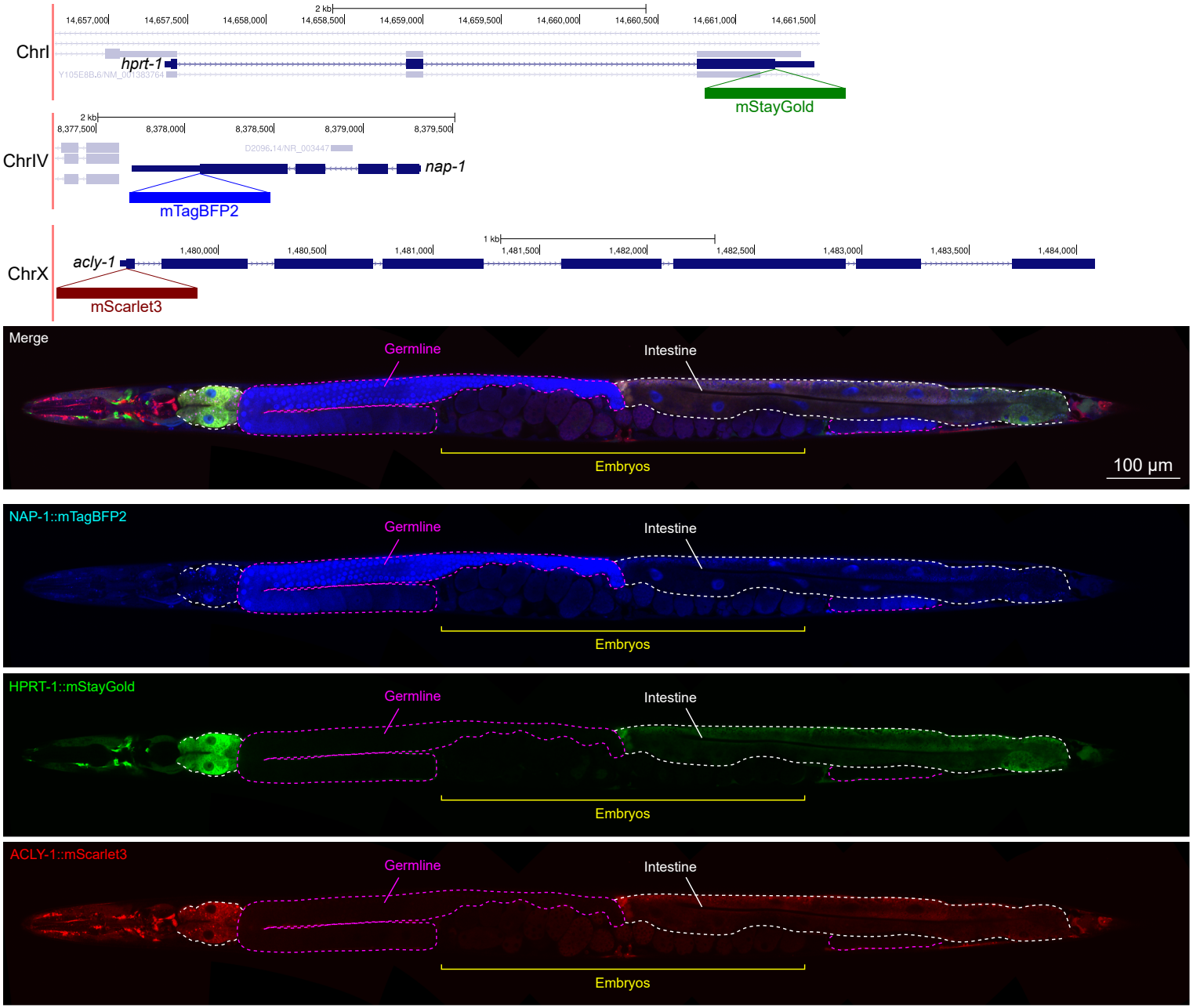


Figure 6

A

Group 10: *hprt-1*::mStayGold (I); *nap-1*::mTagBFP2 (IV); mScarlet3::*acly-1* (X)



B

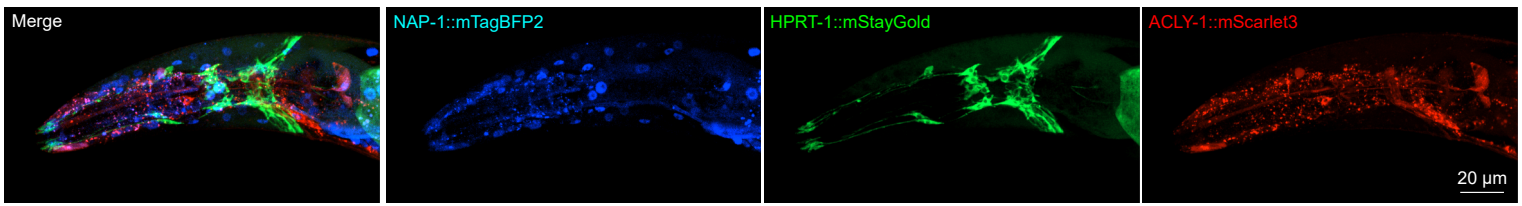


Figure 7

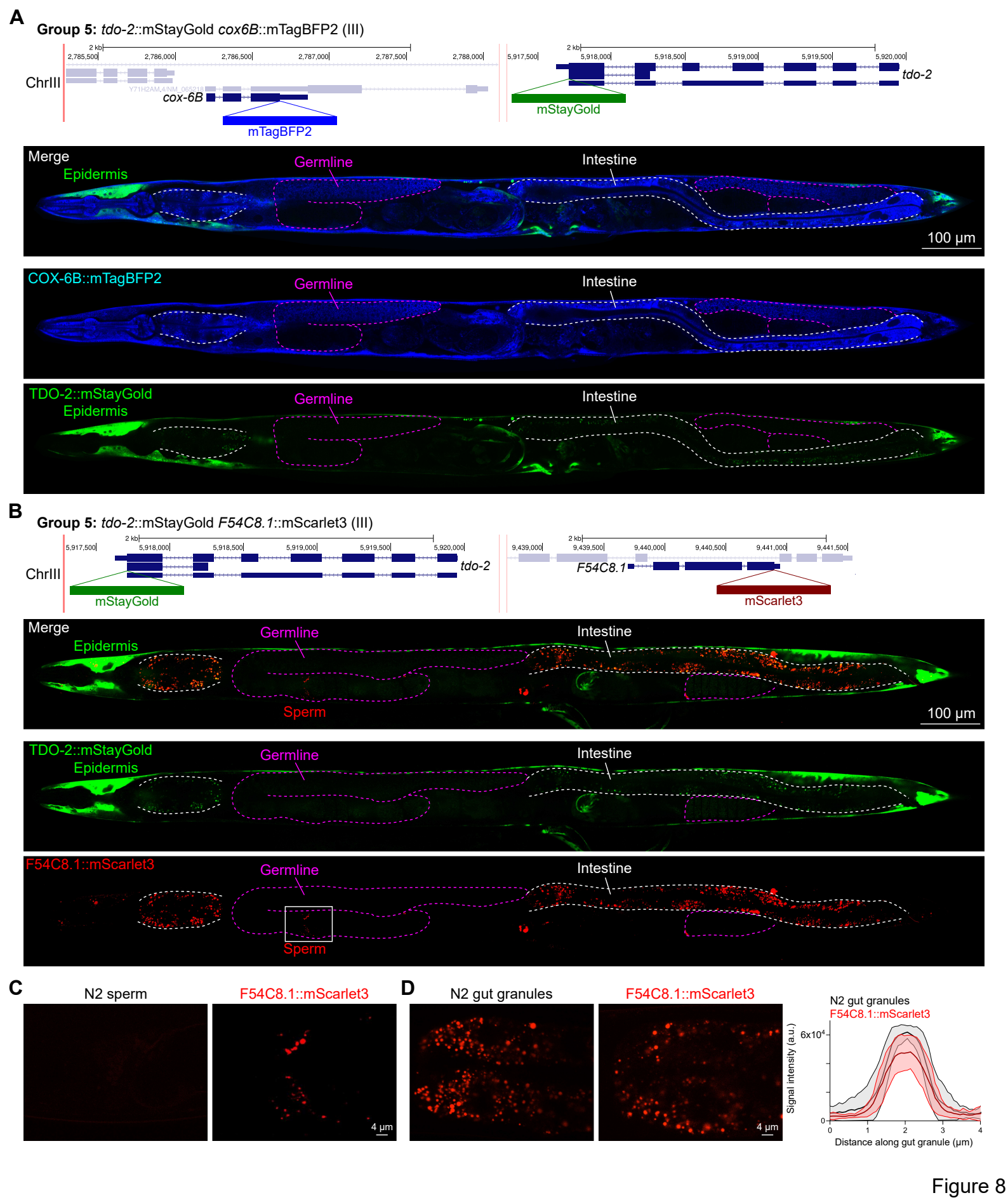

A
Group 5: tdo-2::mStayGold cox6B::mTagBFP2 (III)
ChrIII
cox-6B
mTagBFP2
tdo-2
mStayGold
Merge
Epidermis
Germline
Intestine
100 µm
COX-6B::mTagBFP2
TDO-2::mStayGold
B
Group 5: tdo-2::mStayGold F54C8.1::mScarlet3 (III)
F54C8.1
mScarlet3
Sperm
F54C8.1::mScarlet3
C
N2 sperm
4 µm
D
N2 gut granules
Signal intensity (a.u.)
Distance along gut granule (µm)

Figure 8

Group 7: *tpi-1*::mStayGold *flad-1*::mScarlet3 (II); *hsp-1*::mTagBFP2 (IV)

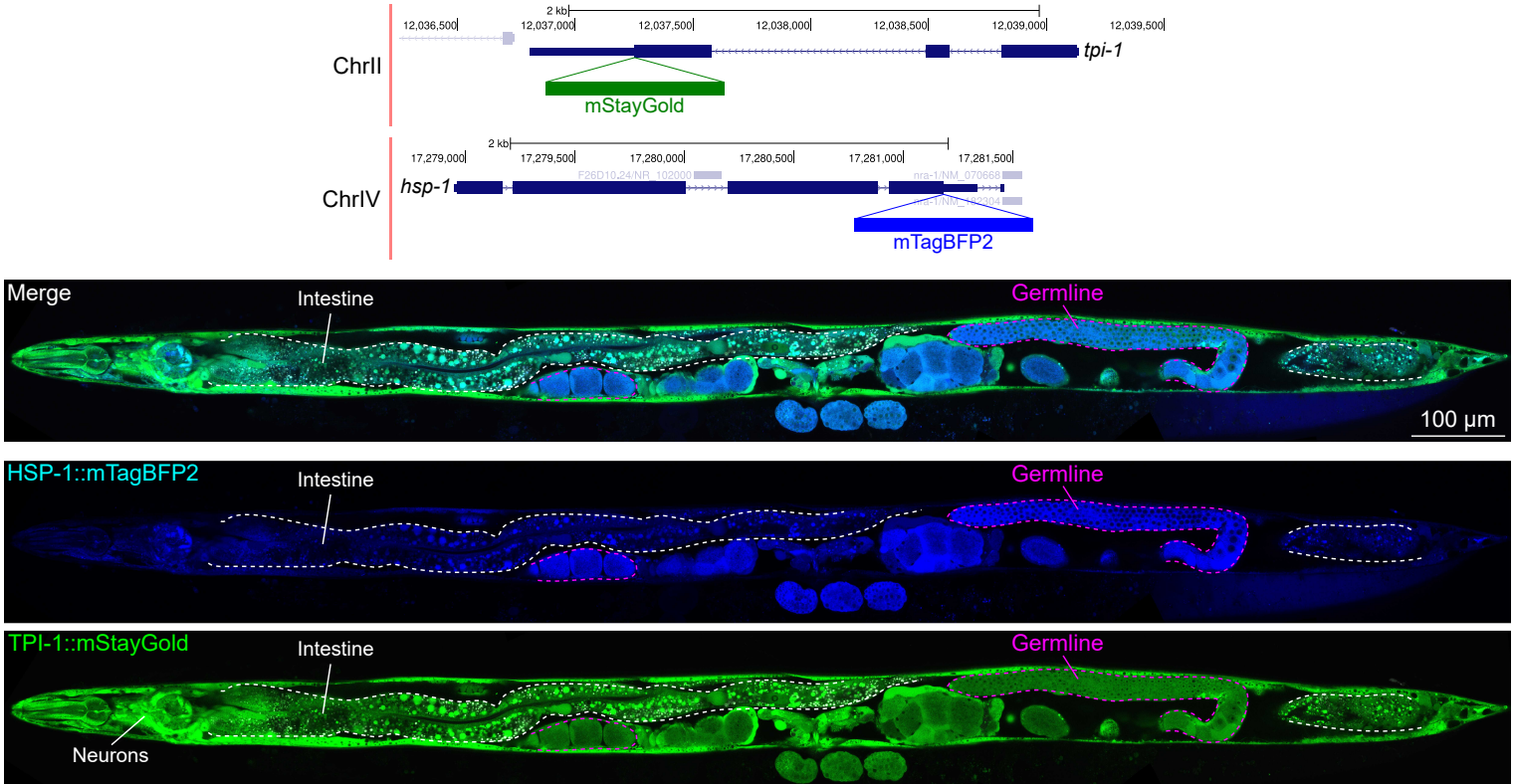


Figure 9

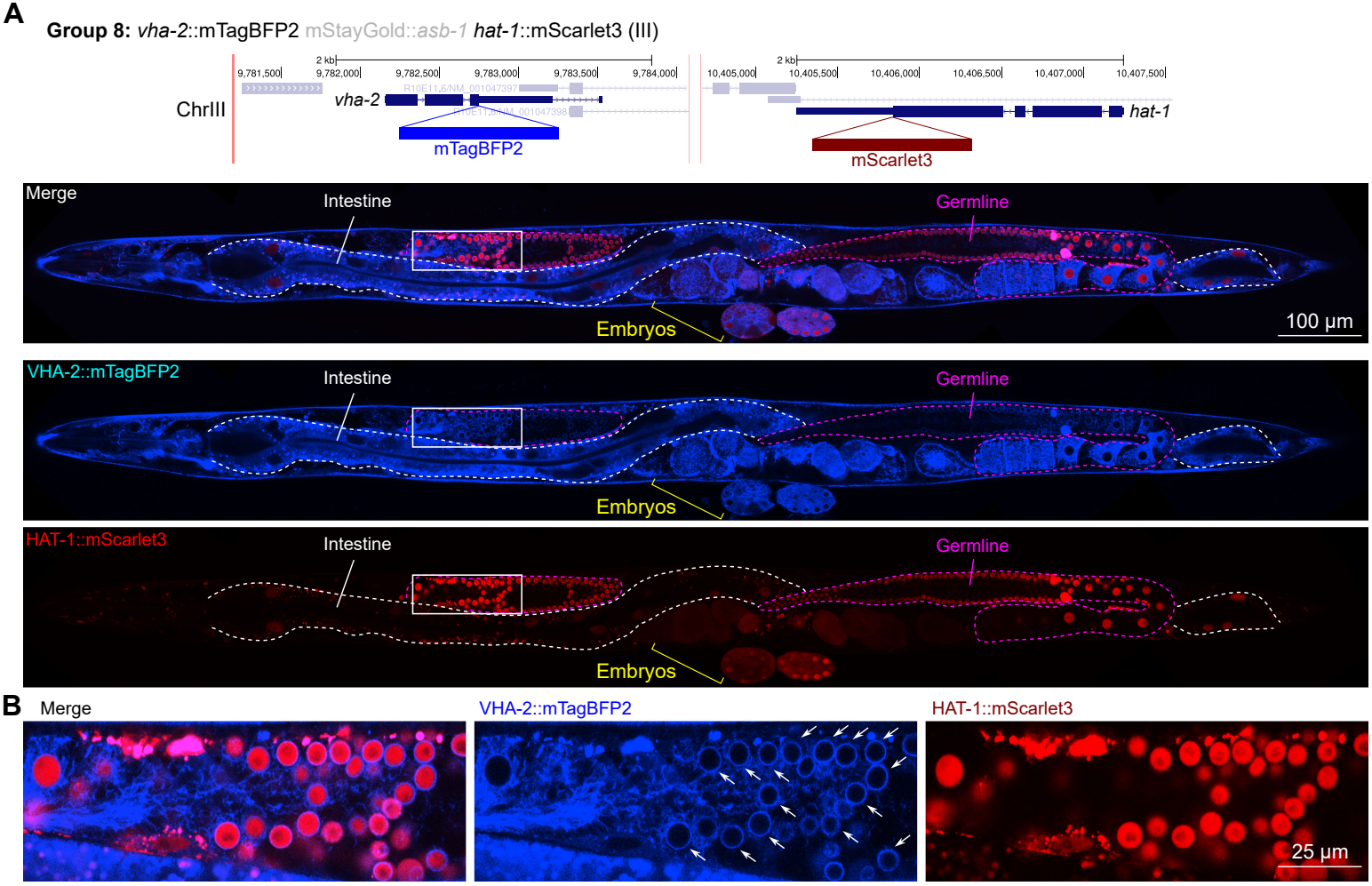

A
Group 8: vha-2::mTagBFP2 mStayGold::asb-1 hat-1::mScarlet3 (III)
ChrIII
vha-2
mTagBFP2
hat-1
mScarlet3
Merge
Intestine
Germline
Embryos
100 µm
VHA-2::mTagBFP2
HAT-1::mScarlet3
B
Merge
VHA-2::mTagBFP2
HAT-1::mScarlet3
25 µm

Figure 10

Table 1: Genes tagged

Group	Gene	Chr	Class	Function	WormBase L4 Fl	Terminus	Tag	Success?	RNAi Phen	Tag dissecting visible?	Rationale
1	*Y82E9BR.3*	III	ATP synthase membrane subunit c	ATP synthesis & ETC	7295.1	N	mTBFP2	No	Let		Highest expressed ETC component
	cts-1	III	citrate synthase	Citrate cycle TCA cycle	465.7	C	mSG	Yes	Let	Yes	Highest expressed citrate cycle
	cah-4	X	carbonic anhydrase	Nitrogen metabolism	104.8	C	mSc3	Yes	Let	Yes	Highest expressed nitrogen fixation
2	*eef-1A.1*	III	elongation factor 1 alpha 1	Translation	6059	C	mTBFP2	Yes	Let	Yes	Highest expressed translation elongation factor
	ZK632.9	III	Nuclear protein 1	Nucleus	452.4	N	mSG	Yes	Stp	Yes	Among highest expressed pred. nuclear proteins, uncharacterized
	acdh-10	X	acyl-CoA dehydrogenase	Metabolism	73.5	C	mSc3	Yes	Ste	Yes	Fatty acid metabolism, Acyl CoA DeHydrogenase
3	*par-5*	IV	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein	Signaling	2121.8	C	mTBFP2	Yes	Let	Yes	Highest expressed 14-3-3 gene
	Y22D7AL.10	III	heat shock protein family E	Proteostasis	363.5	C	mSG	Yes	Let	Yes	Uncharacterized, highly expressed heat shock protein
	hxk-1	I	glucokinase	Glycolysis / Gluconeogenesis	46.3	C	mSc3	Yes	Let	Yes	Initial enzyme of glycolysis (glucose --> g6p)
4	*act-4*	X	actin beta	Cytoskeleton	1661.6	C	mTBFP2	No	Let		Highest expressed actin
	pyr-1	II	carbamoyl-phosphate synthetase 2, aspartate transcarbamylase, and dihydroorotase	Pyrimidine metabolism	58.9	C	mSG	Yes	Let	Yes	Initial enzyme of pyrimidine synthesis
	fasn-1	I	fatty acid synthase	Fatty acid biosynthesis	38.1	C	mSc3	No	Let		Converts acetyl CoA to fatty acids
5	*cox-6b*	III	cytochrome c oxidase subunit 6B	ATP synthesis & ETC	1650.1	C	mTBFP2	Yes	Let	Dim F1, F2 yes	Highest expressed cytochrome c oxidase
	tdo-2	III	tryptophan 2,3-dioxygenase	W degradation	201.3	C	mSG	Yes	Gro	Yes	Tryptophan degradation
	F54C8.1	III	hydroxyacyl-CoA dehydrogenase	Fatty acid elongation	34.2	C	mSc3	Yes	Let	Yes	Uncharacterized
6	*pab-1*	I	poly(A) binding protein	Transcript reg	1017.1	C	mTBFP2	Yes	Let	Yes	Highest expressed PAB
	dld-1	IV	dihydrolipoamide dehydrogenase	Glycolysis / Gluconeogenesis	199.7	C	mSG	Yes	Let	Yes	Mitochondrial acetyl coa synthesis from pyruvate
	set-1	III	lysine methyltransferase	Histone modification	27.1	C	mSc3	No	Let		Highest expressed histone modifier
7	*hsp-1*	IV	heat shock protein family A	Proteostasis	998.7	C	mTBFP2	Yes	Let	Yes	Highest expressed heat shock protein
	tpi-1	II	triosephosphate isomerase	Glycolysis / Gluconeogenesis	180.1	C	mSG	Yes		Yes	Glycolysis, cytosolic
	flad-1	II	flavin adenine dinucleotide synthetase	Riboflavin metabolism	23.7	Internal, exo	mSc3	No	Gro		Only FAD synthetase (cofactor for histone demethylation by LSD/KDM family)
8	*vha-2*	III	ATPase H+ transporting V0 subunit c	Proton transport	732.3	C	mTBFP2	Yes	Let	Yes	Highest expressed V-ATPase
	asb-1	III	ATP synthase peripheral stalk-membrane subunit b	Oxidative phosphorylation	157.1	N	mSG	No	Let		ATP synthase
	hat-1	III	histone acetyltransferase 1	Histone modification	23.3	C	mSc3	Yes	Let	Dim	Histone acetyltransferase
9	*gdh-1*	IV	glutamate dehydrogenase	Glutamate / glutamine metabolism	565.2	C	mTBFP2	Yes	Ste	Yes	akg synthesis, (cofactor for histone demethylation by JmjC family)
	kars-1	II	lysyl-tRNA synthetase	Aminoacyl-tRNA biosynthesis	153.4	C	mSG	Yes	Let	Yes	Highest expressed aminoacyl tRNA synthetase
	cbp-1	III	CREB binding protein	Histone modification	22.7	C	mSc3	Yes	Let	Dim	CREBBP/p300, broad transcriptional activator, ubiquitous
10	*nap-1*	IV	nucleosome assembly protein	Nucleus	500.6	C	mTBFP2	Yes	Let	Yes	Uncharacterized, predicted role in nucleosome assembly
	hprt-1	I	phosphoribosyl transferase	Purine metabolism	101	C	mSG	Yes		Yes	Purine salvage pathway, IMP synthesis
	acly-1	X	ATP citrate lyase	Citrate cycle TCA cycle	18	N	mSc3	Yes	Let	Yes	Citrate to acetyl CoA and OAA