### Title

In vivo mapping of tissue- and subcellular-specific proteomes in Caenorhabditis elegans.




### Journal

Science Advances, 3(5)



### Authors

Reinke, Aaron
Mak, Raymond
Troemel, Emily
et al.



### Publication Date

2017-05-01

### DOI

10.1126/sciadv.1602426

Peer reviewed

PROTEOMICS

# In vivo mapping of tissue- and subcellular-specific proteomes in *Caenorhabditis elegans*

**Aaron W. Reinke,* Raymond Mak, Emily R. Troemel, Eric J. Bennett**



**Multicellular organisms are composed of tissues that have distinct functions requiring specialized proteomes. To define the proteome of a live animal with tissue and subcellular resolution, we adapted a localized proteomics technology for use in the multicellular model organism *Caenorhabditis elegans*. This approach couples tissue- and location-specific expression of the enzyme ascorbate peroxidase (APX), which enables proximity-based protein labeling in vivo, and quantitative proteomics to identify tissue- and subcellular-restricted proteomes. We identified and localized more than 3000 proteins from strains of *C. elegans* expressing APX in either the nucleus or cytoplasm of the intestine, epidermis, body wall muscle, or pharyngeal muscle. We also identified several hundred proteins that were specifically localized to one of the four tissues analyzed or specifically localized to the cytoplasm or the nucleus. This approach resulted in the identification both of proteins with previously characterized localizations and of those not known to localize to the nucleus or cytoplasm. Further, we confirmed the tissue- and subcellular-specific localization of a subset of identified proteins using green fluorescent protein tagging and fluorescence microscopy, validating our in vivo proximity-based proteomics technique. Together, these results demonstrate a new approach that enables the tissue- and subcellular-specific identification and quantification of proteins within a live animal.**

## INTRODUCTION

Animal development and function rely on the coordinated expression of proteins in specific tissues and the correct localization of those proteins to specific subcellular compartments. Understanding both the tissue and the subcellular localization of a protein can be critical to revealing its function. Because of the fundamental importance of understanding protein localization, several experimental approaches have been established to globally define tissue- and compartment-specific protein expression. Comparisons among different existing approaches to measure protein localization in animals are summarized in table S1 and described in detail below.

One widely used approach for determining protein localization relies on generating fluorescent protein fusions with proteins of interest and analyzing protein localization using microscopy. A seminal study using the single-celled yeast *Saccharomyces cerevisiae* determined the subcellular localization of most of the proteome (*1*). This fluorescent tagging approach has been subsequently applied at the genome scale to the bacteria *Escherichia coli* and *Caulobacter crescentus* (*2*–*4*). Because of the cellular complexity in animals, global determination of protein localization has been more challenging. Genome-wide fluorescent tagging approaches have been initiated to analyze protein localization in the animals *Caenorhabditis elegans* and *Drosophila melanogaster*. These efforts are impressive; but because of the large number of proteins and difficulty in generating transgenic animals, the most comprehensive attempts thus far have only localized 1 to 2% of the interrogated proteome (*5*, *6*).

Another approach to define tissue- and subcellular-specific proteomes relies on biochemical isolation of tissues, followed by mass spectrometry to identify proteins. This approach has been widely used to define large-scale tissue maps for the human and mouse proteomes based on dissection of specific organs, followed by mass spectrometry analysis (*7*, *8*), but such studies lack cellular resolution within tissues, which can be composed of multiple different cell types (*9*). In addition, these techniques are difficult to perform with small organisms where tissue dissection is challenging, such as with *C. elegans*. Cellular compartment-specific proteomes have also been generated (*10*, *11*), but biochemical isolation techniques can result in a loss of integrity of isolated subcellular compartments, leading to incomplete or inaccurate spatial information (*12*).

To address the drawbacks associated with tissue dissection, a class of approaches has been used to define protein expression in the individual tissues and cells of live animals. These approaches rely on the tissue-specific expression of a modified transfer RNA synthetase that selectively incorporates unnatural amino acids with chemical handles into the proteome. The labeled proteins can then be purified, and mass spectrometry can be used to identify proteins from specific tissues. Several studies have successfully used this approach in *C. elegans* and *D. melanogaster* (*13*, *14*). Although these are a promising class of approaches, they can have limited sensitivity because the incorporation of the unnatural amino acid is reported to be less than 1% per codon (*14*). These approaches also lack the ability to detect protein subcellular localization directly.

Recently, a proteomic technique has been developed that enables subcellular protein localization by labeling proteins in discrete locations of live cells (without the need for biochemical fractionation). Known as spatially restricted enzymatic tagging, this method allows for proteins in specific cellular compartments to be tagged with a chemical handle in vivo (*15*). This approach relies on the localized expression of soybean ascorbate peroxidase (APX), which, in the presence of hydrogen peroxide ($H_2O_2$) and biotin-phenol, catalyzes the formation of a phenolic radical that can covalently modify proximal proteins with biotin. These proteins can then be isolated with streptavidin beads and then identified and quantified using mass spectrometry. The efficacy of this method was demonstrated in human cells, where APX was shown to be active in a large number of subcellular compartments (*15*). This approach was also recently used to identify proteins localized to the mitochondria in dissected fly tissues (*12*).

Division of Biological Sciences, Section of Cell and Developmental Biology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA.
*Corresponding author. Email: awreinke@gmail.com; areinke@ucsd.edu

These studies demonstrate the potential of this technique for identification of proteins in live animals.

The nematode *C. elegans* offers an ideal system for the application of spatially restricted enzymatic tagging in a live animal. *C. elegans* is a simple multicellular model organism that has only 959 cells organized into conserved tissues such as muscle and intestine (*16*). *C. elegans* has provided the basis for fundamental discoveries in signaling, development, and neurobiology but is lacking a global description of protein localization in its various tissues (*17*). To generate a tissue- and subcellular-specific map of protein localization in *C. elegans*, we expressed APX localized to two subcellular compartments (cytoplasm and nucleus) in each of four tissues (intestine, epidermis, body wall muscle, and pharyngeal muscle). Subsequent isolation of biotinylated proteins and identification by quantitative mass spectrometry allowed us to quantitatively compare proteins detected in the cytoplasm and nucleus within each tissue to provide a catalog of protein expression specific to either subcellular compartment within specific tissues. Together, these results demonstrate a global approach to characterize tissue- and compartment-specific proteomes in vivo.

## RESULTS

### Development of spatially restricted enzymatic tagging in the intestine of *C. elegans*

To develop a system for identifying proteins with a high degree of temporal, tissue, and subcellular resolution in a live animal, we adapted the use of spatially restricted enzymatic tagging to the nematode *C. elegans*, focusing first on the intestine (Fig. 1). We first generated transgenic animals that express APX as a single genomic copy in *C. elegans* using the MosSCI method (*18*). The enzyme was fused to green fluorescent protein (GFP) for visualization, localized to the cytoplasm using a nuclear export signal (NES), and specifically expressed in the intestine using the *spp-5* promoter (table S2). We confirmed the intestinal cytoplasm-specific expression of this strain, GFP-APX-NES, by fluorescence microscopy (fig. S1). As a negative control, we also generated a strain in a similar manner that expresses GFP in the intestine without the APX enzyme.

To test the activity of APX in *C. elegans*, we grew a single plate of 30,000 synchronized animals on bacteria until the fourth larval (L4) stage (44 hours) at 20°C. Populations of animals were grown for both the strain expressing APX and the negative control strain expressing GFP without APX. These animals were then washed off the plate and treated with the biotin-phenol substrate for 1 hour, followed by the addition of $H_2O_2$ for 2 min. The labeling reaction was then quenched, and proteins from these animals were extracted and purified with streptavidin beads. Proteins were then eluted from the beads, separated on SDS–polyacrylamide gel electrophoresis gels, and visualized with Oriole staining. In the absence of APX expression, endogenously biotinylated proteins were detected (*19*). The presence of APX did not result in the expected increase of biotinylated proteins compared to the control GFP-only–expressing animals (Fig. 2A).

On the basis of a lack of APX-mediated biotin labeling and our observation that APX was properly expressed, we hypothesized that the concentration of the biotin-phenol substrate in the worm intestinal cells was inadequate to facilitate efficient protein labeling. To investigate whether labeling efficiency could be improved by increasing the permeability of the *C. elegans* cuticle to biotin-phenol, we knocked down the expression of the *bus-8* gene. BUS-8 is a glycosyltransferase involved in cuticle development, and reduction of BUS-8 function has been shown to increase small-molecule permeability by decreasing cuticle integrity (*20*). Therefore, we fed animals bacteria that express double-stranded RNA against *bus-8* to induce RNA interference (RNAi) knockdown of *bus-8* expression. Under these conditions, APX-expressing animals treated with the biotin-phenol substrate displayed increased biotin tagging of proteins compared to control animals that do not express APX (Fig. 2A). Therefore, decreasing cuticle integrity appears to increase the availability of the substrate, leading to increased biotinylation of cellular proteins by APX. Because of this substantial increase in biotinylation efficiency, we performed all subsequent experiments by growing animals on bacteria expressing the *bus-8* RNAi clone to increase cuticle permeability and APX-mediated protein biotinylation.

We then investigated whether the biotinylation reaction detailed above depends on the previously described components for an APX-mediated reaction: the APX enzyme, the biotin-phenol, and $H_2O_2$. First, we observed a slight increase in background biotinylation when the biotin-phenol was added to the control animals that do not express APX (Fig. 2B), indicating that a low level of biotinylation occurs without the enzyme. However, there was a substantial increase in labeling when animals expressing APX were exposed to biotin-phenol and $H_2O_2$ (Fig. 2B), demonstrating that efficient biotinylation is greatly potentiated by APX expression. We also found that both biotin-phenol and $H_2O_2$ were required for efficient labeling (Fig. 2B), consistent with APX-mediated biotinylation in human cells also being dependent on $H_2O_2$. Notably, in our experiments, animals are incubated with biotin-phenol for 1 hour, whereas $H_2O_2$ is only added for a period of 2 min before it is quenched. Because $H_2O_2$ is required for labeling, this result indicates that the labeling reaction is rapid and occurs within only 2 min in *C. elegans*.

### Biotin labeling of proteins in specific locations within *C. elegans*

To determine whether protein biotinylation and analysis could be performed in other tissues and compartments in *C. elegans*, we created a panel of strains expressing the APX enzyme in different tissues and subcellular locations. In addition to a version of the protein localized to the cytoplasm, we created another version where APX is localized to the nucleus using a nuclear localization signal (NLS). In addition to the intestine, we expressed the enzyme in three other tissues using the following tissue-specific promoters: epidermis (*dpy-7*), body wall muscle (*myo-3*), and pharyngeal muscle (*myo-2*) (table S2). In total, we generated eight strains expressing the APX enzyme. This panel of strains and the negative control strain were grown and treated with biotin-phenol and $H_2O_2$ to label proteins as described above. The APX-mediated biotinylation of proteins in each of these tissues was examined, revealing a clear increase in biotinylation in every location compared to the negative control strain, with the exception of the nuclear-localized enzyme in the pharyngeal muscle (Fig. 2C). Of the tissues analyzed in our panel, this location represents the smallest tissue, and thus it is likely that labeled proteins could not be detected over background biotinylation levels.

To confirm the location specificity of the labeling, we used fluorescence microscopy to visualize where proteins were biotinylated within the animal. After treatment of the animals with biotin-phenol and $H_2O_2$, we fixed and stained the animals with anti-GFP antibodies to localize the fusion protein and with fluorescent streptavidin to localize biotinylated proteins. Fluorescence microscopy was used to analyze

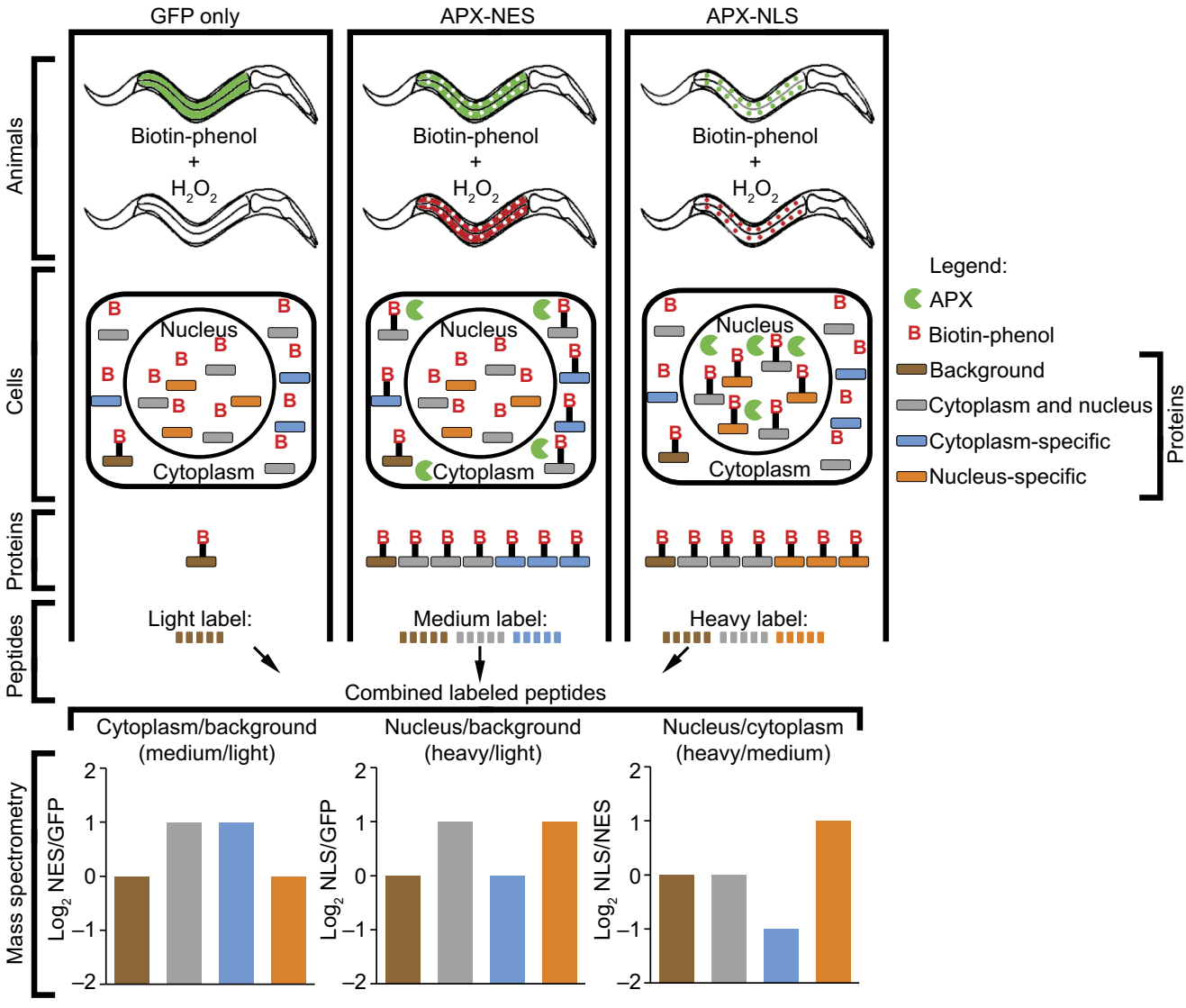


**Fig. 1. Overview of an approach to identify tissue- and subcellular-specific protein expression in *C. elegans*.** Schematic of spatially restricted enzymatic tagging in *C. elegans*. Animal strains that express the APX enzyme in either the cytoplasm or the nucleus in a tissue-specific manner (such as in the intestine as illustrated here) are generated. Animals are treated with biotin-phenol that diffuses into cells. The APX enzyme, in the presence of $H_2O_2$ and biotin-phenol, catalyzes the formation of a phenoxyl radical that covalently labels the neighboring proteins with biotin (red "B" with black bars) (*15*). Thus, biotin labeling (labeled in red) of proteins occurs in whichever specific tissues and subcellular locations the enzyme is expressed. Three strains are used to measure protein localization in each tissue: GFP only, GFP-APX-NES, and GFP-APX-NLS. Spatially restricted enzymatic tagging is performed using these three strains, and then the proteins are extracted and purified using streptavidin beads. These purified proteins are then digested into peptides and labeled using reductive dimethyl labeling for quantitative comparisons between the three strains. The peptides from each sample are then combined, and peptide ratios in each sample are measured using mass spectrometry. The peptide ratios can then be used to determine whether the protein is detected over background and enriched in either the nucleus or the cytoplasm.

these stained animals, and we found that the location of the biotin labeling was dependent on the tissue and compartment where APX was expressed (Fig. 3). Although labeling could not be detected on gels when the enzyme was localized to the nucleus of the pharyngeal muscle, efficient and specific labeling was observed in this location by microscopy (Fig. 3). The control GFP-only strain lacking APX did not display biotin labeling, demonstrating the specificity of the approach (Fig. 3). The combined results from examining protein extracts and from microscopy indicate the ability to label proteins in each of the eight locations that we tested. Together, this demonstrates the efficacy of in vivo spatially restricted enzymatic tagging in *C. elegans*.

## Identification of *C. elegans* cytoplasmic and nuclear proteins expressed in specific tissues using mass spectrometry

Having confirmed the tissue and cell compartment specificity of our in vivo APX-mediated proximity tagging approach, we then set out to identify proteins that are present in the cytoplasm and nucleus and in specific tissues using mass spectrometry. We used a quantitative strategy to rapidly and accurately compare proteins isolated from each tissue and between different subcellular locations. To identify cytoplasmic and nuclear proteins, we used a set of three strains for each tissue. We used the strains expressing APX-NES and APX-NLS in each tissue and the negative control strain that expresses GFP without

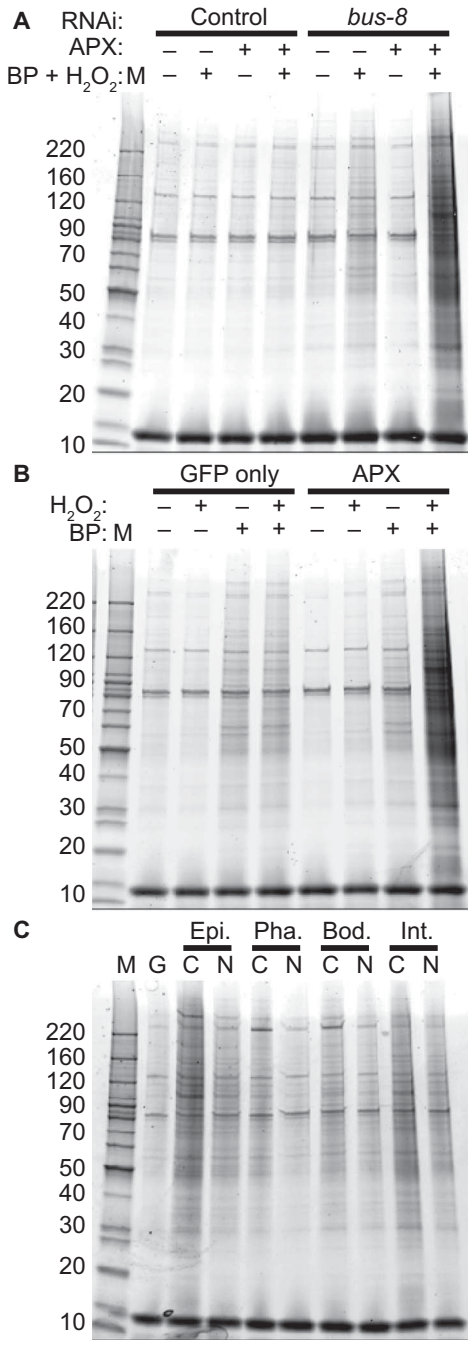


**Fig. 2. Efficient, spatially restricted enzymatic tagging in *C. elegans* is dependent on biotin-phenol, $H_2O_2$, and *bus-8* RNAi.** (**A** to **C**) Streptavidin-purified proteins from *C. elegans* protein extracts were visualized with Oriole staining. (A) Animals expressing GFP-APX (APX+) in the intestinal cytoplasm or a control GFP-only strain (APX–) were grown on plates with either control (L4440) or *bus-8* RNAi. Animals were either treated or untreated with biotin-phenol (BP) and $H_2O_2$. Protein markers are indicated (labeled "M"). (B) Animals expressing GFP-APX in the intestinal cytoplasm or a GFP control strain grown on plates with *bus-8* RNAi were treated with either $H_2O_2$ or biotin-phenol, or both. (C) Strains of *C. elegans* expressing GFP-APX in either the cytoplasm (C) or the nucleus (N) of the epidermis (Epi.), pharyngeal muscle (Pha.), body wall muscle (Bod.), or intestine (Int.). A strain expressing GFP only (G) is the negative control. All strains were grown on plates with *bus-8* RNAi and treated with biotin-phenol and $H_2O_2$.

APX in the intestine (Fig. 1). For the three samples in each tissue set, proteins were labeled by APX-mediated biotinylation and isolated as described above. Proteins bound to streptavidin-agarose from the three strains (GFP-APX-NES, GFP-APX-NLS, and GFP only) were then digested with trypsin.

For quantitative comparisons among the three samples in each tissue set, peptides from each sample were labeled with a different isotopic tag using reductive dimethyl labeling (Fig. 1) (*21*). For each tissue set, differentially labeled reductive dimethylated peptides from each sample were mixed in equal proportion before analysis by high-resolution mass spectrometry. Samples from each tissue set were prepared and analyzed in triplicate. To evaluate the ability of this approach to identify proteins in the APX samples above proteins in the control GFP-only samples, we initially performed control experiments using the three strains of the intestine tissue set (GFP-APX-NES, GFP-APX-NLS, and GFP only). As a control experiment, we pooled peptide samples from each of the three strains, separated them into three identical pools before dimethyl labeling, and, upon remixing in equal amounts, found that less than 5% of all proteins displayed quantitative ratios greater than twofold between pools. This result establishes a base false discovery rate of less than 5% for our experimental method (fig. S2A). In contrast, when peptide samples from the three strains were labeled by reductive dimethyl labeling individually before mixing, more than 90% of proteins from the APX-NES and APX-NLS samples had ratios over the control GFP-only sample greater than twofold (fig. S2B). These control experiments demonstrated the effectiveness of using reductive dimethyl labeling and mass spectrometry to quantitatively identify APX-mediated biotinylated proteins and established a twofold threshold as being able to differentiate proteins above background, which we used in subsequent analyses.

We then identified proteins from each of the experimental strains using the criterion that each protein has a ratio of at least twofold over background in two of the three biological replicate experiments (see Materials and Methods, Supplementary Materials, fig. S3, and table S3). Using this criterion, we identified between 108 and 2484 proteins for each of the eight strains (Fig. 4A). The proteins identified and ratios between pairs of strains in the same tissue set are reported in table S4. The largest tissues, the intestine and epidermis, had the most identified proteins, followed by the body wall muscle and then the pharyngeal muscle (table S5). More proteins were identified from the cytoplasm than from the nucleus for each tissue. A total of 3180 proteins were identified in at least one of the eight strains (Fig. 4B).

## Identification of tissue- and subcellular-specific *C. elegans* protein expression

We then investigated which of the proteins that we identified displayed tissue-specific expression. To identify proteins that are tissue-specific, we compared proteins detected at twofold above background in one of the four tissues in two of the three replicates, but not detected twofold above background in any of the replicates in the other three tissues. This resulted in the identification of 338 proteins that were specific to only one tissue (Fig. 4C). To assess the accuracy of their inferred tissue-specificity, we compared these proteins to an existing comprehensive data set of tissue-specific mRNA expression (*22*). This data set was generated from experimental expression data, and prediction scores were made for each gene in each of the four tissues we tested. The set of proteins we identified as being specific for each tissue had the highest average mRNA expression scores

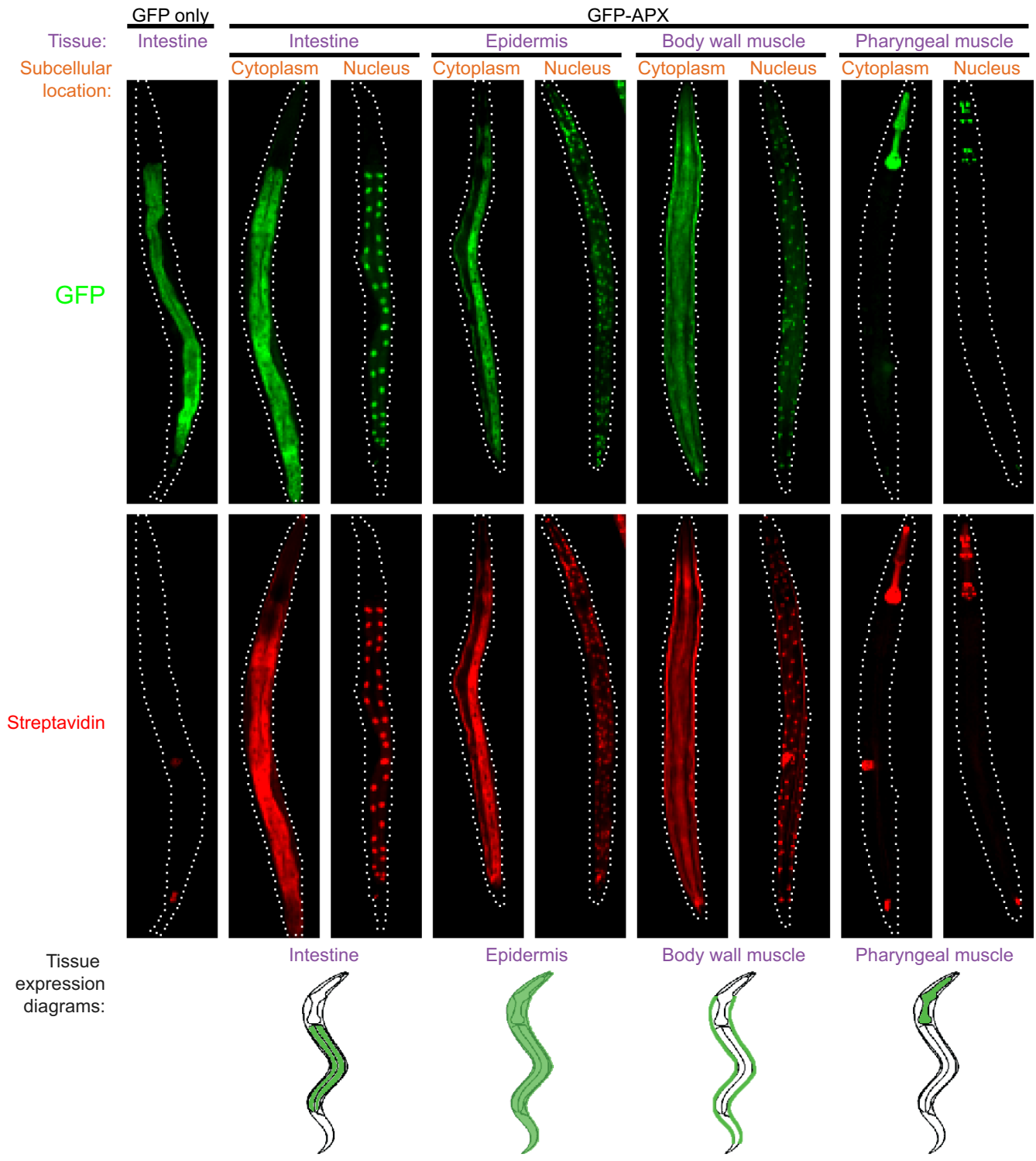


**Fig. 3. APX-mediated biotin labeling in vivo displays tissue and subcellular specificity.** Spatially restricted enzymatic tagging was performed on strains of *C. elegans* expressing APX in the cytoplasm or nucleus of the intestine, epidermis, body wall muscle, or pharyngeal muscle as indicated. A strain expressing GFP without APX was used as a negative control. Animals were fixed and stained for GFP (top, green) to determine the localization of the enzyme and streptavidin (middle, red) to determine the location of protein biotinylation. Representative images are shown for each strain. Animals are aligned so that the anterior is up. Tissue expression diagrams show the location of each tissue in *C. elegans* (bottom).

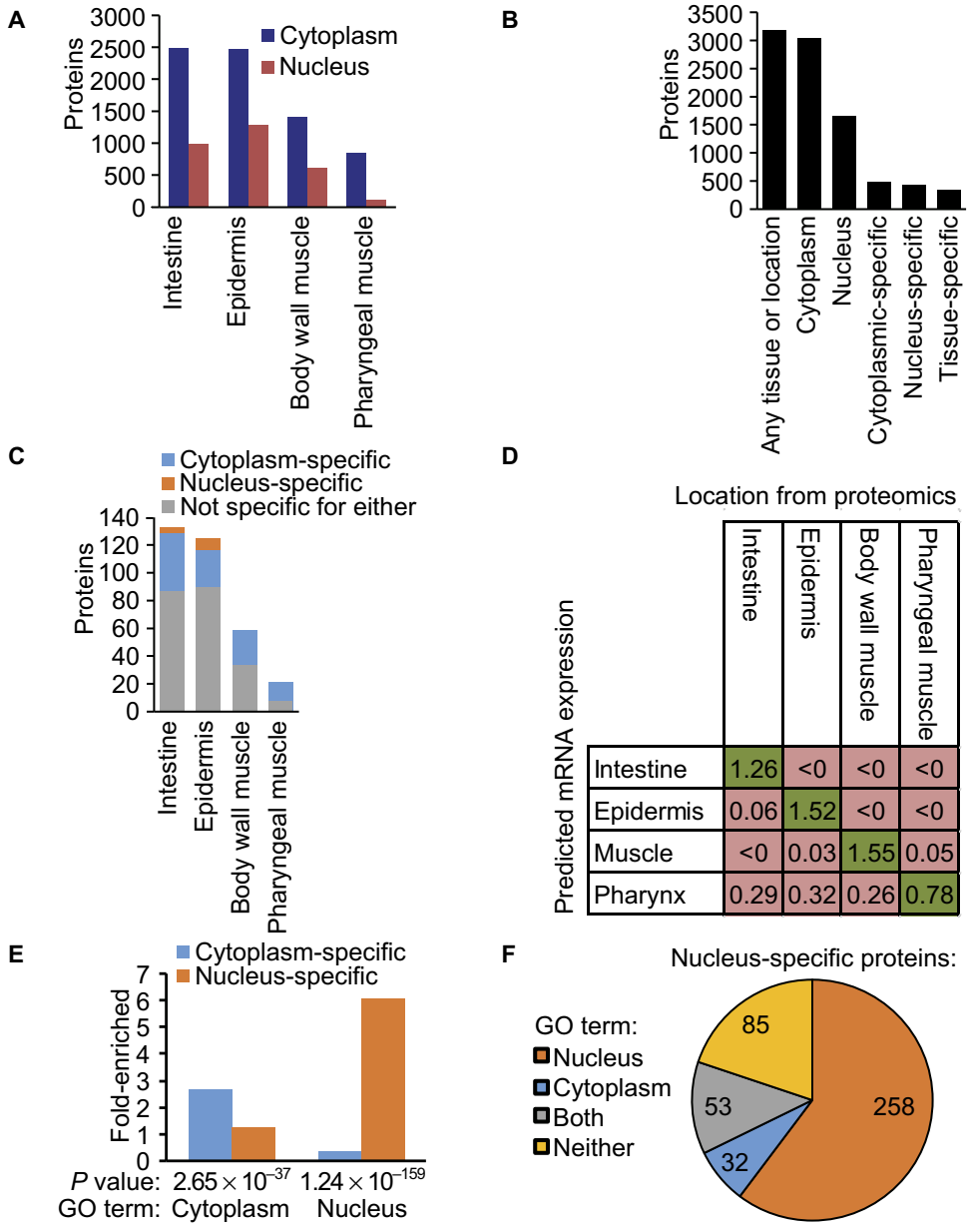


**Fig. 4. Identification of *C. elegans* proteins with tissue- and subcellular-specific localizations.** (**A**) The number of proteins identified above background from the tissue or subcellular location for each of the eight locations indicated. (**B**) The total number of proteins identified in our experiments that were detected in different locations or detected as being specific to a location. (**C**) The number of proteins we identified in the indicated tissue but not in the other three tissues. For each tissue, three categories of proteins are shown: those that are specific to the cytoplasm (orange), those that are specific to the nucleus (blue), and those that are not specific to either compartment (gray). (**D**) Comparison of the identified tissue-specific proteins to a data set of predicted mRNA expression (*22*). The data presented are the average of all the mRNA expression prediction scores for each tissue-specific protein in each of the four tissues. Higher prediction scores are more likely to be expressed in that tissue. Each column represents proteins identified as specific to that tissue, compared to the predicted mRNA expression of the tissue in each row. The highest average score in each column is shaded in green, and all other scores in the column are shaded in red. (**E**) Gene Ontology (GO) term enrichment analysis of proteins identified in our experiments as specific to either the cytoplasm or the nucleus. (**F**) Pie chart of the nucleus-specific proteins we identified in our experiment and whether they have a GO term location of either the nucleus or the cytoplasm, both, or neither.

in that tissue compared to the other three, supporting the accuracy of our technique (Fig. 4D). We identified the largest number of proteins from the intestine and epidermis as being tissue-specific. These tissues also had the lowest mRNA expression scores for those proteins identified as specific to other tissues. We identified the fewest tissue-specific proteins from the pharyngeal muscle; this tissue also had higher mRNA expression scores for proteins identified as being specific to other tissues. This lowered accuracy is likely due to identifying fewer total proteins in the pharyngeal muscle compared to the other tissues (Fig. 4A). Thus, some proteins that we describe as specific to non-pharyngeal tissues may actually be expressed in the pharyngeal muscle as well, but were undetected in the pharyngeal muscle because of the

lowered sensitivity of our technique in this small tissue. This result suggests that these analyses are more useful for the large tissues where we identified more proteins. However, because the pharyngeal-specific proteins we describe do have the highest expression score in the pharynx compared to the other tissues, it demonstrates that this analysis is still useful for identifying tissue-specific proteins even in smaller tissues.

We then examined the subcellular specificity of the proteins we identified by comparing the ratios of protein levels in the nucleus compared to protein levels in the cytoplasm. We normalized the quantitative ratios between the APX-NES and APX-NLS samples and used them to identify proteins that were enriched in either location (fig. S4 and table S6). This comparison of proteins enriched twofold resulted in the identification of 486 proteins specific to the cytoplasm and 428 proteins specific to the nucleus (Fig. 4E). We tested a number of thresholds and found that using more stringent cutoffs above twofold does not greatly enhance specificity (fig. S5); for this reason, we used twofold cutoffs for our analysis. To assess the accuracy of these location assignments, we performed GO term enrichment analysis using PANTHER (*23*). Here, we found that the proteins identified as being cytoplasm- or nucleus-specific were highly enriched for proteins previously annotated to be cytoplasmic or nuclear (Fig. 4E). Of the 428 proteins identified as nucleus-specific, 117 were not previously known to be nuclear-localized (Fig. 4F). We also examined the compartment specificity of the proteins we classified as being tissue-specific and identified 107 of these proteins as being cytoplasm-specific and 12 as being nucleus-specific (Fig. 4C).

To validate our results, we used fluorescence microscopy to confirm the localization of several proteins that we identified to be expressed in specific tissues or subcellular locations using our quantitative proteomics approach. We chose seven proteins that we measured with high confidence to be either nucleus- or cytoplasm-specific or specific to one of the four tissues (fig. S6). These seven proteins had no previous experimentally determined location described in WormBase (www.wormbase.org). Using TransgeneOme (*5*) constructs that contain C-terminal GFP fusions of each protein expressed under the native promoter, we generated strains of transgenic animals overexpressing each protein. We found that each of these seven test proteins localized to the corresponding tissue or subcellular location that we identified by spatially restricted enzymatic tagging (Fig. 5). These results confirm the accuracy of our approach and demonstrate the efficacy of using in vivo proximity-based labeling methods and quantitative mass spectrometry to identify proteins with tissue-specific and/or subcellular compartment-specific localization. Overall, we present a robust method that can be applied to detect in vivo protein localization in an unbiased manner within intact animals and provide a resource of proteins with specific locations in *C. elegans* (table S3).

## DISCUSSION

Here, we describe an approach that allows for the determination of in vivo protein localization in an intact animal through the use of spatially restricted enzymatic tagging. To our knowledge, this is the first in vivo localization of a large number of proteins with subcellular resolution in a live animal. Using spatially restricted enzymatic tagging in *C. elegans*, we provide one of the largest systematic identifications of proteins that have tissue- or compartment-specific localizations. We identified a total of 3180 proteins, 1132 of which are localized to a specific tissue or subcellular compartment. This resource represents

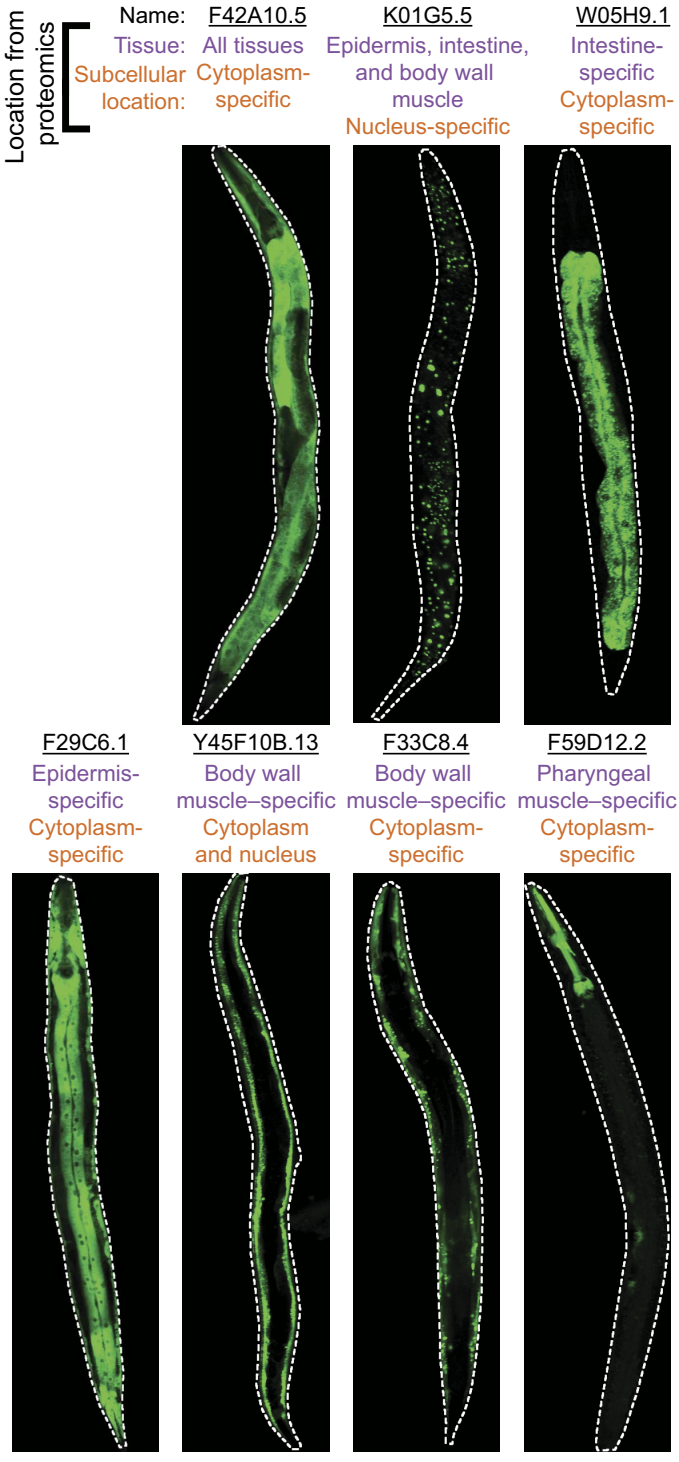


**Fig. 5. Validation of identified protein locations using fluorescently tagged proteins.** Strains of transgenic *C. elegans* expressing GFP-tagged proteins identified to be tissue- or location-specific in our study. Animals were grown to the L4 stage, and representative images displaying protein localization are shown. The protein name is listed above each construct-expressing strain. The tissue and subcellular localization determined from our proteomic data is listed below the protein name. Animals are aligned so that the anterior is up.

an advancement from previous studies aimed at identifying protein localization using GFP tagging, in which the location of 230 proteins was characterized (*5*). A different approach using unnatural amino acid labeling in the pharyngeal muscle in *C. elegans* identified 43 proteins that were greater than twofold enriched above background levels (*13*). In contrast, our technique identified 887 proteins in total from either the cytoplasm or the nucleus of the pharyngeal muscle. This technique complements several approaches that have been used in *C. elegans* to detect mRNA levels in specific tissues (*22*, *24*–*26*). Although our approach is less sensitive than these methods, it directly detects protein levels and can provide information on their subcellular localization.

Despite our success in using spatially restricted enzymatic tagging to identify tissue- and subcellular-specific protein expression, there are some limitations to this approach that warrant discussion. The number of proteins identified is only ~7 to 21% of the number of mRNAs demonstrated to be expressed in the same tissue (table S5) (*24*), although there are regulatory processes that affect how efficiently the mRNA is translated into protein, which results in mRNA and protein levels not being strongly correlated (*27*). In addition, fewer proteins were identified in smaller tissues than in larger tissues, and in the nucleus than in the cytoplasm. These concerns may be addressed in the future by using recently developed versions of APX that have increased sensitivity (*28*). Moreover, because the variability of identified proteins was greater in smaller tissues (fig. S3), the measurement of additional replicates could be used to increase the sensitivity of protein detection. Another issue is that our approach relies on knocking down *bus-8* using RNAi, which causes improper development of the cuticle and locomotion defects (*20*). However, we were able to confirm the specific location of a number of proteins using transgenic analysis in wild-type animals, demonstrating that protein localization in *bus-8*–defective animals appears to be largely similar to that in wild-type animals. In addition, there are potentially other mutants or chemical approaches that could be used to improve the accessibility of the biotin-phenol substrate and may lessen their physiological impact on the animals.

The methodology we describe here can be expanded to obtain greater tissue and subcellular resolution. Spatially restricted enzymatic tagging has also been reported to work in other subcellular locations in human cells, including the endoplasmic reticulum, mitochondria, and plasma membrane (*15*). We were able to identify proteins from all eight locations to which we localized the enzyme. Thus, it is likely that this approach could be applied to a number of other tissues and subcellular locations in *C. elegans*.

The technique we describe could also be useful for a number of other applications. This technique could be used to determine protein translocation between the cytoplasm and the nucleus under different growth conditions. Through the use of a yeast GFP-tagged library, 71 cytoplasmic proteins were shown to localize to the nucleus under starvation conditions, demonstrating that large numbers of proteins translocate in response to stress (*29*). Using our described approach, these types of translocation events could be measured globally in live animals, using a smaller number of strains to investigate the response to various stress conditions. Although our experiments were focused on quantifying differences between the cytoplasm and nucleus, the quantitative labeling scheme we used is flexible and could be used to directly compare levels of proteins between different tissues. In addition, this approach could also be adapted to measure posttranslational modifications, such as phosphorylation. This would allow comparing differences in post-translational modifications of proteins between different compartments and between different tissues (*30*). This approach can also be applied to identify proteins from pathogenic or symbiotic microbes that localize to different host tissues and subcellular locations (*31*). Spatially restricted enzymatic tagging has now been reported in *D. melanogaster*, *C. elegans*, and human cells and thus can likely be used in any organism where transgenic techniques exist and biotin-phenol can be delivered.

## MATERIALS AND METHODS

### Cloning and generation of strains

The protein sequence of soybean APX with the W41F mutation (*15*) was optimized for *C. elegans* expression using DNAWorks to design primers (*32*). Primers were annealed using a two-step polymerase chain reaction (PCR) method, and double-stranded DNA was cloned into Gateway vector pDONR 221 using BP Clonase II (Thermo Fisher). This construct was modified using Gibson cloning (*33*) with an N-terminal fusion of GFP. This construct was additionally modified to encode the C-terminal NES (LQLPPLERLTLD) and NLS (PKKKRKVDPKKKRKVDPKKKRKV) by encoding these tags into primers, amplifying the plasmid with PCR, and ligating the PCR product. Upstream regions of the following *C. elegans* genes were used as promoters: *dpy-7* (epidermis), *spp-5* (intestine), *myo-2* (pharyngeal muscle), and *myo-3* (body wall muscle). These promoters were cloned into the 5′ plasmid pDONR P4-P1R using BP Clonase II (Thermo Fisher). Multisite Gateway cloning was used to generate targeting constructs using LR Clonase II plus (Thermo Fisher) to combine one of the four promoter plasmids, one of the two APX containing plasmids, the 3′ plasmid pDONR P2R-P3 vector containing the 3′ region of *unc-54*, and the destination vector pCFJ150. These targeting constructs along with the Mos1 transposase and marker plasmids were injected into *unc-119* mutants from the strain EG6699 (*34*). Non-*Unc* worms were recovered, and each transgenic strain was backcrossed three times into the wild-type N2 strain. The homozygote was used in subsequent experiments. Transgenic strains expressing TransgeneOme GFP-tagged proteins as extrachromosomal arrays were generated by injecting constructs into EG6699 and selecting non-unc animals. The following amount of DNA was injected for each construct: 100 ng/μl of each construct for F33C8.4, Y45F10B.13, F59D12.2, and F29C6.1; 50 ng/μl of the construct with 50 ng/μl pBSK for F42A10.5; and 10 ng/μl of the construct with 90 ng/μl pBSK for W05H9.1 and K01G5.5. All strains used in the study are listed in table S2. All *C. elegans* strains were maintained using standard procedures (*35*).

### Spatially restricted enzymatic tagging in *C. elegans*

Populations of animals were grown and bleached to recover eggs, which were then hatched to generate first larval stage (L1) synchronized animals (*35*). About 30,000 L1 animals of each strain in 2.5 ml of M9 buffer were added to a 15-cm RNAi plate seeded with HT115 bacteria expressing a *bus-8* RNAi feeding clone (*36*). Animals were protected from light and grown to the L4 stage on these plates for 44 hours at 20°C. To recover animals, each plate was washed with M9T (M9/0.1% Tween 20). The recovered animals were washed once with M9T. These animals were then placed into 1.5-ml tubes in a total of 100 μl of M9T. To each sample, we added 900 μl of the labeling solution (0.1% Tween 20, M9, and 3.3 mM biotin-phenol, synthesized as previously described) (*15*). Samples were incubated for 1 hour at

22° to 24°C on an end-over-end rotator. To activate biotin labeling, 10 µl of 100 mM $H_2O_2$ was added for 2 min. To quench the reaction, 500 µl of quench buffer (M9, 0.1% Tween 20, 10 mM sodium azide, 10 mM sodium ascorbate, and 5 mM Trolox) was added. Samples were then washed four times with 1 ml of quench buffer. After the last wash, the remaining buffer was removed, and 800 µl of lysis buffer [150 mM NaCl, 50 mM tris (pH 8), 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 10 mM sodium azide, protease complete tablet (Roche), 10 mM sodium ascorbate, 5 mM Trolox, and 1 mM phenylmethylsulfonyl fluoride] was added. Animals were then immediately frozen dropwise in liquid $N_2$.

To extract proteins, frozen worm pellets were ground to a fine powder in liquid $N_2$. To generate supernatants, these protein extracts were then centrifuged for 10 min at 21,000*g* at 4°C. The supernatant was then filtered over a desalting column with a 7000 molecular weight cutoff (Pierce). The protein concentrations of the extracts were measured using a Pierce 660-nm Protein Assay and normalized. We added 25 µl of high-capacity streptavidin agarose resin (Pierce) in a total of 700 µl of lysis buffer to 450 to 550 µg of each sample. Extracts were incubated with beads for 1 hour on an end-over-end rotator. Beads were then washed five times with 1 ml of lysis buffer, three times with 1 ml of 8 M urea/10 mM tris (pH 8), and three times with 1 ml of phosphate-buffered saline (PBS). The liquid was removed from the beads, and 100 µl of trypsin (0.1 µg/µl; Promega)/100 mM triethylammonium bicarbonate was added to each sample and incubated at 37°C for 24 hours.

These peptides were then differentially labeled with reductive dimethyl labeling as previously described (*21*). Briefly, each of the three samples in the set was labeled with a different isotopic tag that differed by 4 Da. To generate samples with a light tag, 4 µl of 4% (v/v) $CH_2O$ and 4 µl of 600 mM $NaBH_3CN$ were added. The other tags were generated in a similar way, with $CD_2O$ and $NaBH_3CN$ being used for the medium tag and $C^{13}D_2O$ and $NaBD_3CN$ being used for the heavy tag. Samples were then incubated for 1 hour with mixing on an end-over-end rotator. To quench the reaction, 16 µl of 1% (v/v) ammonia was added to each sample. The samples were then acidified by adding 8 µl of formic acid. The three samples in each set were then combined.

### Gel analysis of biotinylated proteins

From the samples prepared as described above, 15% of beads were removed before digestion. Liquid was removed from the beads, and 20 µl of Laemmli buffer with 2 mM biotin was added. Samples were heated for 10 min at 95°C. Fifteen microliters of each sample was loaded onto a 4 to 20% polyacrylamide gradient gel (Bio-Rad). Gels were then stained with Oriole fluorescent gel stain (Bio-Rad) to visualize proteins.

### Microscopy

To analyze biotin labeling with immunohistochemistry, worms were fixed using Bouin's tube fixation method (*37*). Fixed worms were stained overnight with 1:500 anti-GFP mouse antibody (Roche) in block buffer (PBS, 0.5% Triton X-100, and 1% bovine serum albumin). Worms were then washed with block buffer and stained overnight with 1:500 fluorescein isothiocyanate–conjugated anti-mouse secondary antibody (Invitrogen) and 1:500 streptavidin Alexa Fluor 568 (Thermo Fisher) in block buffer. Worms were then washed in block buffer and imaged using a Zeiss LSM 700 confocal microscope. Live worms expressing GFP-tagged TransgeneOme constructs (*5*) were grown to the L4 stage, treated with 1 mM levamisole, and then imaged as described above.

### Sample analysis by mass spectrometry

Before analysis by liquid chromatography (LC)–tandem mass spectroscopy (MS/MS), peptides were desalted by solid-phase extraction using in-house prepared $C_{18}$ StageTips (*38*) and reconstituted in 5% formic acid and 5% acetonitrile (ACN). All samples were analyzed in triplicate using a Q Exactive mass spectrometer (Thermo Fisher Scientific). The following is a generalized nanoflow high-performance LC (HPLC) and data acquisition method that is representative of individual analyses. Peptides were first separated by reversed-phase chromatography using a fused silica microcapillary column (75-µm inner diameter, 18 cm) packed with $C_{18}$ silica (ReproSil-Pur 120 C18-AQ; 1.9 µm, Dr. Maisch GmbH) using an in-line nanoflow EASY-nLC 1000 ultrahigh HPLC system (Thermo Fisher Scientific). Peptides were eluted over a 100-min 0 to 30% ACN gradient, followed by a 5-min 30 to 60% ACN gradient, and a 5-min 60 to 95% ACN gradient, with a final 10-min isocratic step at 0% ACN for a total run time of 120 min at a flow rate of 250 nl/min. All gradient mobile phases contained 0.1% formic acid. MS/MS data were collected in data-dependent mode using a top 10 method with a full MS mass range from 400 to 1800 mass/charge ratio, a resolution of 70,000, and an automatic gain control target of $3 \times 10^6$. MS2 scans were triggered when an ion intensity threshold of $1 \times 10^5$ was reached with a maximum injection time of 60 ms. Peptides were fragmented using a normalized collision energy setting of 25. A dynamic exclusion time of 20 s was used, and the peptide match setting was disabled. Singly charged ions, charge states above 6, and unassigned charge states were excluded.

### Peptide and protein identification and quantification

The resultant RAW files were converted into mzXML format using the ReAdW.exe (version 4.3.1) program. The SEQUEST search algorithm (version 28) was used to search MS/MS spectra against a concatenated target-decoy database composed of forward and reverse sequences from the reviewed UniProtKB/Swiss-Prot FASTA *C. elegans* database combined with the UniProtKB *E. coli* (K12 strain) database, and with common contaminant proteins appended. Each mzXML file was searched in triplicate with the following parameters: 20 parts per million precursor ion tolerance and 0.01-Da fragment ion tolerance; Trypsin (1 1 KR P) was set as the enzyme; up to three missed cleavages were allowed; and dynamic modification of 15.99491 Da on methionine (oxidation). For searches with light and medium reductive dimethyl labels, additional dynamic modifications of 4.0224 Da on lysine and peptide N termini and static modifications of 28.0313 Da on lysine and peptide N termini were included. For searches with light and heavy reductive dimethyl labels, additional dynamic modifications of 8.04437 Da on lysine and peptide N termini and static modifications of 28.0313 Da on lysine and peptide N termini were included. For searches with medium and heavy reductive dimethyl labels, additional dynamic modifications of 4.02193 Da on lysine and peptide N termini and static modifications of 32.05374 Da on lysine and peptide N termini were included. Peptide matches were filtered to a peptide false discovery rate of 2% using the linear discriminant analysis. Proteins were further filtered to a false discovery rate of 2%, peptides were assembled into proteins using maximum parsimony, and only unique and razor peptides were retained for subsequent analysis. All peptide heavy/light,

medium/light, and heavy/medium ratios with a signal-to-noise ratio of above 5 were used for assembled protein quantitative ratios.

### Analysis of mass spectrometry data

We classified a protein as being identified above background (that is, present in a particular tissue) if it had a greater than twofold ratio of NLS or NES over the GFP-only strain in two of the three replicates for a given tissue. The requirements for a protein to be classified as being tissue-specific were as follows: (i) if it has a greater than twofold ratio of NLS/GFP in two of the three replicates or a greater than twofold ratio of NES/GFP in two of the three replicates in one of the tissues and (ii) if it is not detected in any of the other tissues with a greater than twofold ratio of NLS/GFP or NES/GFP in any of the three replicates. These identified proteins were compared to the predicted mRNA expression scores (*22*). Cytoplasm- or nucleus-specific proteins were determined by comparing the ratio between the NES and NLS samples. Because more peptides were detected in the NES sample than in the NLS sample from each tissue, the NLS/NES ratios were adjusted using total intensity normalization. The total spectral counts were summed in each NLS and NES sample, and the ratio was used as a scaling factor that was then multiplied by the NLS/NES ratio of each protein to calculate the scaled NLS/NES ratio values. The total spectral counts for each sample and the calculated scaling factors for each replicate are reported in table S6. Proteins that were detected above background (see above for criteria) and with a greater than twofold NLS/NES adjusted ratio were classified as nucleus-specific, and those with a greater than twofold NES/NLS adjusted ratio were classified as cytoplasm-specific. GO term enrichment was performed using PANTHER (*23*). Proteins to validate tissue-specific expression using GFP-tagged constructs were selected on the basis of the criteria of (i) having a greater than twofold ratio of NLS/GFP in all three replicates or a greater than twofold ratio of NES/GFP in all three replicates and (ii) not having NLS/GFP or NES/GFP ratios greater than 0 in any of the replicates for the other three tissues. Proteins to validate subcellular specificity using GFP-tagged constructs were selected based on the following criteria: having NES/GFP and NES/NLS ratios greater than twofold in all three replicates in at least three tissues (cytoplasm-specific) and having NLS/GFP and NLS/NES ratios greater than twofold in all three replicates in at least three tissues (nucleus-specific). All data for the identified proteins and ratios between samples are reported in table S3.

## SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/3/5/e1602426/DC1

Supplementary Discussion

fig. S1. Localization of APX to the intestinal cytoplasm of *C. elegans*.

fig. S2. Use of triplex reductive dimethylated samples to quantitatively compare GFP-APX-NES, GFP-APX-NLS, and GFP-only samples.

fig. S3. Reproducibility of the number of proteins identified from each location.

fig. S4. Identification of cytoplasm- or nucleus-localized proteins.

fig. S5. Effect of enrichment threshold on the identification of cytoplasm- or nucleus-localized proteins.

fig. S6. Quantitative mass spectrometry ratios of proteins selected for validation.

table S1. Comparison of existing techniques to determine protein localization in animals.

table S2. List of strains used in this study.

table S3. Pearson correlation coefficients of replicate samples.

table S4. Mass spectrometry data of spatially restricted enzymatic tagging in *C. elegans*.

table S5. Number of proteins measured compared to mRNA expressed in each tissue.

table S6. Total spectral counts for all identified proteins in each sample and NLS/NES scaling factors.

## REFERENCES AND NOTES


1. W.-K. Huh, J. V. Falvo, L. C. Gerke, A. S. Carroll, R. W. Howson, J. S. Weissman, E. K. O'Shea, Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003).
2. M. Kitagawa, T. Ara, M. Arifuzzaman, T. Ioka-Nakamichi, E. Inamoto, H. Toyonaga, H. Mori, Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli K*-12 ORF archive): Unique resources for biological research. *DNA Res.* **12**, 291–299 (2005).
3. N. J. Kuwada, B. Traxler, P. A. Wiggins, Genome-scale quantitative characterization of bacterial protein localization dynamics throughout the cell cycle. *Mol. Microbiol.* **95**, 64–79 (2015).
4. J. N. Werner, E. Y. Chen, J. M. Guberman, A. R. Zippilli, J. J. Irgon, Z. Gitai, Quantitative genome-scale analysis of protein localization in an asymmetric bacterium. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 7858–7863 (2009).
5. M. Sarov, J. I. Murray, K. Schanze, A. Pozniakovski, W. Niu, K. Angermann, S. Hasse, M. Rupprecht, E. Vinis, M. Tinney, E. Preston, A. Zinke, S. Enst, T. Teichgraber, J. Janette, K. Reis, S. Janosch, S. Schloissnig, R. K. Ejsmont, C. Slightam, X. Xu, S. K. Kim, V. Reinke, A. F. Stewart, M. Snyder, R. H. Waterston, A. A. Hyman, A genome-scale resource for in vivo tag-based protein function exploration in *C. elegans*. *Cell* **150**, 855–866 (2012).
6. M. Sarov, C. Barz, H. Jambor, M. Y. Hein, C. Schmied, D. Suchold, B. Stender, S. Janosch, V. V. KJ, R. T. Krishnan, A. Krishnamoorthy, I. R. S. Ferreira, R. K. Ejsmont, K. Finkl, S. Hasse, P. Kämpfer, N. Plewka, E. Vinis, S. Schloissnig, E. Knust, V. Hartenstein, M. Mann, M. Ramaswami, K. VijayRaghavan, P. Tomancak, F. Schnorrer, A genome-wide resource for the analysis of protein localisation in *Drosophila*. *eLife* **5**, e12068 (2016).
7. T. Geiger, A. Velic, B. Macek, E. Lundberg, C. Kampf, N. Nagaraj, M. Uhlen, J. Cox, M. Mann, Initial quantitative proteomic map of 28 mouse tissues using the SILAC mouse. *Mol. Cell. Proteomics* **12**, 1709–1722 (2013).
8. M.-S. Kim, S. M. Pinto, D. Getnet, R. S. Nirujogi, S. S. Manda, R. Chaerkady, A. K. Madugundu, D. S. Kelkar, R. Isserlin, S. Jain, J. K. Thomas, B. Muthusamy, P. Leal-Rojas, P. Kumar, N. A. Sahasrabuddhe, L. Balakrishnan, J. Advani, B. George, S. Renuse, L. D. N. Selvan, A. H. Patil, V. Nanjappa, A. Radhakrishnan, S. Prasad, T. Subbannayya, R. Raju, M. Kumar, S. K. Sreenivasamurthy, A. Marimuthu, G. J. Sathe, S. Chavan, K. K. Datta, Y. Subbannayya, A. Sahu, S. D. Yelamanchi, S. Jayaram, P. Rajagopalan, J. Sharma, K. R. Murthy, N. Syed, R. Goel, A. A. Khan, S. Ahmad, G. Dey, K. Mudgal, A. Chatterjee, T.-C. Huang, J. Zhong, X. Wu, P. G. Shaw, D. Freed, M. S. Zahari, K. K. Mukherjee, S. Shankar, A. Mahadevan, H. Lam, C. J. Mitchell, S. K. Shankar, P. Satishchandra, J. T. Schroeder, R. Sirdeshmukh, A. Maitra, S. D. Leach, C. G. Drake, M. K. Halushka, T. S. K. Prasad, R. H. Hruban, C. L. Kerr, G. D. Bader, C. A. Iacobuzio-Donahue, H. Gowda, A. Pandey, A draft map of the human proteome. *Nature* **509**, 575–581 (2014).
9. C. V. Jongeneel, M. Delorenzi, C. Iseli, D. Zhou, C. D. Haudenschild, I. Khrebtukova, D. Kuznetsov, B. J. Stevenson, R. L. Strausberg, A. J. G. Simpson, T. J. Vasicek, An atlas of human gene expression from massively parallel signature sequencing (MPSS). *Genome Res.* **15**, 1007–1014 (2005).
10. J. Li, T. Cai, P. Wu, Z. Cui, X. Chen, J. Hou, Z. Xie, P. Xue, L. Shi, P. Liu, J. R. Yates III, F. Yang, Proteomic analysis of mitochondria from *Caenorhabditis elegans*. *Proteomics* **9**, 4539–4553 (2009).
11. M. Wühr, T. Güttler, L. Peshkin, G. C. McAlister, M. Sonnett, K. Ishihara, A. C. Groen, M. Presler, B. K. Erickson, T. J. Mitchison, M. W. Kirschner, S. P. Gygi, The nuclear proteome of a vertebrate. *Curr. Biol.* **25**, 2663–2671 (2015).
12. C.-L. Chen, Y. Hu, N. D. Udeshi, T. Y. Lau, F. Wirtz-Peitz, L. He, A. Y. Ting, S. A. Carr, N. Perrimon, Proteomic mapping in live *Drosophila* tissues using an engineered ascorbate peroxidase. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 12093–12098 (2015).
13. K. P. Yuet, M. K. Doma, J. T. Ngo, M. J. Sweredoski, R. L. J. Graham, A. Moradian, S. Hess, E. M. Schuman, P. W. Sternberg, D. A. Tirrell, Cell-specific proteomic analysis in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 2705–2710 (2015).
14. T. S. Elliott, F. M. Townsley, A. Bianco, R. J. Ernst, A. Sachdeva, S. J. Elsässer, L. Davis, K. Lang, R. Pisa, S. Greiss, K. S. Lilley, J. W. Chin, Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. *Nat. Biotechnol.* **32**, 465–472 (2014).
15. H.-W. Rhee, P. Zou, N. D. Udeshi, J. D. Martell, V. K. Mootha, S. A. Carr, A. Y. Ting, Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* **339**, 1328–1331 (2013).
16. Z. F. Altun, L. A. Herndon, C. A. Wolkow, C. Crocker, R. Lints, D. H. Hall, *Handbook of C. elegans Anatomy* (Wormatlas, 2016).
17. D. L. Riddle, T. Blumenthal, B. J. Meyer, J. R. Priess, *The Biological Model* (Cold Spring Harbor Laboratory Press, 1997).
18. C. Frøkjær-Jensen, M. W. Davis, C. E. Hopkins, B. J. Newman, J. M. Thummel, S.-P. Olesen, M. Grunnet, E. M. Jorgensen, Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat. Genet.* **40**, 1375–1383 (2008).
19. S. L. Ooi, J. G. Henikoff, S. Henikoff, A native chromatin purification system for epigenomic profiling in *Caenorhabditis elegans*. *Nucleic Acids Res.* **38**, e26 (2010).
20. F. A. Partridge, A. W. Tearle, M. J. Gravato-Nobre, W. R. Schafer, J. Hodgkin, The *C. elegans* glycosyltransferase BUS-8 has two distinct and essential roles in epidermal morphogenesis. *Dev. Biol.* **317**, 549–559 (2008).

21. P. J. Boersema, R. Raijmakers, S. Lemeer, S. Mohammed, A. J. R. Heck, Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc.* **4**, 484–494 (2009).
22. M. D. Chikina, C. Huttenhower, C. T. Murphy, O. G. Troyanskaya, Global prediction of tissue-specific gene expression and context-dependent gene networks in *Caenorhabditis elegans*. *PLOS Comput. Biol.* **5**, e1000417 (2009).
23. H. Mi, A. Muruganujan, J. T. Casagrande, P. D. Thomas, Large-scale gene function analysis with the PANTHER classification system. *Nat. Protoc.* **8**, 1551–1566 (2013).
24. W. C. Spencer, G. Zeller, J. D. Watson, S. R. Henz, K. L. Watkins, R. D. McWhirter, S. Petersen, V. T. Sreedharan, C. Widmer, J. Jo, V. Reinke, L. Petrella, S. Strome, S. E. Von Stetina, M. Katz, S. Shaham, G. Rätsch, D. M. Miller III, A spatial and temporal map of *C. elegans* gene expression. *Genome Res.* **21**, 325–341 (2011).
25. R. Hunt-Newbury, R. Viveiros, R. Johnsen, A. Mah, D. Anastas, L. Fang, E. Halfnight, D. Lee, J. Lin, A. Lorch, S. McKay, H. M. Okada, J. Pan, A. K. Schulz, D. Tu, K. Wong, Z. Zhao, A. Alexeyenko, T. Burglin, E. Sonnhammer, R. Schnabel, S. J. Jones, M. A. Marra, D. L. Baillie, D. G. Moerman, High-throughput in vivo analysis of gene expression in *Caenorhabditis elegans*. *PLOS Biol.* **5**, e237 (2007).
26. D. Dupuy, N. Bertin, C. A. Hidalgo, K. Venkatesan, D. Tu, D. Lee, J. Rosenberg, N. Svrzikapa, A. Blanc, A. Carnec, A.-R. Carvunis, R. Pulak, J. Shingles, J. Reece-Hoyes, R. Hunt-Newbury, R. Viveiros, W. A. Mohler, M. Tasan, F. P. Roth, C. Le Peuch, I. A. Hope, R. Johnsen, D. G. Moerman, A.-L. Barabási, D. Baillie, M. Vidal, Genome-scale analysis of in vivo spatiotemporal promoter activity in *Caenorhabditis elegans*. *Nat. Biotechnol.* **25**, 663–668 (2007).
27. C. Vogel, E. M. Marcotte, Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **13**, 227–232 (2012).
28. S. S. Lam, J. D. Martell, K. J. Kamer, T. J. Deerinck, M. H. Ellisman, V. K. Mootha, A. Y. Ting, Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat. Methods* **12**, 51–54 (2015).
29. M. Breker, M. Gymrek, M. Schuldiner, A novel single-cell screening platform reveals proteome plasticity during yeast stress responses. *J. Cell Biol.* **200**, 839–850 (2013).
30. J. V. Olsen, M. Mann, Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol. Cell. Proteomics* **12**, 3444–3452 (2013).
31. A. W. Reinke, K. M. Balla, E. J. Bennett, E. R. Troemel, Identification of microsporidia host-exposed proteins reveals a repertoire of large paralogous gene families and rapidly evolving proteins. bioRxiv 10.1101/056788 (2016).
32. D. M. Hoover, J. Lubkowski, DNAWorks: An automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res.* **30**, e43 (2002).
33. D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison III, H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
34. C. Frøkjær-Jensen, M. W. Davis, M. Ailion, E. M. Jorgensen, Improved *Mos1*-mediated transgenesis in *C. elegans*. *Nat. Methods* **9**, 117–118 (2012).
35. T. Stiernagle, Maintenance of *C. elegans*. *WormBook*; www.wormbook.org.
36. R. S. Kamath, A. G. Fraser, Y. Dong, G. Poulin, R. Durbin, M. Gotta, A. Kanapin, N. Le Bot, S. Moreno, M. Sohrmann, D. P. Welchman, P. Zipperlen, J. Ahringer, Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**, 231–237 (2003).
37. J. S. Duerr, Immunohistochemistry. *WormBook*; www.wormbook.org.
38. J. Rappsilber, M. Mann, Y. Ishihama, Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).


**Acknowledgments:** We thank K. Balla, R. Luallen, and K. Reddy for providing helpful comments on the manuscript. We also thank N. Kosa and R. Ardecky for providing aid in purifying the biotin-phenol. **Funding:** A.W.R. is a Monsanto Fellow of the Life Sciences Research Foundation. This work was supported by GM114139, AG052622, and a Burroughs Wellcome Fund fellowship to E.R.T. and by DP2GM119132, a Hellman Fellowship, and a New Scholar Award from the Ellison Medical Foundation to E.J.B. **Author contributions:** A.W.R. designed, conducted, and analyzed the experiments and co-wrote the paper. R.M. performed the mass spectrometry analysis. E.J.B. designed the experiments, performed the mass spectrometry analysis, and co-wrote the paper. E.R.T. provided mentorship and guidance to the project and co-wrote the paper. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 3 October 2016
Accepted 9 March 2017
Published 10 May 2017
10.1126/sciadv.1602426

**Citation:** A. W. Reinke, R. Mak, E. R. Troemel, E. J. Bennett, In vivo mapping of tissue- and subcellular-specific proteomes in *Caenorhabditis elegans*. *Sci. Adv.* **3**, e1602426 (2017).