Specific aspartyl and calpain proteases are required for neurodegeneration in *C. elegans*

Popi Syntichaki, Keli Xu, Monica Driscoll & Nektarios Tavernarakis

Supplementary Information

Supplementary Methods

GeneBank accession numbers of the aspartyl protease sequences used for the analysis are: Human hCathepsinD, NP_001900; mouse mCathepsinD, NP_034113; Drosophila dRE02351p, AAL48533; yeast yPep4p, NP_015171; ASP-1 (ORF: Y39B6A.20), AAF19442; ASP-2 (ORF: T18H9.2), AAA83331; ASP-3 (ORF: H22K11.1), AAK39240; ASP-4 (ORF: R12H7.2), NP_510191; ASP-5 (ORF: F21F8.3), AAC47991; ASP-6 (ORF: F21F8.7), AAC47996; Y39B6A.23, CAD31820; Y39B6A.24, CAD31821; Y39B6A.22, CAD31819; ZK384.3, T26490; Y39B6A.21, CAD31818; F59D6.3, T31770; F59D6.2, T31771; F21F8.4, T29411; C11D2.2, T32964; F21F8.6, T29409; F28A12.4, T28954; C15C8.3, T19309; F21F8.2, T29413 and K10C2.3, T25812. Accession numbers of the calpain sequences are: Human hm-Calpain, NP_001739; human hCalpain1, NP_008989; mouse mCalpain1, NP_031626; Drosophila dCalpain, CAA55297; Drosophila dSol, P27398; CLP-1 (ORF: C06G4.2), P34308; CLP-2 (ORF: T04A8.16), T24431; CLP-3 (ORF: Y47H10A.1), NP_493052; CLP-4 (ORF: Y39A3CL.5), AAF60521; TRA-3/CLP-5 (ORF: LLC1.1), AAB60256; CLP-6 (ORF: Y77E11A.10), AAF36079; CLP-7 (ORF: Y77E11A.11), AAF36080; T11A5.6, T24825; W05G11.4, T32871; F44F1.3, T22177; F47F6.5, NP_493933; F44F1.1, T22175; M04F3.4, NP_491447; Y53H1B.6, NP_492903; H25P06.4, T23122; W04A4.4, T26132;and T21H3.3, T31737.

BLAST searches[1] were performed with the National Center for Biotechnology Information web based servers (NCBI; http://www.ncbi.nlm.nih.gov/BLAST/). Multiple sequence alignments were

generated using the ClustalW algorithm[2] and displayed with SeqVu (The Garvan Institute of Medical Research, Sydney, Australia) and BoxShade 3.21 (EMBNet; http://www.ch.embnet.org/software/BOX_form.html). Dendrograms showing phylogenetic distance relationships among aspartyl and calpain proteases were constructed with the neighbor-joining method[3], based on pair wise distance estimates of the expected number of amino acid replacements per site, and visualized with TreeView (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html). Prosite searches for conserved motifs in protein sequences were performed with the Expert Protein Analysis System (ExPASy) proteomics web based server at the Swiss Institute of Bioinformatics (SIB; http://www.expasy.ch/tools/scanprosite/).

Supplementary References

1. Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402. (1997).
2. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-4680. (1994).
3. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406-425. (1987).
4. Tcherepanova, I., Bhattacharyya, L., Rubin, C. S. & Freedman, J. H. Aspartic proteases from the nematode *Caenorhabditis elegans*. Structural organization and developmental and cell-specific expression of *asp-1*. *J Biol Chem* **275**, 26359-26369 (2000).
5. Moldoveanu, T. et al. A Ca(2+) switch aligns the active site of calpain. *Cell* **108**, 649-660 (2002).

Legends

Figure 5. ClustalW-generated, multiple alignment of nematode aspartyl proteases. Representative members from other species are also included for comparison (hCathepsinD, human; mCathepsinD, mouse; dRE02351p, Drosophila; yPep4p, yeast). The symbol # in the consensus line denotes the conserved lysosomal targeting, N-glycosylation site (asparagine 71 in ASP-1), which is absent in ASP-3 and ASP-4, as well as the potential N-glycosylation site in ASP-3 and ASP-4, common in non-lysosomal cathepsin E proteases. Of the 20 predicted nematode proteins included in the alignment, at least 6 appear to be expressed (ASP-1 to 6)[4].

Figure 6. Dendrogram showing phylogenetic relationships among aspartyl proteases aligned in Figure 5. ASP-3 and ASP-4, shown in bold, are required for neurodegeneration and cluster within a distinct branch that also includes mammalian proteases. Red branches indicate nematode proteins (green for ASP-3 and ASP-4). Non-nematode proteins are in blue. Branch lengths represent the expected number of amino acid replacements per site (0.1 in the scale bar).

Figure 7. *mec-4(d)*-induced neurodegeneration is suppressed in aspartyl protease-deficient genetic backgrounds and is restored by elevated expression of ASP-3 and ASP-4. **a**, Images of vacuolated touch receptor neurons (indicated by arrows) in animals of the indicated genetic background. *mec-4(d)* L1 larvae exhibit large prominent vacuoles in mid-body and tail areas corresponding to dying ALM and PLM touch receptor neurons respectively. In *cad-1(j1);mec-4(d)*, *daf-4(e1364);mec-4(d)* and *unc-52(su250);mec-4(d)* double mutants, where neurodegeneration is significantly suppressed, not only is the number of vacuoles lower, but incident vacuoles are also qualitatively different; they are smaller and less pronounced than in *mec-4(d)* single mutants. **b**, Introduction of ASP-3 and ASP-4 restores neurodegeneration in

daf-4(e1364);mec-4(d) and *unc-52(su250);mec-4(d)* double mutants. Bars signify degenerating touch receptors at L1 stage. *Ex*[*asp-3*] and *Ex*[*asp-4*] restored cell death in *daf-4(e1364);mec-4(d)* and *unc-52(su250);mec-4(d)* double mutants (n=100, $P<0.001$, unpaired *t*-test). Values for *mec-4(d)*, *daf-4(e1364);mec-4(d)* and *unc-52(su250);mec-4(d)* controls are those in Fig. 1a and are included here for comparison.

Figure 8. Multiple alignment of nematode calpain proteases. Representative members from other species are also included for comparison (hm-Calpain, human; hCalpain1, human; mCalpain1, mouse; dCalpain, Drosophila; dSol, Drosophila *Small optic lobes* protease). The symbol # in the consensus line denotes the conserved catalytic residues (glutamine 269, cysteine 275, histidine 431, asparagine 455 and tryptophan 457 in CLP-1)[5]. Of the 17 predicted nematode proteins included in the alignment, 7 show significant identity to mammalian calpains over their entire length and have been annotated as such in WormBase (http://www.wormbase.org).

Figure 9. Dendrogram showing distance relationships among calpain proteases aligned in Figure 8. CLP-1 and TRA-3, shown in bold, are involved in neurodegeneration. Red branches indicate nematode proteins (green for CLP-1 and TRA-3). Non-nematode proteins are in blue. Branch lengths represent the expected number of amino acid replacements per site (0.1 in the scale bar).

Figure 10. Details of protease gene expression in *C. elegans*. Aspartyl proteases ASP-3 and ASP-4 appear to be highly expressed in the intestine of adult animals and at relatively lower levels in many other tissues including body wall muscles, the hypodermis, neurons, the uterus and others. Expression of both genes is detectable in two-fold stage embryos. Calpain *clp-1* is strongly expressed in the nervous system, muscles and the hypodermis. Expression is also

observed in the intestine, the canal cell and the pharynx. **a**, images of animals carrying ASP-3::GFP. **b**, images of animals carrying ASP-4::GFP. **c**, images of animals carrying p_{clp-1}GFP. Various cells and organs are indicated by arrows.

Table 1. The conserved catalytic pentad of amino acid residues in calpains (glutamine 269, cysteine 275, histidine 431, asparagine 455 and tryptophan 457 in CLP-1)[5] is intact in CLP-1, CLP-2, TRA-3, CLP-6 and CLP-7 (shown in bold). The rest of the 17 nematode sequences are missing one or more specific residues (+/-: presence/absence of catalytic residues).

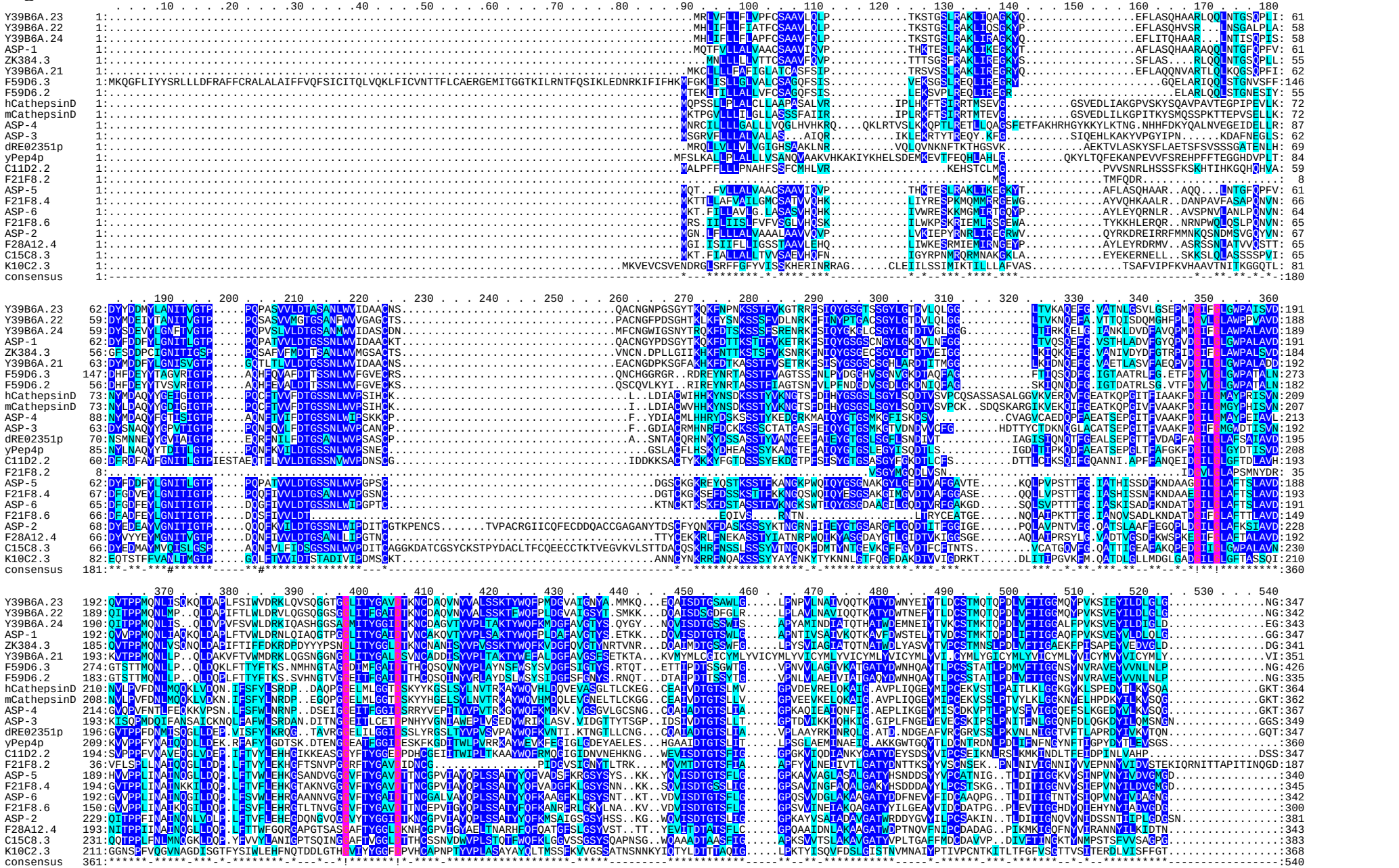
		Q269	C275	H431	N455	W457
1	**CLP-1**	+	+	+	+	+
2	**CLP-2**	+	+	+	+	+
3	CLP-3	-	-	-	-	+
4	CLP-4	+	+	-	-	-
5	**TRA-3**	+	+	+	+	+
6	**CLP-6**	+	+	+	+	+
7	**CLP-7**	+	+	+	+	+
8	F44F1.3	+	+	+	+	-
9	F44F1.1	+	+	+	+	-
10	F47F6.5	+	+	+	-	-
11	H25P06.4	-	-	+	-	-
12	M04F3.4	-	-	-	-	-
13	T11A5.6	-	+	+	+	+
14	T21H3.3	-	-	-	-	-
15	W04A4.4	-	-	+	-	-
16	W05G11.4	-	+	+	+	+
17	Y53H1B.6	-	-	+	-	-

Table 2. Neurodegeneration inflicted by overexpression of *asp-4* cannot be bypassed by calpain protease deficiency.

Strain	***%[a]***
Ex[p$_{mec-4}$ASP-4]	13.9±1.4
Ex[p$_{mec-4}$ASP-4] *clp-1(RNAi)*	13.2±1.8
Ex[p$_{mec-4}$ASP-4] *tra-3(RNAi)*	12.8±1.6
Ex[p$_{mec-4}$ASP-4] *clp-1(RNAi);tra-3(RNAi)*	13.5±2.1

[a] Percentage (± standard deviation) of animals with vacuolated touch receptor neurons at L1 stage.

Figure 5



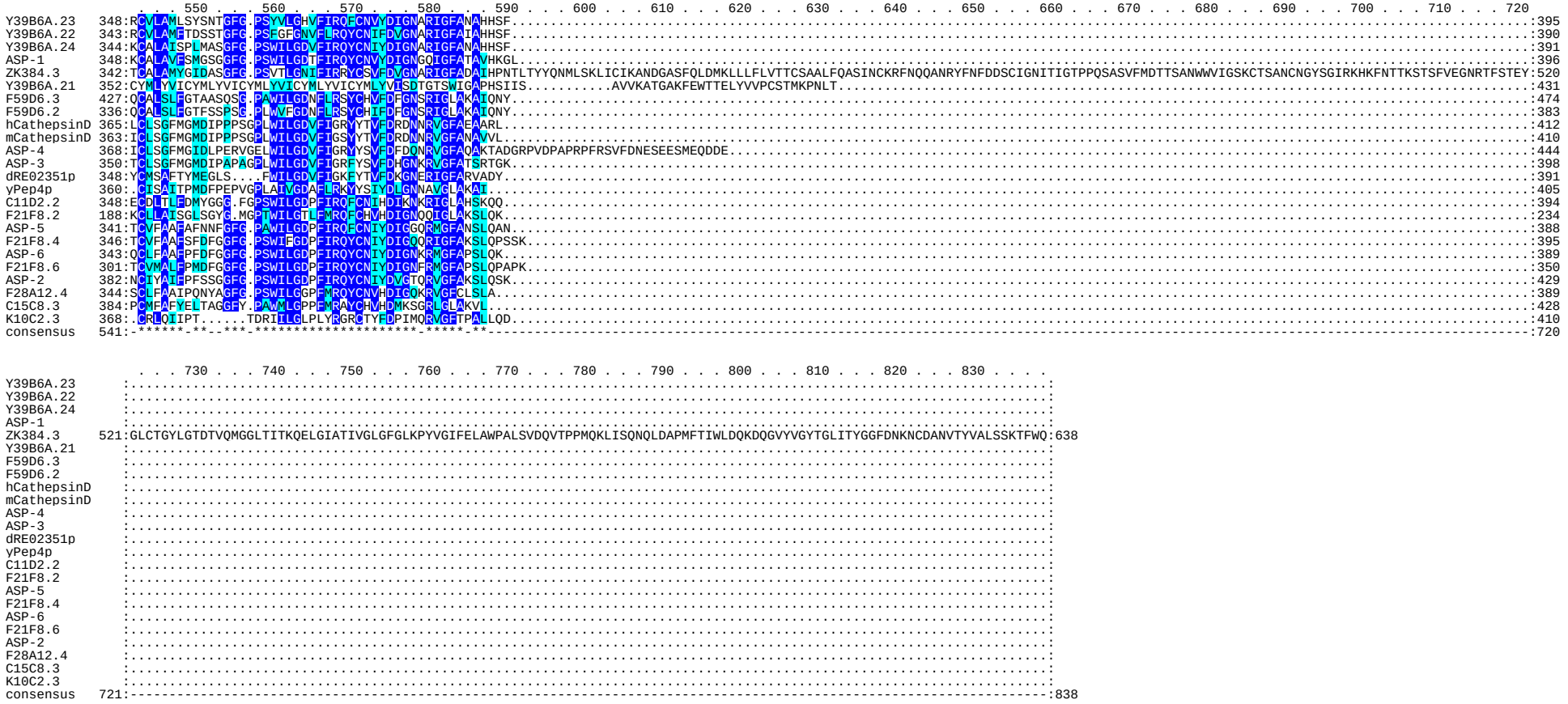


Figure 6

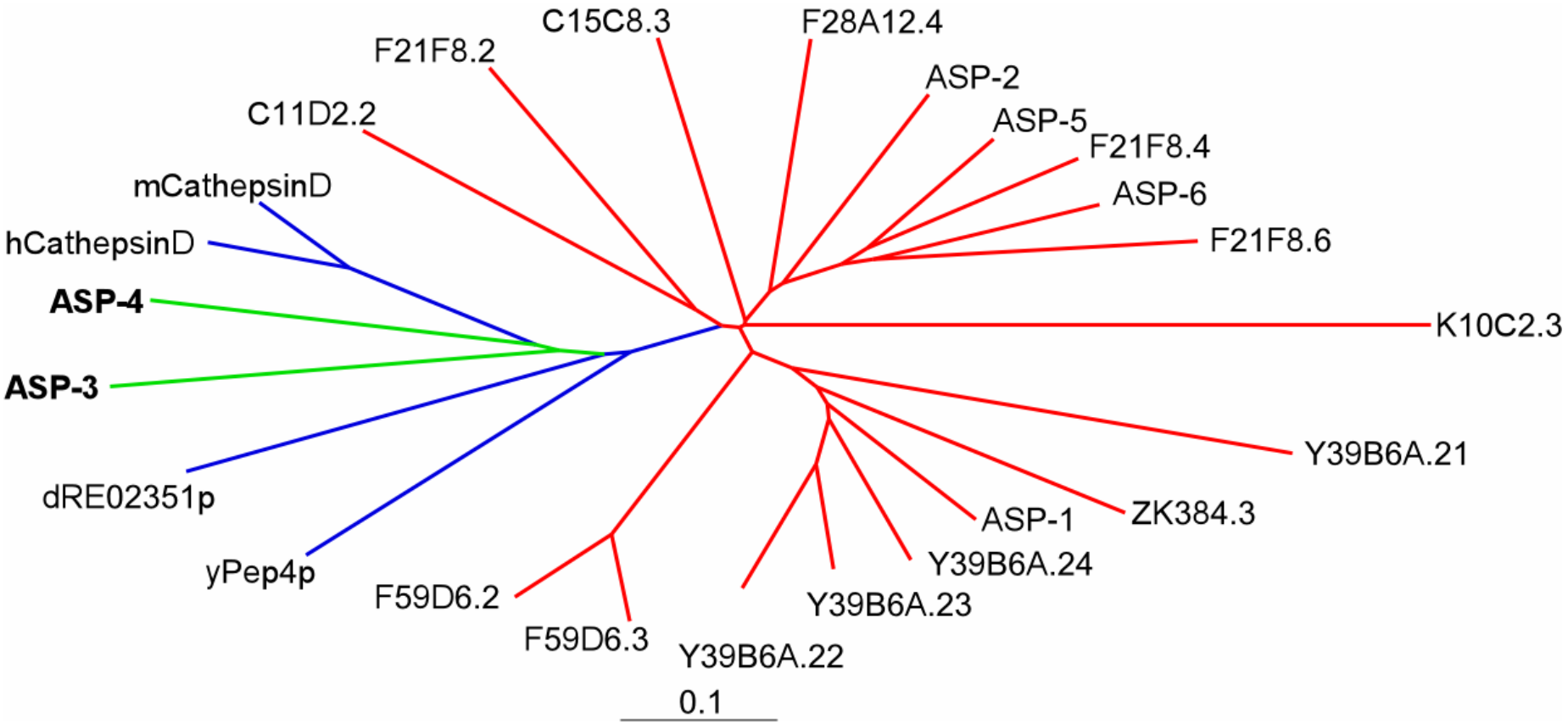

C15C8.3
F28A12.4
F21F8.2
ASP-2
C11D2.2
ASP-5
F21F8.4
mCathepsinD
ASP-6
hCathepsinD
F21F8.6
ASP-4
K10C2.3
ASP-3
Y39B6A.21
dRE02351p
ASP-1
ZK384.3
Y39B6A.24
yPep4p
F59D6.2
Y39B6A.23
F59D6.3
Y39B6A.22
0.1

Figure 7

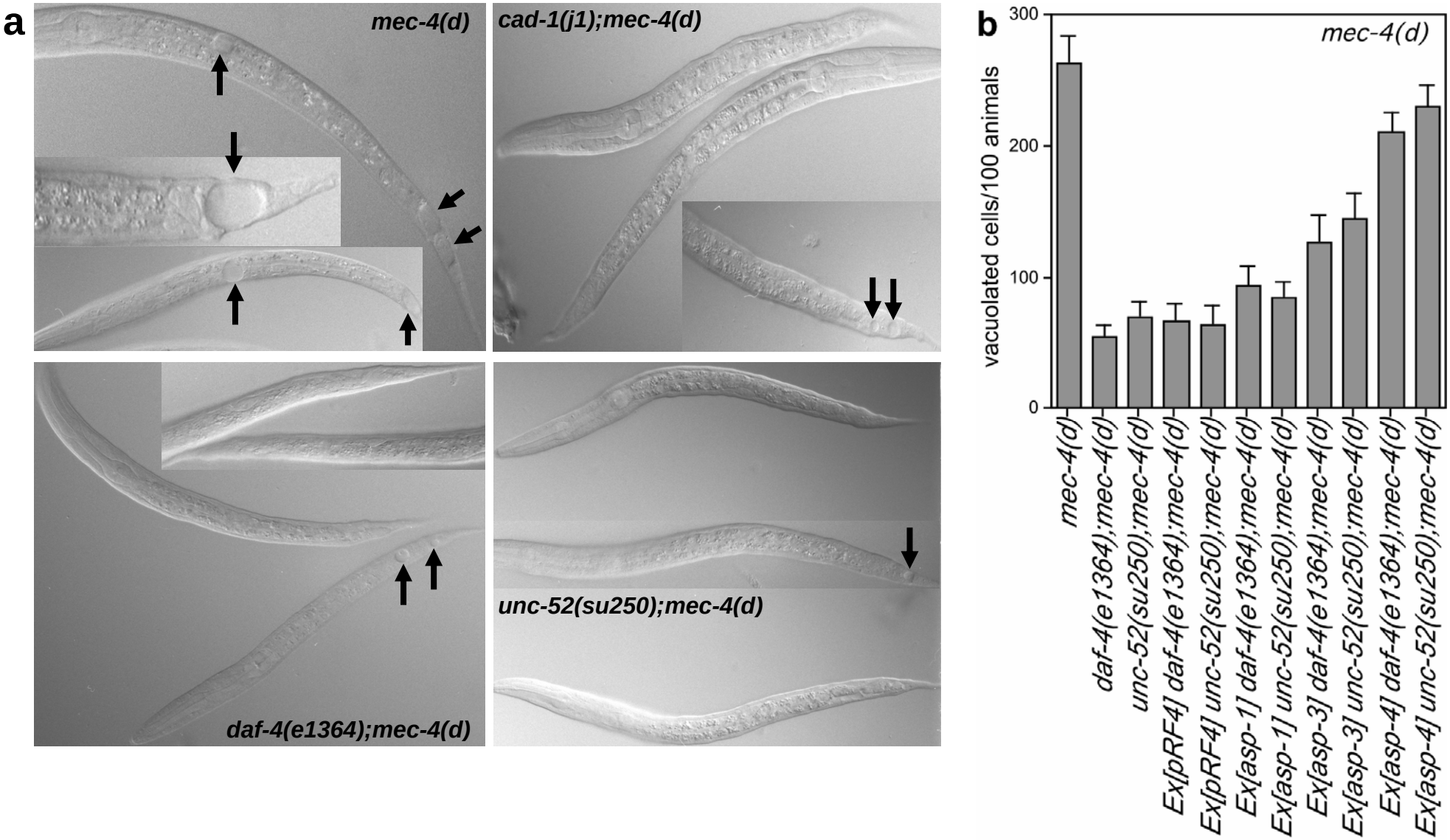
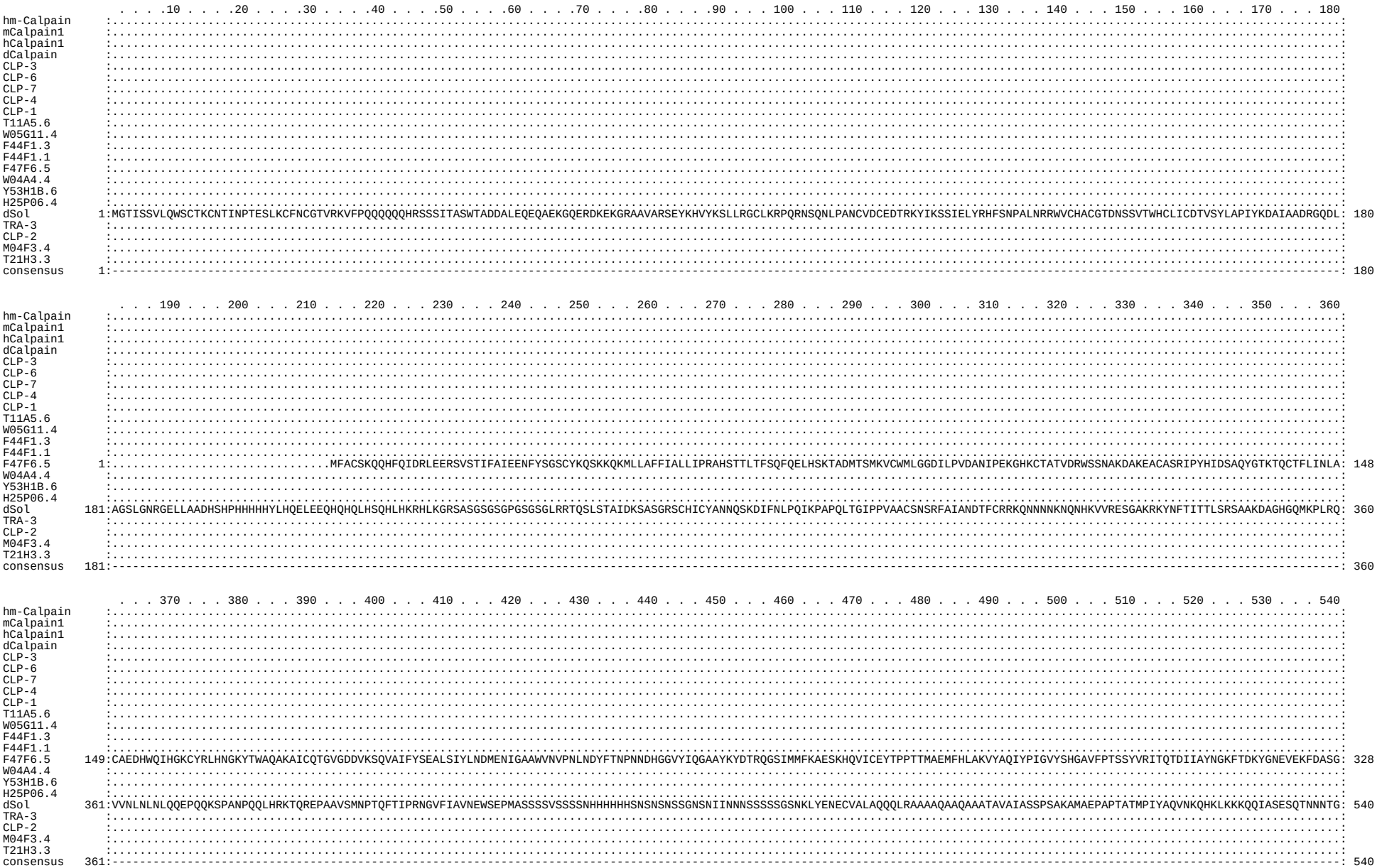
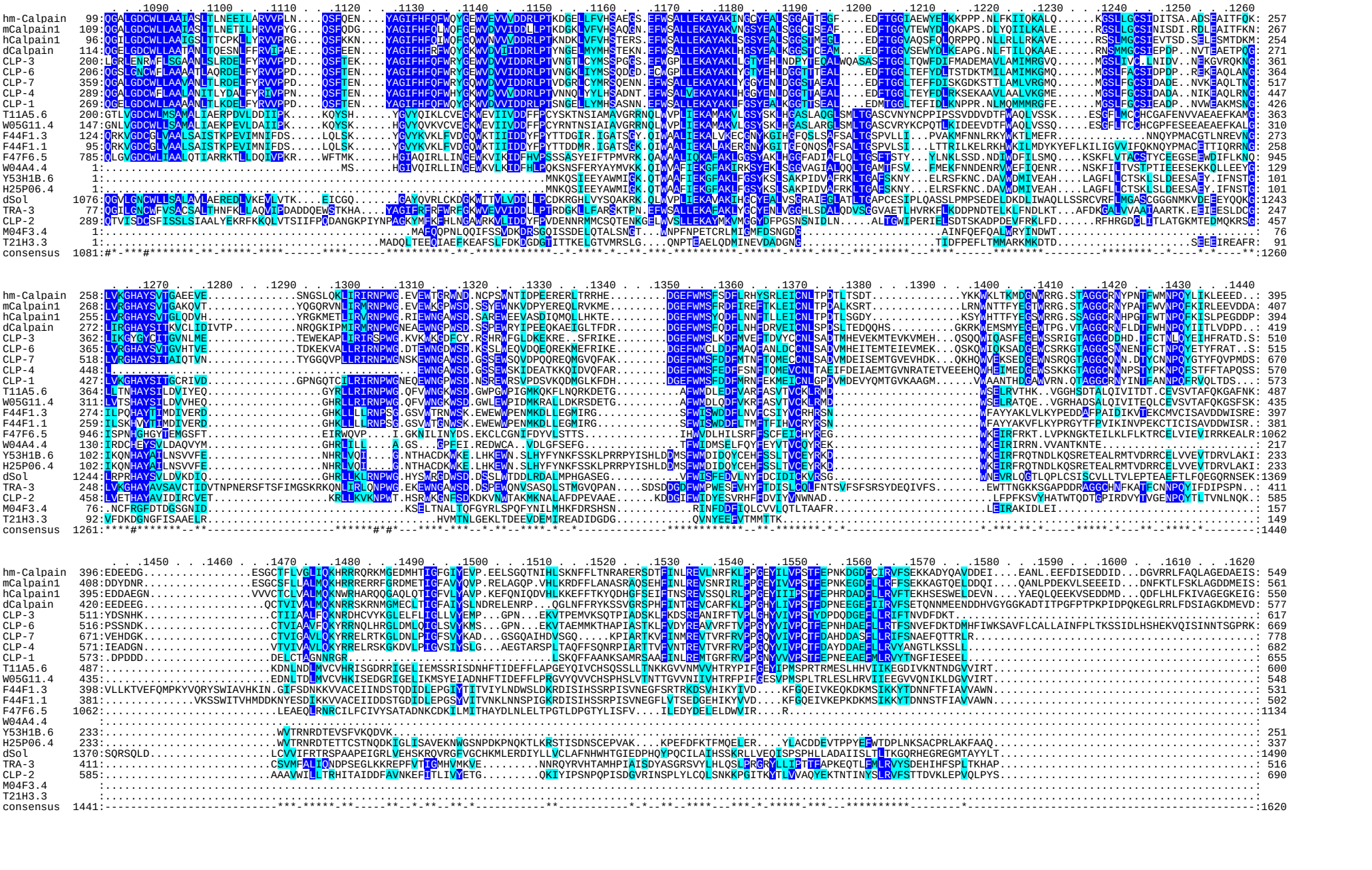

a
mec-4(d)
cad-1(j1);mec-4(d)
daf-4(e1364);mec-4(d)
unc-52(su250);mec-4(d)
b
mec-4(d)
vacuolated cells/100 animals
300
200
100
0
mec-4(d)
daf-4(e1364);mec-4(d)
unc-52(su250);mec-4(d)
Ex[pRF4] daf-4(e1364);mec-4(d)
Ex[pRF4] unc-52(su250);mec-4(d)
Ex[asp-1] daf-4(e1364);mec-4(d)
Ex[asp-1] unc-52(su250);mec-4(d)
Ex[asp-3] daf-4(e1364);mec-4(d)
Ex[asp-3] unc-52(su250);mec-4(d)
Ex[asp-4] daf-4(e1364);mec-4(d)
Ex[asp-4] unc-52(su250);mec-4(d)

Figure 8


hm-Calpain
mCalpain1
hCalpain1
dCalpain
CLP-3
CLP-6
CLP-7
CLP-4
CLP-1
T11A5.6
W05G11.4
F44F1.3
F44F1.1
F47F6.5
W04A4.4
Y53H1B.6
H25P06.4
dSol
TRA-3
CLP-2
M04F3.4
T21H3.3
consensus
MGTISSVLQWSCTKCNTINPTESLKCFNCGTVRKVFPQQQQQQHRSSSITASWTADDALEQEQAEKGQERDKEKGRAAVARSEYKHVYKSLLRGCLKRPQRNSQNLPANCVDCEDTRKYIKSSIELYRHFSNPALNRRWVCHACGTDNSSVTWHCLICDTVSYLAPIYKDAIAADRGQDL
MFACSKQQHFQIDRLEERSVSTIFAIEENFYSGSCYKQSKKQKMLLAFFIALLIPRAHSTTLTFSQFQELHSKTADMTSMKVCWMLGGDILPVDANIPEKGHKCTATVDRWSSNAKDAKEACASRIPYHIDSAQYGTKTQCTFLINLA
AGSLGNRGELLAADHSHPHHHHHYLHQELEEQHQHQLHSQHLHKRHLKGRSASGSGSGPGSGSGLRRTQSLSTAIDKSASGRSCHICYANNQSKDIFNLPQIKPAPQLTGIPPVAACSNSRFAIANDTFCRRKQNNNNKNQNHKVVRESGAKRKYNFTITTLSRSAAKDAGHGQMKPLRQ
CAEDHWQIHGKCYRLHNGKYTWAQAKAICQTGVGDDVKSQVAIFYSEALSIYLNDMENIGAAWVNVPNLNDYFTNPNNDHGGVYIQGAAYKYDTRQGSIMMFKAESKHQVICEYTPPTTMAEMFHLAKVYAQIYPIGVYSHGAVFPTSSYVRITQTDIIAYNGKFTDKYGNEVEKFDASG
VVNLNLNLQQEPQQKSPANPQQLHRKTQREPAAVSMNPTQFTIPRNGVFIAVNEWSEPMASSSSVSSSSNHHHHHHSNSNSNSSGNSNIINNNSSSSSGSNKLYENECVALAQQQLRAAAAQAAQAAATAVAIASSPSAKAMAEPAPTATMPIYAQVNKQHKLKKKQQIASESQTNNNTG
11



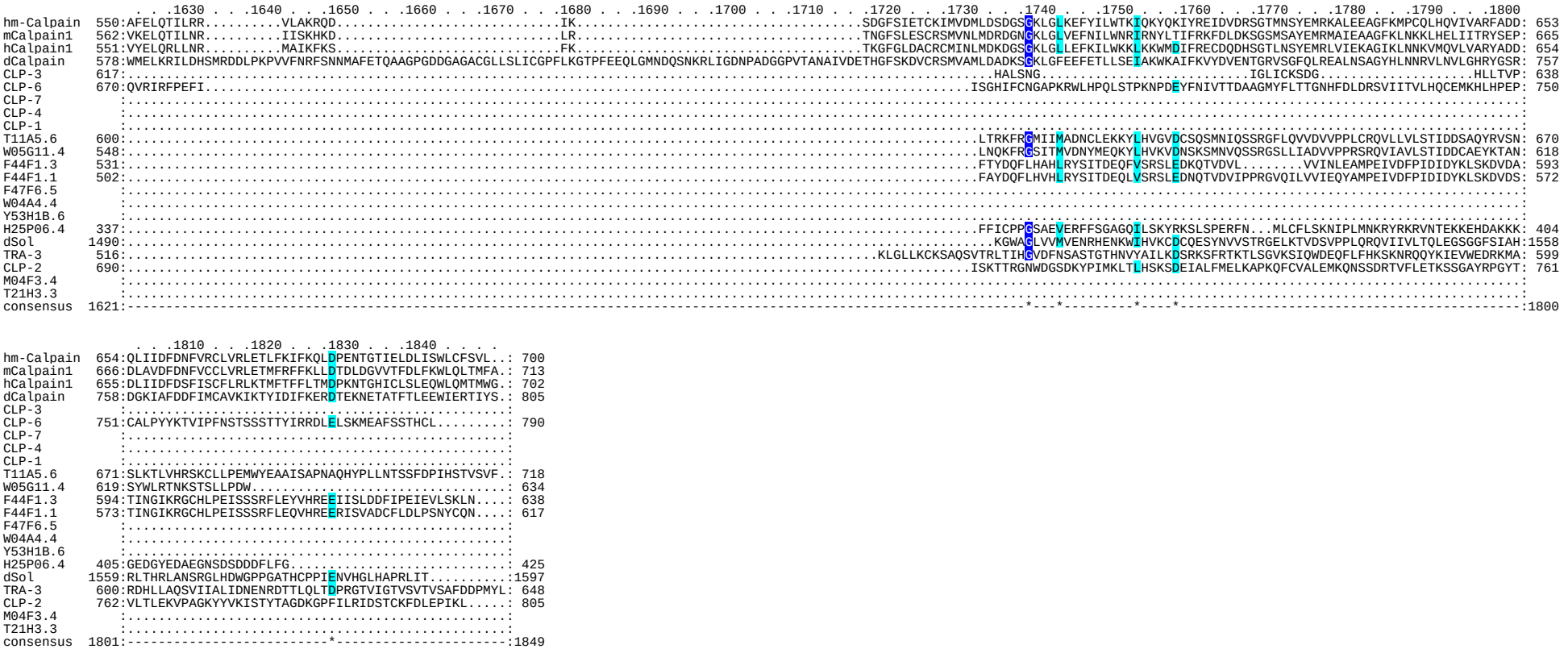


Figure 9

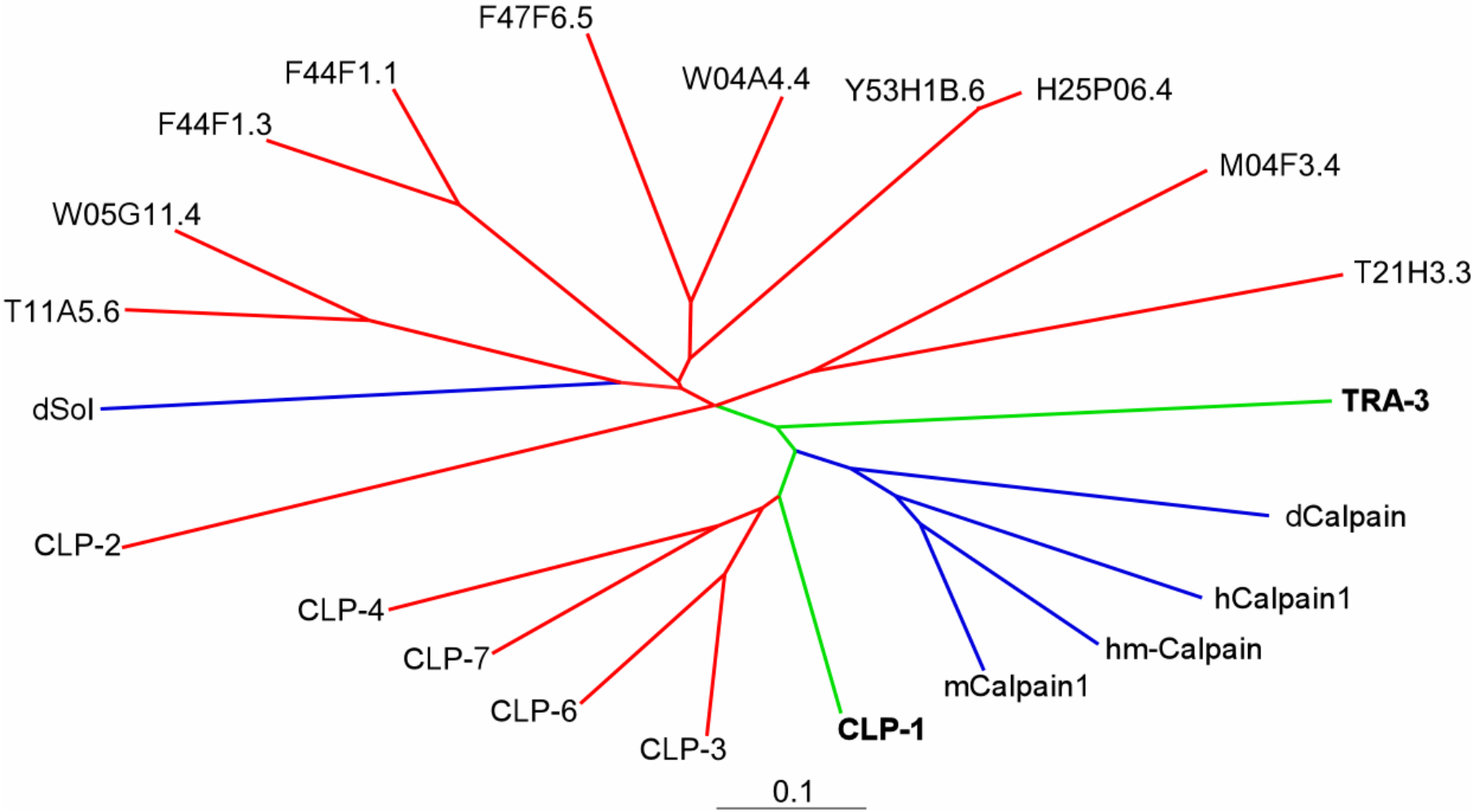
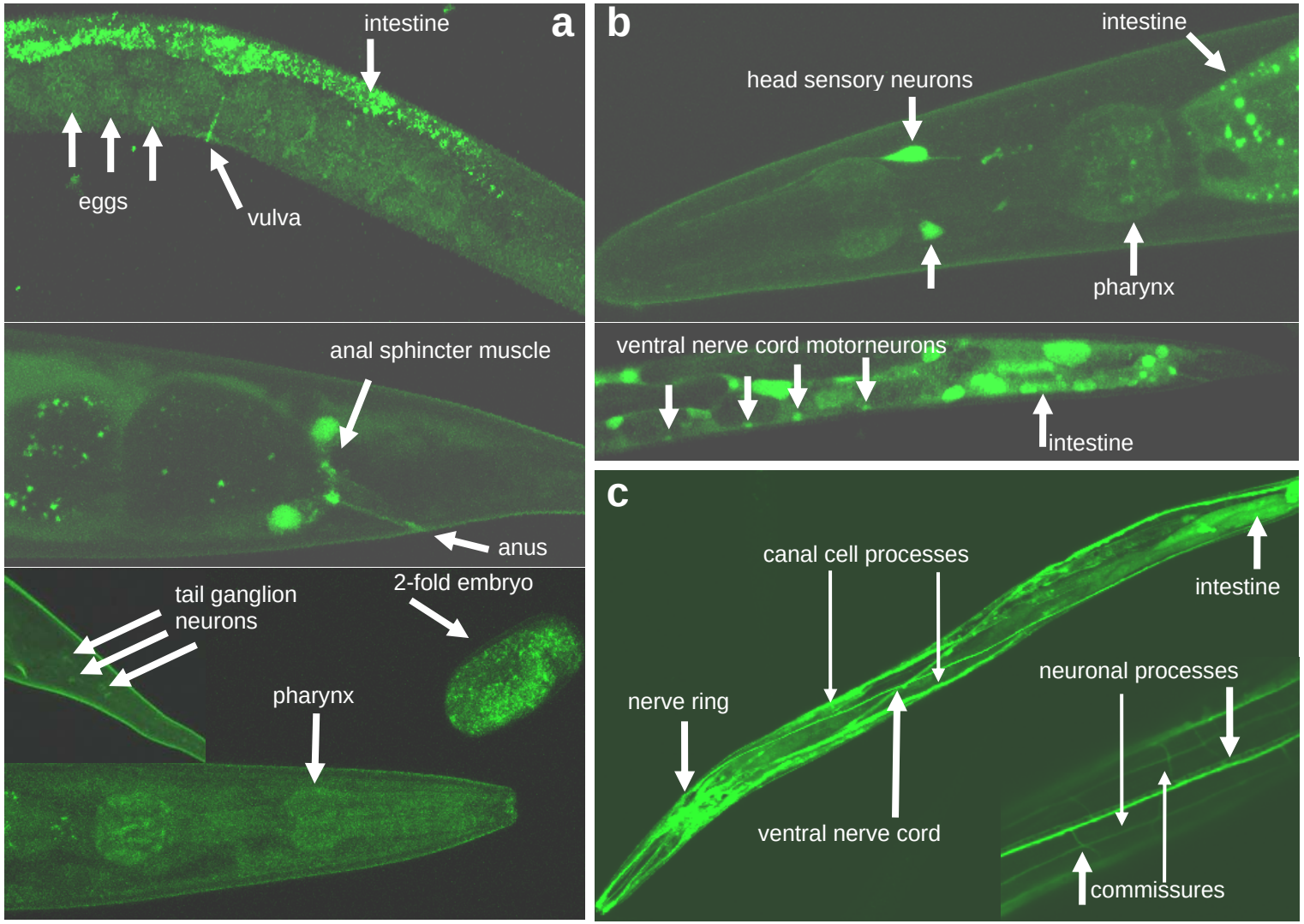

F47F6.5
F44F1.1
W04A4.4
Y53H1B.6
H25P06.4
F44F1.3
M04F3.4
W05G11.4
T21H3.3
T11A5.6
dSol
TRA-3
dCalpain
CLP-2
hCalpain1
CLP-4
hm-Calpain
CLP-7
mCalpain1
CLP-6
CLP-1
CLP-3
0.1

Figure 10


a
intestine
eggs
vulva
anal sphincter muscle
anus
tail ganglion neurons
2-fold embryo
pharynx
b
intestine
head sensory neurons
pharynx
ventral nerve cord motorneurons
intestine
c
canal cell processes
intestine
nerve ring
neuronal processes
ventral nerve cord
commissures