PORTLAND PRESS

Research Article

Uncovering the essential roles of glutamate carboxypeptidase 2 orthologs in *Caenorhabditis elegans*

Lucie Panska[1,2,*], Stepanka Nedvedova[1,3,*], Vojtech Vacek[1], Daniela Krivska[1,3], Lukas Konecny[1,4], Filip Knop[5], Zsofia Kutil[6], Lubica Skultetyova[6], Adrian Leontovyc[7], Lenka Ulrychova[4,7], Judy Sakanari[8], Masako Asahina[9], Cyril Barinka[6], Marie Macurkova[5] and Jan Dvorak[1,2,7]

[1]Department of Zoology and Fisheries, Center of Infectious Animal Diseases, Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences Prague, Kamycka 129, Prague 165 00, Czech Republic; [2]Faculty of Environmental Sciences, Center of Infectious Animal Diseases, Czech University of Life Sciences in Prague, Kamycka 129, Prague 165 00, Czech Republic; [3]Department of Chemistry, Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences Prague, Kamycka 129, Prague 165 00, Czech Republic; [4]Department of Parasitology, Faculty of Science, Charles University, Vinicna 7, Prague 2 128 00, Czech Republic; [5]Department of Cell Biology, Faculty of Science, Charles University, Vinicna 7, Prague 2 128 00, Czech Republic; [6]Laboratory of Structural Biology, Institute of Biotechnology of the Czech Academy of Sciences, BIOCEV, Prumyslova 595, Vestec 252 50, Czech Republic; [7]Institute of Organic Chemistry and Biochemistry, The Czech Academy of Sciences, Flemingovo n. 2, Prague 160 00, Czech Republic; [8]Department of Pharmaceutical Chemistry, University of California, San Francisco, 1700 4th Street, CA 94158, USA.; [9]Department of Physiology, University of California, San Francisco, 600 16th Street, CA 94143, U.S.A.

Correspondence: Lucie Panska (panska@fzp.czu.cz)

OPEN ACCESS

Human glutamate carboxypeptidase 2 (GCP2) from the M28B metalloprotease group is an important target for therapy in neurological disorders and an established tumor marker. However, its physiological functions remain unclear. To better understand general roles, we used the model organism *Caenorhabditis elegans* to genetically manipulate its three existing orthologous genes and evaluate the impact on worm physiology. The results of gene knockout studies showed that *C. elegans* GCP2 orthologs affect the pharyngeal physiology, reproduction, and structural integrity of the organism. Promoter-driven GFP expression revealed distinct localization for each of the three gene paralogs, with *gcp-2.1* being most abundant in muscles, intestine, and pharyngeal interneurons, *gcp-2.2* restricted to the phasmid neurons, and *gcp-2.3* located in the excretory cell. The present study provides new insight into the unique phenotypic effects of GCP2 gene knockouts in *C. elegans*, and the specific tissue localizations. We believe that elucidation of particular roles in a non-mammalian organism can help to explain important questions linked to physiology of this protease group and in extension to human GCP2 involvement in pathophysiological processes.

*These authors contributed equally to this work.

Introduction

The M28 protease family is an intensively studied group of proteolytic enzymes mainly due to its significance in human pathologies (MEROPS) [1]. Most of the research attention is devoted to the M28B subfamily because of its importance in pathological processes in humans. However, members of this subfamily are ubiquitously expressed across all phyla, from unicellular organisms to plants and animals. The unique feature of this di-zinc metalloprotease group is the presence of enzymes with either amino- or carboxypeptidase activity [1]. In the human genome, five genes encoding M28B paralogs possessing carboxy- or amino-peptidase activities were identified to date [1–3]. Interestingly, transferrin receptors with non-enzymatic function (transferrin receptor protein and transferrin receptor 2 protein) are also structurally related to the common ancestor together with M28B family proteases [4]. The most studied human protease from the M28B subfamily is glutamate carboxypeptidase 2 (GCP2 or GCPII; EC 3.4.17.21),

which is also known as a prostate-specific membrane antigen (PSMA), N-acetylated-α-linked acidic dipeptidase (NAALADase), and folate hydrolase 1 (FOLH1) [4,5].

In humans, GCP2 is a membrane-bound metallopeptidase composed of zinc-dependent homodimers [6]. This transmembrane metalloprotease comprises three structural domains: protease, apical, and C-terminal, which all contribute to the architecture of the bimetallic active site [6]. In addition to zinc ions, human GCP2 also requires Ca^{2+} ions for proper folding, stability, and enzymatic activity [7]. GCP2 is primarily expressed in the nervous system (astrocytes and Schwann cells), kidney [8], and jejunal brush membranes. Lower expression levels were also reported in salivary and lacrimal glands, prostate, heart, pancreas, bladder, skin, breast, liver, lung, colon, and testis [9,10]. Despite well-described expression profiles in human organs, there is an absence of data on GCP2 physiology in most tissues, except for the nervous system and small intestine. In the nervous tissue, GCP2 contributes to the intercellular communication between neurons and glial cells. It achieves this by breaking down N-acetyl-aspartyl-glutamate (NAAG), which is the most prevalent neuropeptide in the mammalian brain. In the digestive system, GCP2 assumes a crucial role in the absorption of dietary folates. By cleaving the C-terminal poly-glutamylated tails of folates, GCP2 enables the absorption of folate mono-glutamate into the bloodstream [11]. In addition to their function as hydrolases/peptidases, several reports allude to a non-proteolytic role(s) of GCP2, such as the roles associated with the anaphase-promoting complex in prostate cancer cells or to activation of the NF-κB signaling pathway in cell proliferation [12]. Human GCP2 is an attractive target for pharmacological interventions in neurodegenerative diseases and cancer. Understanding and dissecting potentially diverse biological functions of the enzyme is, therefore, desirable. Consistent with the role of GCP2 (NAAG hydrolysis), changes in the NAAG levels and/or changes of GCP2 enzymatic activities correlate with pathologic conditions, including Alzheimer's disease, Huntington's disease, amyotrophic lateral sclerosis [13,14], epilepsy, schizophrenia [15], and stroke [16]. The inhibition of GCP2 activity has already been proven as an effective neuroprotective therapy against ischemic brain injury in animal models [17]. GCP2 expression has also been found in the tumors of the neovasculature [18], kidney [19], digestive tract [20] and in various subtypes of bladder cancer [21]. Other members of the M28B family in human are PSMA-L, glutamate carboxypeptidase 3 (GCP3), NAALADase L, and NAALADase L2. Their physiological roles are either partially described or not clarified yet [6,22]. However, NAALADase L was recently characterized as an enzyme with aminopeptidase activity expressed in intestinal tissue [5], and GCP3 was shown to have overlapping substrate specificity and inhibition profile with GCP2 [23,24].

The studies of the physiological functions of this protease group are complicated for many reasons. The most notable obstacle is the presence of several paralogs in a single organism as seen in mammals including humans, therefore, more phylogenetically basal model organisms may offer significant advantages. During evolution, a wide spectrum of M28B proteases evolved by duplications and subsequent speciation occurred as evident from the example of the aforementioned human paralogs [4]. In contrast with vertebrates, invertebrates (including nematodes) possess only a single or few genes coding M28B metalloproteases [25], thus offering opportunities to study ancestral proteases before their diversification. So far, none of the GCP2 orthologs from lower basal organisms has been intensively studied. For that reason, we selected the free-living nematode *Caenorhabditis elegans* as a model to investigate the physiological functions of GCP2 orthologs. *C. elegans* serves as an excellent laboratory model in a variety of research areas due to a well-characterized genome, relatively simple body organization, uncomplicated maintenance in laboratory conditions, and the existence of a host of experimental tools to study its physiology, including transgenic approaches [26]. We focused on all three genes coding subfamily M28B orthologs named CeGCP2.1, CeGCP2.2, and CeGCP2.3 encoded by *gcp-2.1*, *gcp-2.2*, and *gcp-2.3* genes, respectively [27]. Their nomenclature, however, is without any known relevance to the substrate preferences or the enzymatic mode of action and it is solely based on existing gene annotations [28].

In our study, we report the most pronounced phenotypes observed in a member of this metalloprotease group in any animal model to date. After performing a wide range of phenotypic studies, we identified several unique phenotypes that, together with their localization and expression, led to the unexpected findings that all three genes acquired distinct and unique functions.

Materials and methods
Sequence alignment and homology modeling
The amino acid sequence of hsGCPII (NCBI ID: NP_004467.1), and *C. elegans* orthologs deposited in the UniProtKB: GCP2.1 under sequence ID: P91406-1 coded by *gcp-2.1* (R57.1) gene; GCP2.2; under sequence ID: Q93332 coded by *gcp-2.2* (C35C5.2) gene; and GCP2.3 translated from the sequence inaccurately labeled as a pseudogene *gcp-2.3* (C35C5.11) from WormBase under sequence ID: C35C5.11, were used for the sequence alignment and homology

modeling. The amino acid sequences of hsGCPII and all three *C. elegans* orthologs were aligned using Jalview Version 2 [29] and subsequently analyzed for sequence identity and similarity with Sequence Manipulation Suite: Ident and Sim tool [30].

For homology modeling, the 5ELY GCP2 structure determined to high (1.81 Å) resolution limits was selected as a template [31]. The Modeller 9.23 software was used to construct the target-template sequence alignment and to generate five 3D homology models. The best model for each enzyme was selected based on discrete optimized protein energy (DOPE) scores [32]. Finally, the homology models of CeGCP2 variants and hsGCPII structure (PDB code: 3BXM) were superimposed in PyMol and analyzed by visual inspection.

The amino acid sequences of hsGCPII and all three *C. elegans* orthologs were aligned using Jalview Version 2 [29] and subsequently analyzed for sequence identity and similarity with Sequence Manipulation Suite: Ident and Sim tool [30].

Phylogenetic analysis

For phylogenetic analysis, protein sequences of nematode paralogs of the M28B peptidase family were downloaded from the MEROPS database [1] and aligned with sequences acquired from NCBI and WormBase databases [28,33]. The alignment was constructed by MAFFT v. 7.222 and automatically trimmed by BMGE v. 1.12 software using the blosum62 scoring matrix. The phylogenetic tree was constructed in IQ-TREE v. 1.6.1 using the best-fitting model (LG4M) selected by ModelFinder. GCP2 of *Homo sapiens* was used as the outgroup. The resulting tree was graphically enhanced in Figtree v. 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree).

Maintenance of *C. elegans* and backcrossing

All *C. elegans* nematodes were grown and maintained for several generations at 18°C under standard conditions [34]. Worms were kept on nematode growth medium (NGM) plates and fed by OP50 *Escherichia coli* strain, which was performed according to slightly modified instructions in the WormBook [34]. Three *gcp-2* mutant worm strains RB1055 - *gcp-2.1(ok1004)* (CGC), gcp-2.2 - *gcp-2.2(tm6541)* and C35C5.11 - *gcp-2.3(tm5414)* [35], and the wild-type (N2) were used in the present study (Supplementary Table S1). The mutation of the three paralog genes represents loss-of-function mutation.

Before all experimental procedures, knockout worms were backcrossed eight times with N2 males with the consequent recovery of mutant homozygotes to remove possible off-target and other random genetic mutations. Backcrossing was performed according to WormBook protocols [36].

DNA for control genotyping was isolated from a single hermaphrodite after the offspring production. The worms were dissolved in 10 µl of QuickExtract™ DNA Extraction Solution (Biosearch Technologies). Final analyses of homozygous or heterozygous lines were performed by PCR reaction using Ready To Use PCR MasterMix (Central European Biosystems) with forward and reverse primer for the target gene (Supplementary Table S2), and resolved by agarose gel electrophoresis (see Supplementary Figure S1 and Table S3 and detailed protocols therein).

Gene expression assay

Adult worms were washed down from the growing plate with PBS buffer (Sigma-Aldrich) and spun down at 3000 ***g***. Worms were resuspended in 500 µl of TRIzol™ Reagent (ThermoFisher Scientific) and RNA was isolated with RNeasy Mini Kit (Qiagen), according to manufacturer's instructions. Any remaining DNA was removed by TURBO™ DNase (ThermoFisher Scientific). The reverse transcription to obtain cDNA from RNA samples was performed with Super-Script™ III (ThermoFisher Scientific) and Oligo(dT)18 primers, according to the manufacturer's protocol. Resulting cDNA samples were stored at −20 °C until use.

The RT-qPCR analysis of all three paralog genes (*gcp-2.1*, *gcp-2.2*, *gcp-2.3*) and a housekeeping gene *C. elegans tubulin* alpha (TBA1) [37] was performed for each sample in biological triplicates by use of LightCycler® 480 SYBR Green I Master mix (Roche) and evaluated by LightCycler® 480 II cycler (Roche), according to the manufacturer's instructions. Gene-specific primers were selected and evaluated as previously published [38].

Behavioral phenotyping

All behavioral observations of knockout worms were carried out at controlled room temperature (21–22°C). To determine the development speed, worms were age-synchronized according to WormBook protocol [34]. Then, synchronized L1 stages were seeded on the NGM plates under standard cultivation conditions and monitored until they reached the young adult stage. Age-synchronized young adults were placed on individual plates each day until they

ceased reproducing. Fecundity was measured by manual counting of viable progeny and unviable embryos [39]. Feeding behavior was visually scored by counting the number of pharyngeal pumps. A single pharyngeal stroke was defined as one synchronous contraction and relaxation cycle of the corpus and terminal bulb [40]. The worms (L4-stage) were placed on plates under a stereomicroscope Nikon SMZ 25 and recorded for 15 s using a Nikon DS-Ri2 camera. Then the pumping rate was quantified for 10 worms from each strain using a hand-held counter in 15 s frames 10 times for each animal. The mean and standard deviation were counted. The experiment was repeated three times with the same results.

Resistance to osmotic stress

In the present study, *C. elegans* worms were subjected to different levels of hyperosmotic stress by placing them on control NGM plates or NGM plates with 300 or 400 mM NaCl. The aim was to detect both acute (15 min) and chronic responses (over several days) to the stress.

To test for the acute response, L3-L4 worms that had been raised on standard NGM plates were transferred to high-salt plates and examined by touch 15 min later. Any attempt to move in response to the touch with a platinum wire was counted as a positive response.

For the chronic response, synchronized young adult worms ($n=20$) were placed on individual NGM plates with 51 mM NaCl, 300 mM NaCl, and 400 mM NaCl and left to reproduce. The original worms were transferred daily to new plates to separate them from their offspring until the end of their reproduction. The number of surviving worms was counted on specific days post-beginning the assay, which were days 1, 4, 6, 7, 11, 13, 14, 15, 18, and 20. Worms were considered dead if they did not respond to repeated prodding with a platinum wire.

Cuticle permeability assay

Permeability assessment involved staining entire worms with Hoechst 33258 [41]. Briefly, synchronized young adult worms were exposed to a 10 µl/ml Hoechst (Sigma-Aldrich) solution for 30 min at room temperature. Subsequently, they underwent multiple rinses with M9 buffer. Visualization was carried out using the DAPI filter on Eclipse Ni microscope (Nikon). Worms exhibiting a minimum of 15 stained nuclei in the pharynx were classified as permeable. A total of 17–23 animals were examined in two separate trials.

Localization studies by GFP promoter-driven expressions

Promoter regions (entire upstream non-coding regions) of three studied genes were cloned into the plasmids optimized for *C. elegans* GFP expression. Upstream promoters areas for *gcp-2.1*, *gcp-2.2*, and *gcp-2.3* genes in sizes 2.893 bp, 1.107 bp, and 911 bp (Supplementary Figure S2), respectively, were PCR amplified by specific primers (Supplementary Table S2) using Phusion® High-Fidelity PCR Master Mix (ThermoFisher Scientific) from the genomic DNA template and sub-cloned by restriction cloning into GFP expression plasmids pPD95_79 and pPD95_81. pPD95_79 and pPD95_81 kindly provided by Andrew Fire (pPD95_79 - Addgene plasmid #1496; http://n2t.net/addgene:1496; RRID:Addgene_1496and pPD95_81 - Addgene plasmid #1497; http://n2t.net/addgene:1497; RRID:Addgene_1497).

The resulting isolated plasmids were used for microinjections as described previously [42]. The used injection mix was composed of GFP-expression plasmid (10 ng/µl), pRF4:rol-6(su1006) co-injection marker (50 ng/µl) [43], and pBluescript (Addgene) empty plasmid (180 ng/µl) was added to reach a total DNA concentration of 100–200 ng/µl in sterile ddH2O. Agarose pads (2%) and Halocarbon oil 700 (Sigma-Aldrich) were used for immobilization, and a standard M9 medium was used as a recovery buffer [42]. Worm phenotype (Rol of animals) for control co-injection marker (pRF4:rol-6(su1006)) was checked under the microscope. Transformed worms were picked and put on a new seeded plate. Transgenic lines were then maintained according to standard protocols (as described above).

In vivo analysis and imaging of tissue-specific GFP expression in *C. elegans*

Transgenic worms were anesthetized using 20mM sodium azide (NaN3) in M9 medium, placed on a 5% agarose pad on glass slides, and gently covered with cover slides. All images were acquired using a spinning disc confocal microscope (Nikon CSU-W1) using a 60× WI objective (CF Plan Apo VC 60XC WI). Representative images are shown as a projection of X-Y z-stacks using the maximum intensity projection type. Images were analyzed using Huygens Professional 19.10 and open-source ImageJ software [44].

Table 1 Overview of the protein sequence identity and similarity of hsGCP2 and CeGCP2 orthologs

	Identity [%]				Similarity [%]			
	hsGCP2	GCP2.1	GCP2.2	GCP2.3	hsGCP2	GCP2.1	GCP2.2	GCP2.3
hsGCP2	100.00	29.73	27.58	24.90	100.00	48.28	43.83	41.90
GCP2.1	29.73	100.00	34.13	28.43	48.28	100.00	52.64	44.54
GCP2.2	27.58	34.13	100.00	57.25	43.83	52.64	100.00	66.20
GCP2.3	24.90	28.43	57.25	100.00	41.90	44.54	66.20	100.00

Staining of phasmids by DiI

DiI staining of phasmids was performed according to the protocol created by Michael Koelle and published on https://www.wormatlas.org/EMmethods/DiIDiO.htm. DiIC12 (1,1'-Didodecyl-3,3,3',3'-Tetramethylindocarbocyanine Perchlorate) (Invitrogen, D383) was used. DiI is the lipophilic fluorescent dye, which can label phasmid neurons PHA and PHB [45].

Electron microscopy

C. elegans worms, including *E. coli* OP50 bacteria, were scraped on to the high-pressure freezing (HPF) specimen carriers. Bacteria served as a filler, minimizing water content and facilitating freezing [46]. Worms on the carriers were frozen in the high-pressure freezer (Leica EM PACT2). Frozen samples collected on metal carriers were transferred under liquid nitrogen into a pre-frozen cryotube containing 1 ml freeze-substitution solution (2% OsO4 in 100% acetone with 1% lecithin) and finally to a freeze substitution unit (Leica EMAFS) for processing. Samples embedded in Epon EmBed812 resin were ultrasectioned (80nm) with an ultramicrotome (Leica EM UC 6) and placed on copper mesh 300 previously coated with uranyl acetate and lead citrate. Sections were finally examined using a transmission electron microscope (JEOL JEM 2100-Plus 200kV).

Statistical analysis

Evaluation of the statistical data of mutant phenotypes (gene expression, fertility, pharyngeal pumping, effect of high NaCl concentrations) was processed by GraphPad Prism 5.

Results

Primary sequence alignments and homology modeling of *C. elegans* M28 metalloproteases
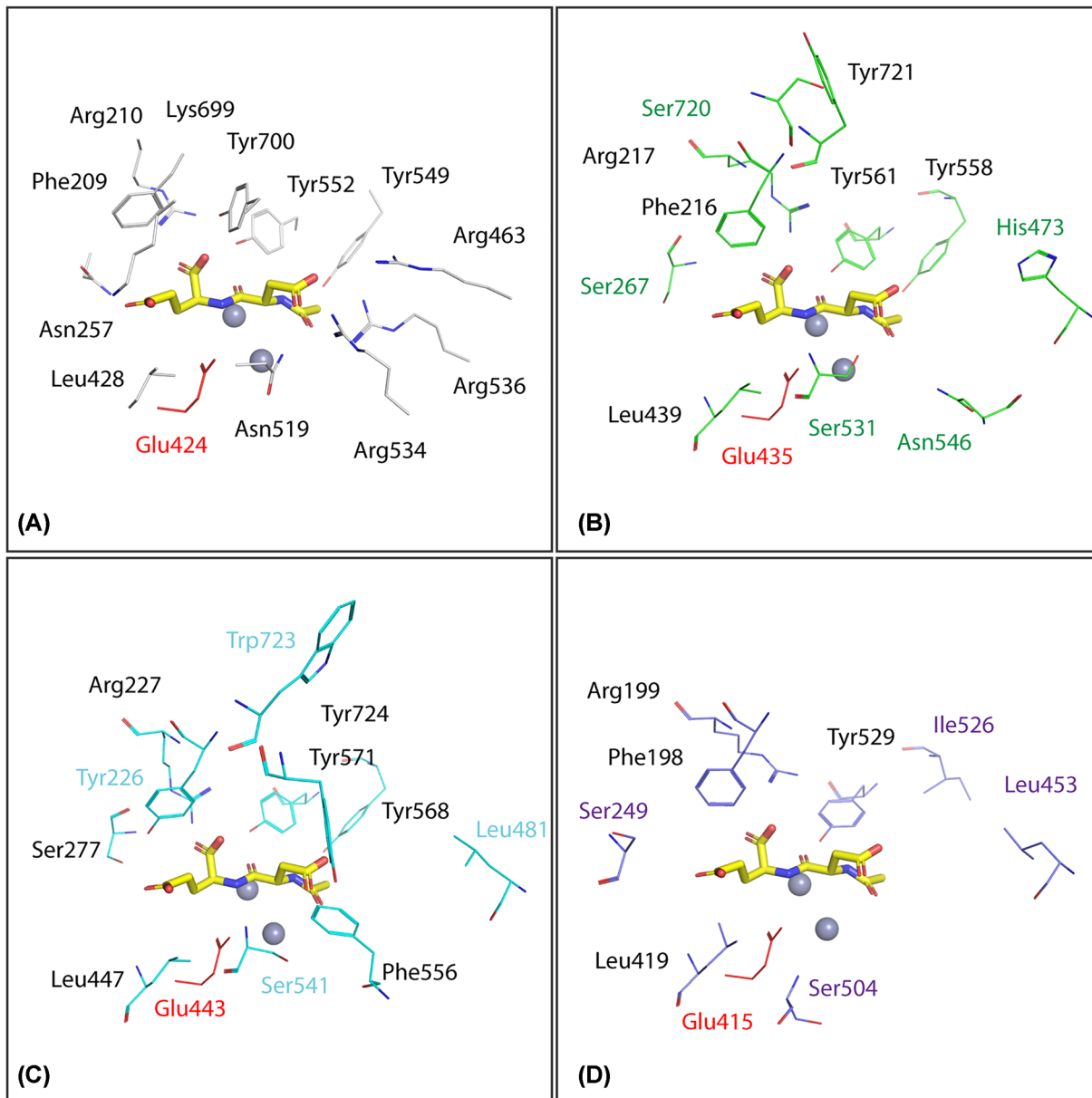
Based on a database search (WormBase, genome assembly WS284), all three genes are located on the X chromosome. *Gcp-2.1* gene is located in the X: 304131..305653 chromosomal region. And the *gcp-2.2* and *gcp-2.3*, are localized adjacent to each other in the X:11537714..11541538 and X:11542138..11545061 regions, respectively [28]. The corresponding proteins are termed CeGCP2.1, 2.2, and 2.3, and this respective terminology will be used throughout this report. All orthologs share the identical domain organization shown in Figure 1, which corresponds to their human GCP2 ortholog (hsGCP2) (also later verified by homology modeling, Figure 2) consisting of transmembrane, protease, apical, and C-terminal dimerization domains [47]. Interestingly, a deletion of 45 amino acids was observed for CeGCP2.3 in C-terminal dimerization domains that could negatively influence the quaternary homodimeric structure of this isoform (Figure 1). Furthermore, three splice variants exist for CeGCP2.1, where the CeGCP2.1b (751 amino acids) isoform lacks the sequence encoding the intracellular part (as compared to the CeGCP2.1a isoform; 770 amino acids), and the shortest CeGCP2.1c isoform (576 amino acids) additionally lacks the C-terminal dimerization domain (Supplementary Figure S3). The results of the protein structure analysis showing sequence identity and similarity of hsGCP2 and CeGCP2 orthologs are presented in Table 1.

To provide structural insight into CeGCP2 variants, we prepared homology models of all three *C. elegans* enzymes and compared them with the X-ray crystal structure of the hsGCP2 (PDB code: 3BXM [48]; Figure 2). In line with the sequence alignment, our homology models reveal the conservation of residues coordinating Zn^{2+} ions in the bimetallic active site of all proteins. Furthermore, Glu424, the proton shuttle residue critical for the catalytic activity of hsGCP2 [49], is also conserved in all orthologs. As both zinc ions and the proton shuttle are essential for GCP2 folding and hydrolytic activity, it is fair to speculate that *C. elegans* orthologs can be enzymatically competent.

Contrary to absolute structural conservation of the zinc-dependent active site, marked differences are observed for residues delineating substrate specificity pockets in the internal cavity of these enzymes, pointing toward their

Figure 1. Primary sequence alignment of CeGCP2.1, CeGCP2.2, CeGCP2.3, and hsGCP2

The intracellular and transmembrane parts of GCP2 are shaded turquoise and magenta, respectively. Protease-like, apical, and C-terminal dimerization domains are shaded red, yellow, and green, respectively. The catalytic acid/base glutamate is marked by red framing with red shading and residues coordinating the active-site zinc ions are marked with a purple frame. Residues delineating specificity pockets are indicated by a gray box with gray shading. The alignment and subsequent identity and similarity analyses are shown in Table. 1.

Figure 2. Comparison of active sites of hsGCP2 and CeGCP2 orthologs

(**A**) Residues delineating the substrate-binding pocket of hsGCP2, are shown in line representation, with atoms colored red, blue, and gray for oxygen, nitrogen, and carbon, respectively. Active-site bound NAAG is shown in stick representation with carbon atoms colored yellow. Zinc ions are shown as gray spheres. Residues in CEGCPs that are structurally equivalent to hsGCP2 ortholog are shown in line representation. (**B–D**) Conserved residues have black labels (B–D), while non-conservative substitutions are colored green, cyan, and violet for CeGCP2.1 (B), CeGCP2.2 (C), and CeGCP2.3 (D), respectively. The critical proton shuttle glutamate residue is highlighted in red. The figure was generated using PyMol 2.4.1.

distinct substrate specificities. For example, hsGCP2 has a preference for substrates that are negatively charged at both P1 and P1' positions [50]. At the P1' position, specificity for glutamate residues is achieved by the intricate network of hydrogen bonds and most notably the presence of positively charged Lys699 [6,49,51,52]. In *C. elegans* orthologs, the Lys699 is substituted by Ser720, Trp723, and Trp644 in CeGCP2.1, 2.2, and 2.3, respectively. Similarly, the arginine patch comprised of side chains of Arg463, Arg534, and Arg536 is critical for the recognition of negatively charged residues in the non-prime specificity pocket of hsGCP2 [47], yet these residues are missing in studied orthologs and/or are replaced by polar/hydrophobic amino acids (Figures 1 and 2). Taken together, GCP2 orthologs are likely
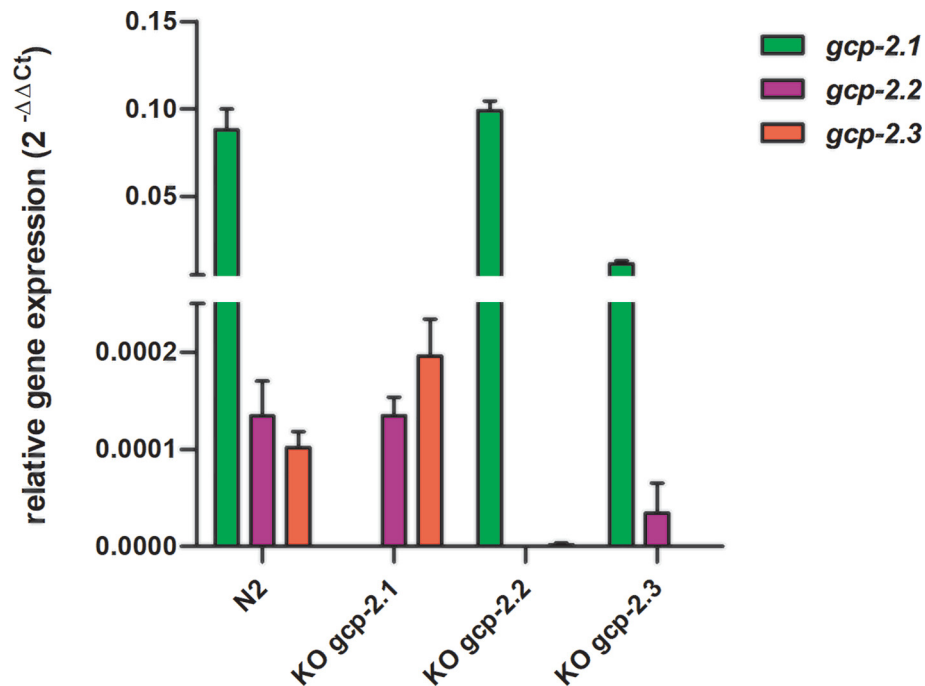
Figure 3. Gene expression of M28B metalloproteases in adult *C. elegans* evaluated by RT-Qpcr
Quantification of *gcp-2.1* (green), *gcp-2.2* (purple), and *gcp-2.3* (red) gene expression levels in wild-type N2 strain and mutants was performed using RT-qPCR; cycle threshold (*C*t) values normalized to *tba-1* as a housekeeping gene, are shown as mean ± standard deviation.

proficient hydrolases with substrate specificities different from the human enzyme, and further studies are warranted to identify their physiological substrates in *C. elegans*.

Phylogenic diversification of M28B metallopeptidases within nematodes

To analyze phylogenetic relationship of M28B peptidases in nematodes, we created a maximum likelihood phylogenetic tree (Supplementary Figure S4). Here, CeGCP2.1 sits in a well-supported clade with other orthologous sequences of the genus *Caenorhabditis*. However, CeGCP2.2 and 2.3 homologs form a clade that branches at the base of group A. This position in the phylogenetic tree shows that CeGCP2.2 and 2.3 are phylogenetically relatively distant from GCP2.1. Nonetheless, they are closely related to each other and most likely originated from gene duplication. These findings suggest that CeGCP2.2 and 2.3 arose by later diversification from the CeGCP2.1. As seen from the gene databases [28], the *gcp-2.1* gene encodes an M28B metalloprotease ancestral for nematodes, while any other gene duplications or modifications result from an adaptive response to the environmental conditions.

Relationship in the gene expression levels of the three *C. elegans* M28B peptidases

To get insight into the physiological roles of M28B peptidases, we first analyzed their gene expression by means of RT-qPCR and then explored the effect of knocking out the *gcp-2.1* (R57.1; allele name – *ok1004*), *gcp-2.2* (C35C5.2; allele name – *tm6541*), or *gcp-2.3* (C35C5.11; allele name – *tm5414*) genes (Supplementary Figure S5) on each other's expression (Figure 3). All three genes including *gcp-2.3*, which is annotated as a pseudogene in WormBase, are expressed (Figure 3). *Gcp-2.1* is the most abundant transcript and *gcp-2.3* has the lowest level of expression in wild-type worms. Based on RT-qPCR data, the expression of the *gcp-2.2* gene was not affected in *gcp-2.1* mutant animals, while the expression of *gcp-2.3* was increased approximately 2-fold in comparison with the wild-type population. The knockout of *gcp-2.2* completely suppressed the expression of the *gcp-2.3* gene to only 1.8% of the wild-type levels. *Gcp-2.3* knockout has significantly reduced the expression of *gcp-2.1* and *gcp-2.2* genes 8- and 4-fold, respectively, when compared to the wild-type control. For those reasons, the impacts of the particular knockouts on the expression of other orthologs must be considered while investigating specific effects on the worm phenotypes, as shown below.
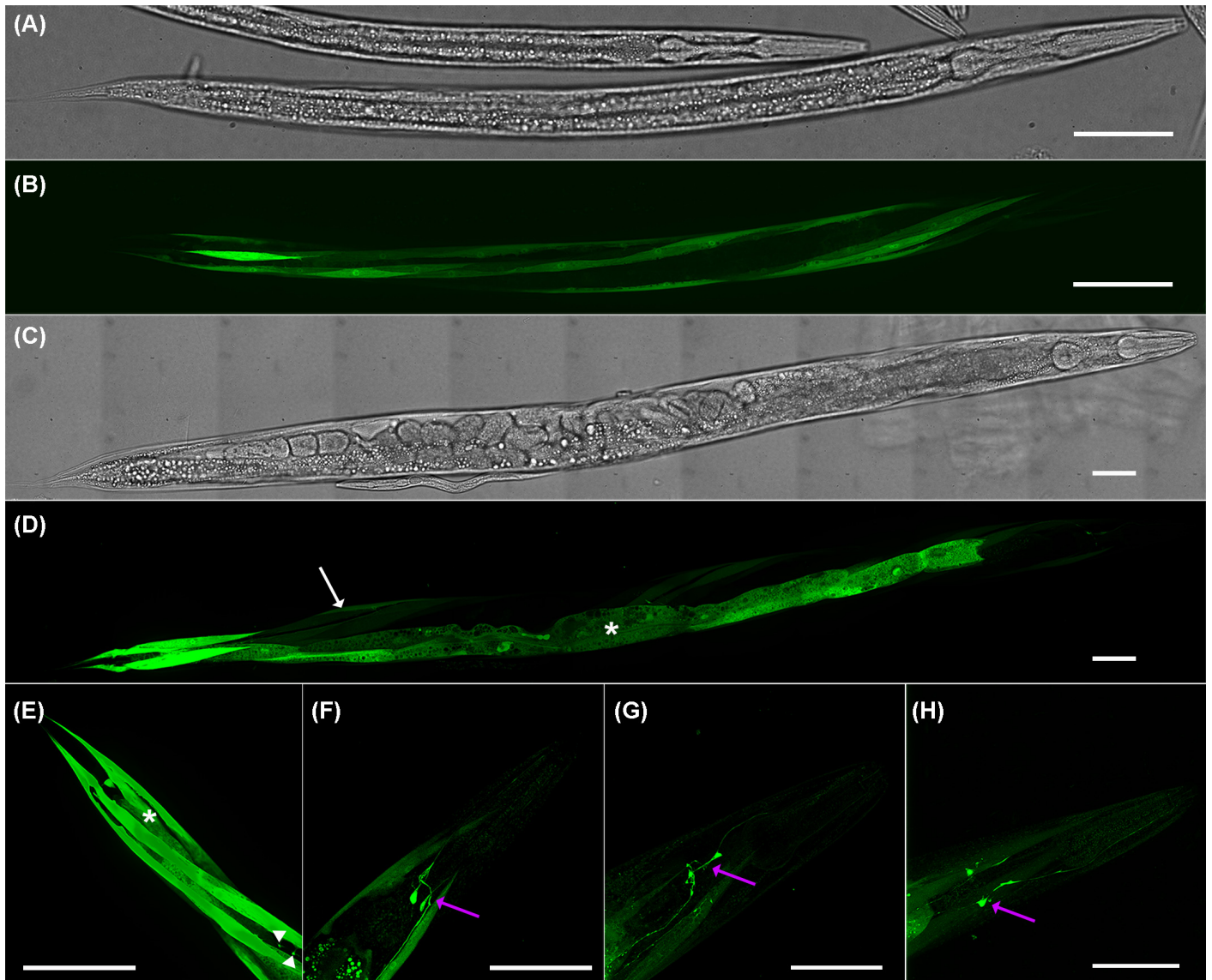
Figure 4. Expression pattern of *gcp-2.1* in the wild-type *C. elegans* visualized by GFP
(**A**) Bright-field image of L3 larvae. (**B**) *gcp2.1p::GFP* expression was localized to somatic muscles of L3 larvae. (**C**) Bright field of the adult worm. (**D**) Strong expression of *gcp-2.1p::GFP* was observed as well in the adult intestine (asterisk), and somatic muscles (arrow). (**E**) Detailed localization of the *gcp-2.1* in the posterior part of the body of wild-type adult *C. elegans*. The expression is visible in the intestine (asterisk), muscles and slender muscle arms reaching the nervous cord (arrowheads). (**F–H**) Expression of *gcp-2.1p::GFP* in the class SAA pharyngeal interneurons and their anterior processes (purple arrows) of adult *C. elegans*. Scale bars represent 50 μm.

Tissue-specific expression of CeGCP2

To determine the expression sites of the three M28B paralogs *in vivo*, we used standard promoter-driven GFP (green fluorescent protein) expression. Upstream regions of each gene used to generate transgenes are described in Supplementary Figure S2. Our data revealed that expression patterns of the three proteases differ markedly, and their expression is limited to distinct tissue structures.

Consistent with the high expression of the *gcp-2.1* gene (Figure 3A), high level of extrachromosomal transgene *gcp-2.1p::GFP* expression was detected in several tissues of the wild-type worms, including the intestine, somatic muscles (Figure 4D), and the nervous system of adult worms (Figure 4E–H). Specifically for the nervous tissue, the *gcp-2.1p::GFP* is mainly expressed in class SAA pharyngeal interneurons (a paired motor neuron localized in the

head that uses acetylcholine signaling) and their anterior processes (Figure 4F–H). L3 larvae, in contrast, displayed expression of *gcp-2.1p::GFP* predominantly in the muscles (Figure 4B).

Lower levels of *gcp-2.2* and *gcp-2.3* gene expression (Figure 3A) correspond to the restricted GFP expression pattern, where each paralog is observed only in a single site: phasmid neurons (PHAL/R, PHBL/R) and their processes (Figure 5B,C) in adults and L3 larvae (Supplementary Figure S6b) for *gcp-2.2p::GFP* and the excretory cell and its extended bilateral canals reaching from the nose to the tail region of the adult (Figure 6B) and larvae (Supplementary Figure S6d) for *gcp-2.3p::GFP* (for an overview see Figure 12).

PHA and PHB, glutamatergic ciliated sensory neurons, are localized in the lumbar ganglia, exposed to the outside environment, and serve as chemosensory cells. Dye-filling with DiI was used for PHAL/R and PHBL/R neuron identification (Supplementary Figure S7b–d) and confirmation of phasmid's functionality in *gcp-2.2* mutants (Supplementary Figure S7f). DiI is a lipophilic fluorescent dye able to selectively incorporate into two pairs of ciliated phasmid neurons in the tail (PHA and PHB) via cilia present at the tips of dendrites [53,54].

Knockout of peptidase paralogs is manifested in different phenotypes

Knockouts of the studied genes resulted in the most apparent phenotypes reported in any experimental animal to date.

The progeny production was changed after the *gcp-2.1* knockout. The number of progeny in the *gcp-2.1* mutant strain was reduced by 25% (Figure 7A and Supplementary Table S4) compared with the wild-type strain but the difference was not statistically significant. Furthermore, manual counting revealed a 50% increase in pharyngeal pumping frequencies in the *gcp-2.1* mutant strain (Figure 7B and Supplementary Table S5). At the same time, *gcp-2.1* knockout did not visibly affect the behavior and movement of the worms. Exposing young adult *gcp-2.1* mutants to higher NaCl concentrations induced acute osmotic stress, leading to a reduction in their locomotor activity. These changes in movement were not different from those observed in the other strains examined in the study (Figure 8). However, when *gcp-2.1* mutants were exposed to elevated NaCl concentration (300 mM) over several days, their survival time significantly decreased in contrast to *gcp-2.3* mutants (Figure 9A and Supplementary Table S6a). Additionally, at a NaCl concentration of 400 mM, *gcp-2.1* mutants displayed a shorter survival time not only in comparison with *gcp-2.3* mutants but also when compared with N2 worms (Figure 9B and Supplementary Table S6b).

Knockout of the *gcp-2.2* gene significantly negatively impacted progeny production rate (60% decrease compared with N2; Figure 7A and Supplementary Table S4) and caused one day shorter progeny production. The increase in the pharyngeal pumping rate was not statistically significant (Figure 7B and Supplementary Table S5). Interestingly, while handling the *gcp-2.2* knockout worms, we noticed they were more fragile compared with other mutants and wild-type worms. Thus, we performed a detailed ultrastructural analysis of the cuticle of all mutant strains using transmission electron microscopy (TEM). Figure 10 shows spongy formations in the basal zone of the cuticle, which were observed only in *gcp-2.2* mutant worms (Figure 10C and Supplementary Figure S8). To corroborate this finding, we conducted TEM on *gcp-2.2 gcp-2.3* double mutant worms. During the microscopic examination of the cuticle, we observed identical ultrastructural alterations in the basal layer of the cuticle, consistent with what was observed in *gcp-2.2* mutants (Supplementary Figure S9). Additionally, when we performed a cuticle permeability assay, it became evident that the *gcp-2.2* mutants exhibited the highest percentage (47%) of worms with Hoechst 33258-stained nuclei in the pharynx when compared to other mutant strains and N2 worms (Figure 11).

For *gcp-2.3* knockout worms, the number of offspring was reduced to approximately 75% of wild-type N2 controls, but the reduction was not statistically significant. (Figure 7A and Supplementary Table S4). On the other hand, the pharyngeal pumping rate remained virtually unchanged and compared with that of N2 worms (Figure 7B and Supplementary Table S5). The significant effect of the gene knockout on the lifespan of the mutants was not demonstrated (Supplementary Figure S10). Additionally, there were no significant differences observed in both acute and chronic responses to osmotic stress when compared with N2 worms (Figures 8 and 9; Supplementary Table S6; for an overview see Table 2).

Discussion

Research into metalloproteases belonging to the M28B subfamily is primarily centered around human GCP2. However, there are still gaps in our understanding of the physiological functions of this protease and its related orthologs [6,22]. By use of promoter-driven GFP expression together with gene knockout mutants, we were able to reveal some of the functions of these enzymes in *C. elegans* including unique phenotypes linked to the particular M28B metalloproteases in *C. elegans* [23,55].
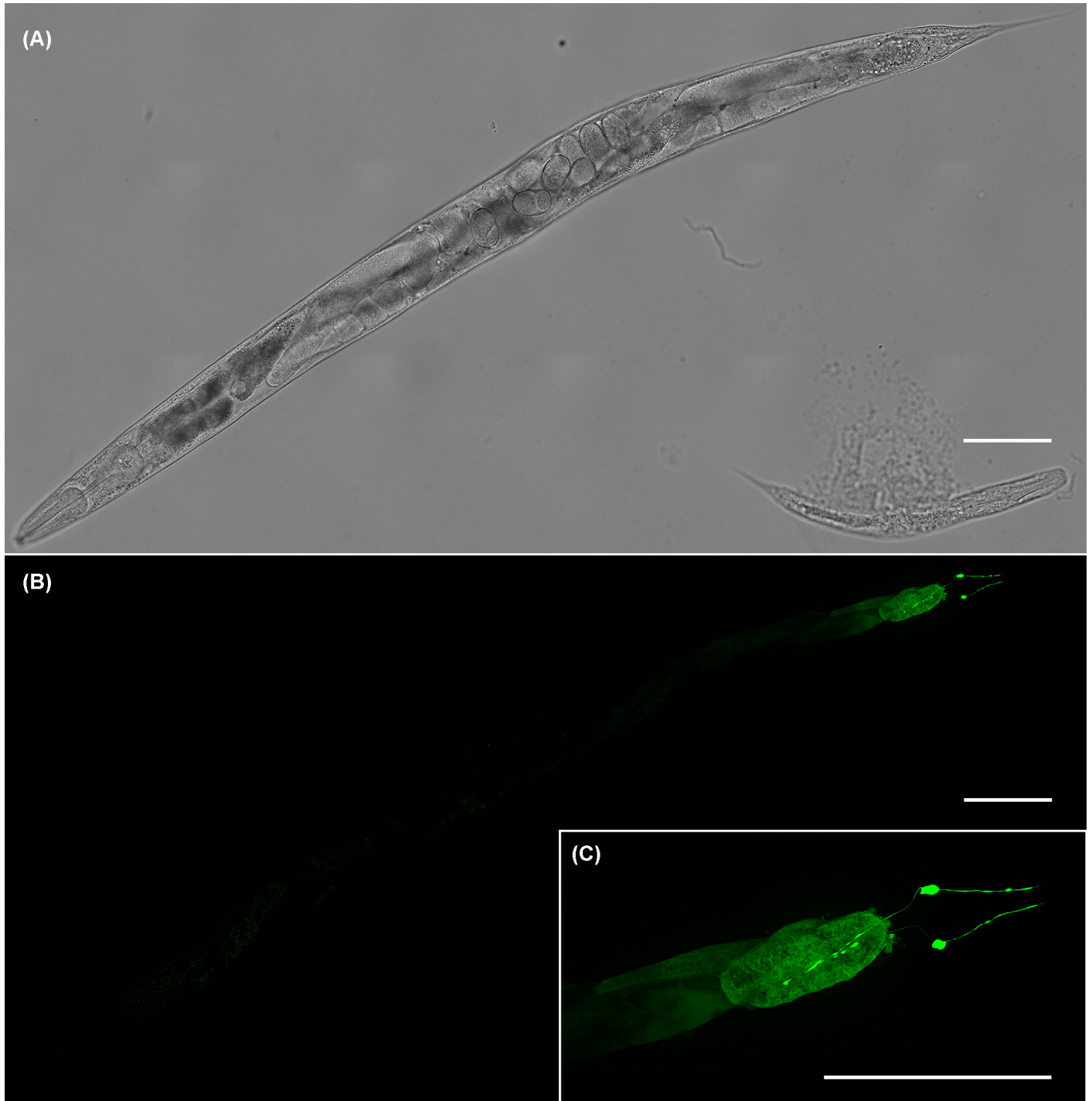
Figure 5. Localization of *gcp-2.2p::GFP* expression in the phasmids neurons (PHAL/R and PHBL/R) and their processes
(**A**) Bright-field image of the adult worm. (**B**) The GFP signal was exclusively observed in phasmids neurons (PHAL/R) and their processes in the tail of adult *C. elegans*. (**C**) GFP signal in PHAL/R and PHBL/R phasmids in detail. Scale bars represent 100 μm.

A phylogenetic comparison of orthologs from the genus *Caenorhabditis* showed that CeGCP2.1 is phylogenetically relatively distant from CeGCP2.2 and CeGCP2.3, which are closely related to each other. The ancestor of CeGCP2.2 and CeGCP2.3 originated in gene duplication derived from CeGCP2.1, probably in the last common ancestor of the group *Caenorhabditidae*. Another consecutive gene duplication led to the speciation of CeGCP2.2 and
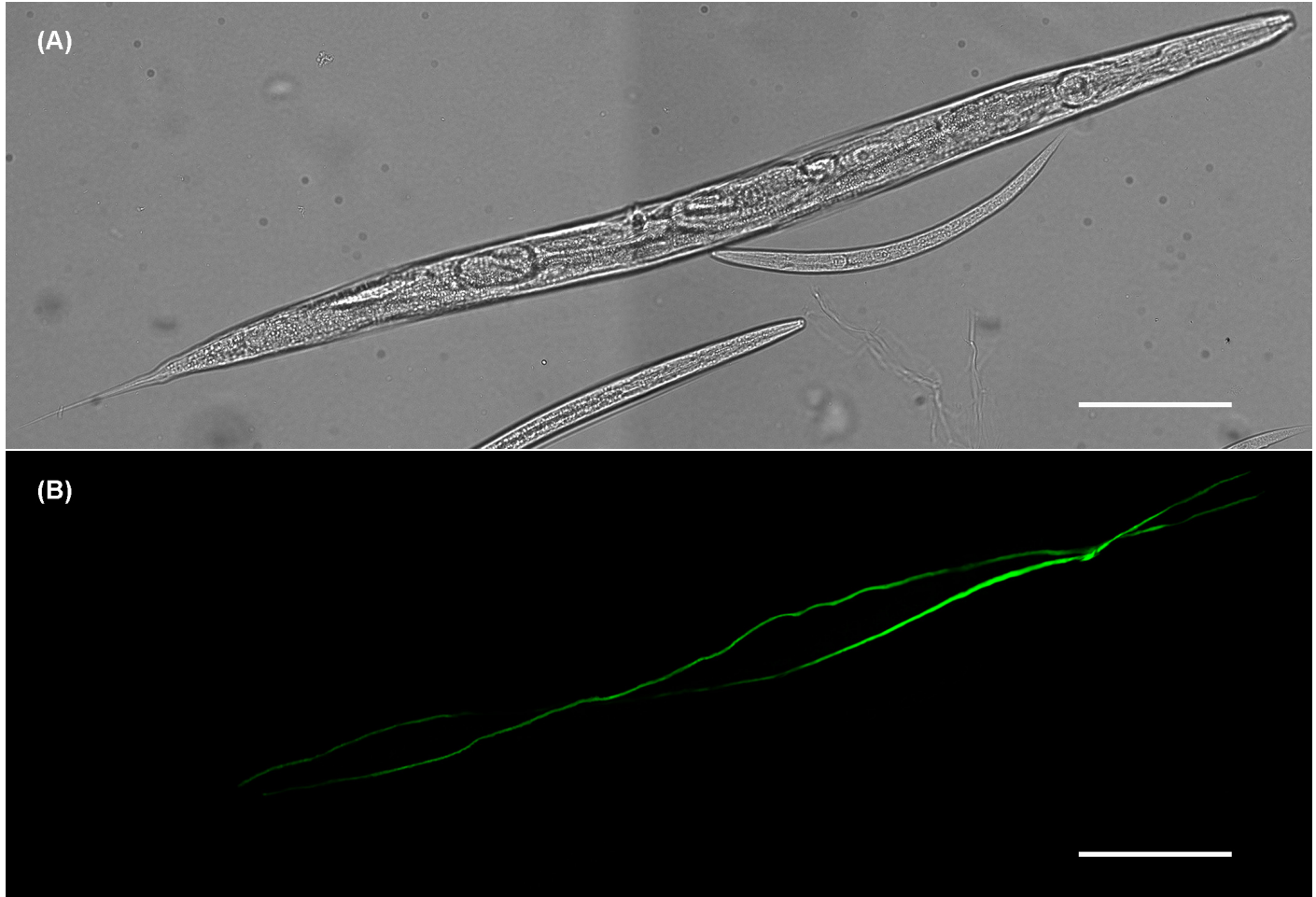
Figure 6. Expression of *gcp-2.3p::GFP* in the excretory cell of adult *C. elegans*
(**A**) Bright-field image of the adult worm. (**B**) GFP expression pattern revealed the exclusive localization of *gcp2.3* in the excretory system (H-shape cell and its processes along the body of the worm). Scale bars represent 100 μm.

Table 2 Summary table of the most prominent phenotypes

Phenotype	GCP2.1	GCP2.2	GCP2.3
Pharyngeal pumping	**The highest increase in the number of pharyngeal pumping/s**	Increased rate of pharyngeal pumping	Lowest pharyngeal pumping rate
Brood size	Reduced by a quarter	**The lowest progeny production**	Reduced by a quarter
Resistance to hyperosmotic stress	Significantly decreased survival time on the plates with high NaCl concentration		
Other observations		**Pronounced fragility, spongy structure in the cuticle**	

Statistically significant changes in phenotype manifestation are marked in bold.

CeGCP2.3, and there is strong evidence that these genes arose early in the evolution of the entire genus *Caenorhabditis*. They are likely more specialized for certain, yet unknown, physiological functions judging from their expression levels and localization. Such duplication events occurred in nematodes several times independently, and still, many species including important parasites possess only a single gene [25], making those enzymes putative drug targets.

Homology modeling showed that residues coordinating Zn^{2+} ions and the proton shuttle Glu424 essential for catalytic activity are conserved in all *C. elegans* GCP2 orthologs. Human GCP2 exhibits a preference for negatively charged substrates containing glutamate [50], while some changes in structural elements of *C. elegans* orthologs indicate possible differences in substrate preference. The residues (Asp387, His377, Asp453, Glu425, His553) that bridge
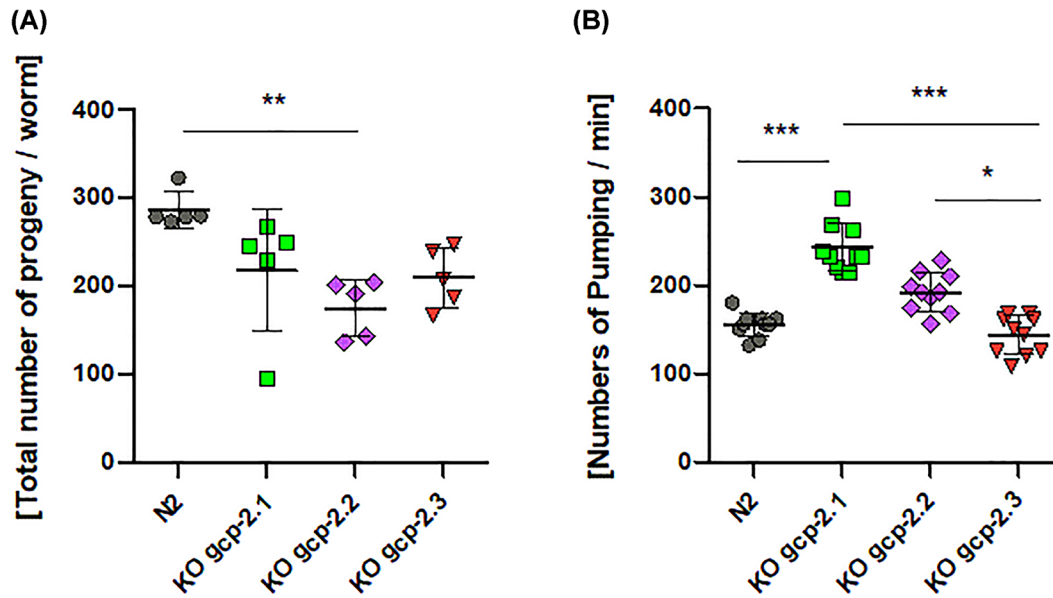
Figure 9. Effect of high NaCl concentrations on the duration of life span of *C. elegans* mutants
Young adult worms were seeded on NGM plates containing various concentrations of NaCl (**A**) 300 mM and (**B**) 400 mM. The count of viable worms was conducted at specific days post-beginning the assay. Worms were classified dead if they did not respond to repeated stimulation with a platinum wire. The data were tested by the Kruskal–Wallis test (nonparametric ANOVA, $P<0.0001$) and Dunn's Multiple Comparison Test (Supplementary Table S6a,b). P-value < 0.001 – Extremely significant (***); P-value 0.001 to 0.01 – Very significant (**); P-value 0.01 to 0.05 – Significant (*).

significant changes: replacement of a positively charged amino acid by a polar uncharged serine, and by an amino acid with a hydrophobic side, in CeGCP2.1 and CeGCP2.2 together with CeGCP2.3, respectively. In the S1 pocket of hsGCP2, Tyr549, and Tyr552 are important hydrophobic residues. Tyr 549 is only substituted for isoleucine in GCP2.3, so even after this substitution the hydrophobicity and thus the structural function is preserved.

The protein alignment of the primary structures of all three studied proteases and hsGCP2 reveals, that CeGCP2.3 lacks the transmembrane domain and has a partial deletion of the C-terminal dimerization domain. This suggests that the CeGCP2.3 protein might be intracellular or potentially even secreted [56–58]. The results of the protein structure analysis showed the closest similarity between CeGCP2.2 and CeGCP2.3, while CeGCP2.2 and CeGCP2.1 are less similar. Also, CeGCP2.1 is more similar to its human ortholog hsGCP2 than to the *C. elegans* paralog CeGCP2.3 (Table 1).

Although *gcp-2.3* is annotated as a pseudogene in the current version of WormBase (WS284) [28], our RT-qPCR data and transcriptomic data (unpublished) together with previously published time-resolved [59] and tissue-specific transcriptomes [60] of *C. elegans* showed that *gcp-2.3* is indeed transcribed. Research of gene expression dynamics throughout ontogenesis has revealed that gcp-2.3 is predominantly expressed during embryonic stages, reaching its highest expression peak at 0.6 DCPM (depth of coverage per million reads). In dauer larvae, it achieves a significant expression level of around 0.45 DCPM. However, in other stages, its expression is considerably reduced, falling below 0.2 DCPM [59].

Transcriptomic data for L2 tissues [59] showed that *gcp-2.2* is expressed in pharyngeal muscles, ciliated sensory neurons, excretory cells, non-seam hypodermis, body wall muscles, and intestine, while expression of *gcp-2.3* was reported in socket cells and intestines. In addition, tissue-specific transcriptomic analysis of adults [60] showed that *gcp-2.2* is mainly expressed in pharyngeal gland cells, RID (single motor neuron, interneuron situated in the dorsal ganglion, innervates dorsal body muscles of *C. elegans*), inner labial neurons, head muscles, GABAergic neurons, head neurons, rectal gland cells, and head ganglion. For *gcp-2.3*, the expression is limited to amphid sheath cells, RID, and phasmid neurons. Contrary to these findings, our promoter-driven GFP reporters revealed different localizations. *Gcp-2.2p::GFP* expression was restricted to phasmids (PHA and PHB) in both larvae and adult worms and *gcp-2.3p::GFP* was exclusively expressed in the excretory cell and its processes of larvae and adult wild-type worms. Discrepancies between the results could be because the above-mentioned transcriptomic datasets served primarily as expression prediction, and they do not provide information about the spatial constraints of expression. Moreover, our design of the GFP-fusion construct for *gcp-2.2* and *gcp-2.3* may not cover all regulatory elements in the different tissues and cells.
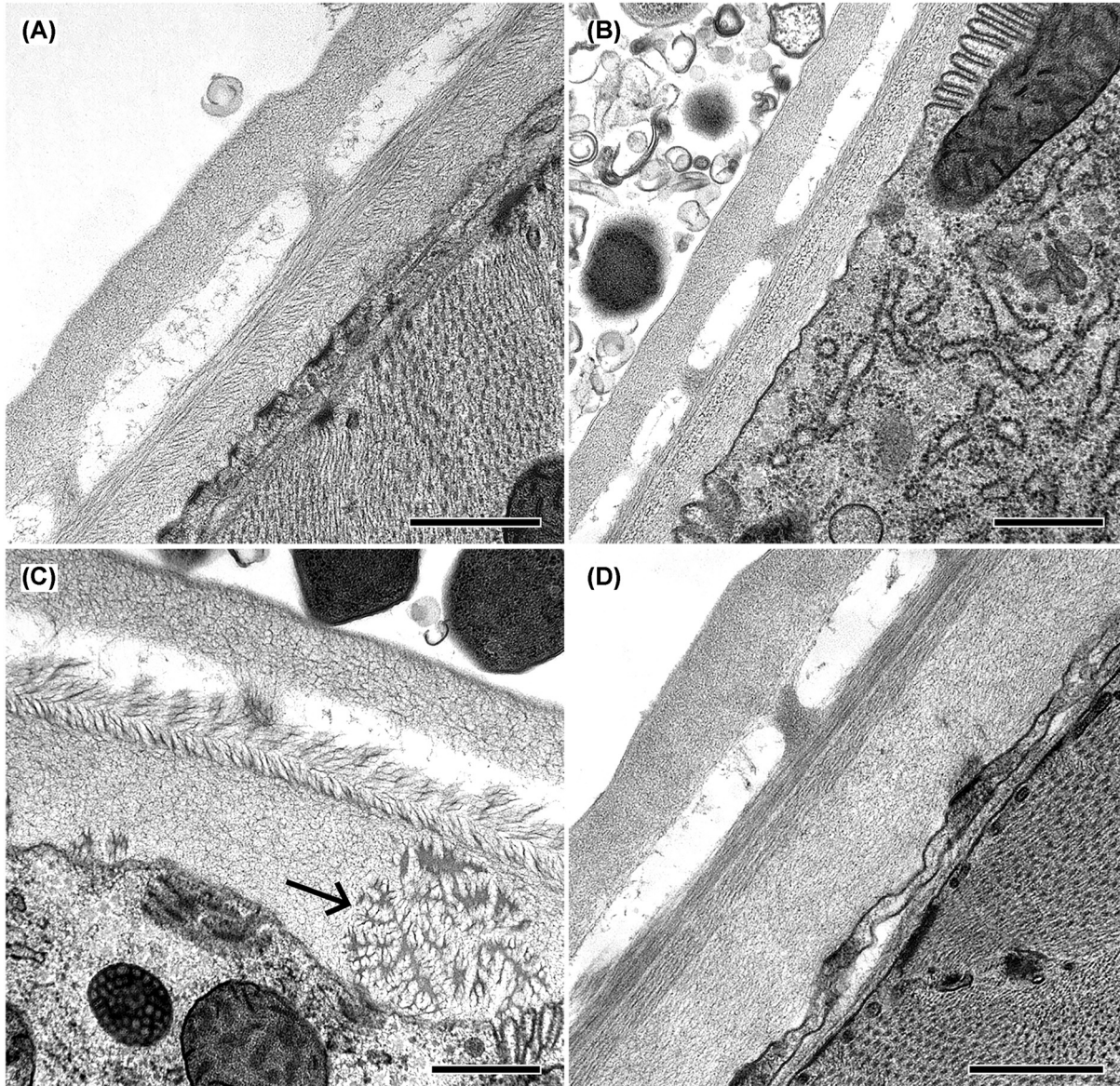
Figure 10. Ultrastructural analysis of the surface structures of *C. elegans*
(**A**) Image of the cuticle of N2 strain. (**B**) Cuticle of *gcp-2.1* mutant adult worm. (**C**) Cuticle of *gcp-2.2* mutant adult worm. For *gcp-2.2* mutants (**C**) unusual spongy structures (arrow) in the basal layer were observed. These structures were not detected in the cuticle of other strains of *C. elegans* included in the present study. (**D**) Cuticle of *gcp-2.3* mutant adult worm. Scale bars indicate 500 nm.

Based on our localization studies, *gcp-2.1* is expressed in pharyngeal interneurons class SAA, somatic muscles, and intestine. SAA class neurons belong to neurons receiving cholinergic inputs [61]. The long, anteriorly directed processes of SAA interneurons probably serve as stretch receptors monitoring the posture of the tip of the head [62]. SAA interneurons interact with the body and the head motor system, and it seems that they can coordinate head and body movement [62]. Indeed, localization of *gcp-2.1* expression in muscles, nervous system, and SAA interneurons could be related to some of our phenotyping results. We found out that *gcp-2.1* mutant worms showed the most significant increase in the number of pharyngeal pumps per minute of all studied paralogs. Even if the pumping rate is modulated by multiple mechanisms [63], the elimination of the *gcp-2.1* gene has an impact on the speed of pumping. It could suggest that CeGCP2.1 plays an important role in pharyngeal pumping modulation. Although the somatic muscles, nervous system, and intestine were dominant locations of the *gcp-2.1* expression, we did not detect any other phenotypic manifestations related to the digestive or locomotor system in *gcp-2.1* mutant worms. Despite
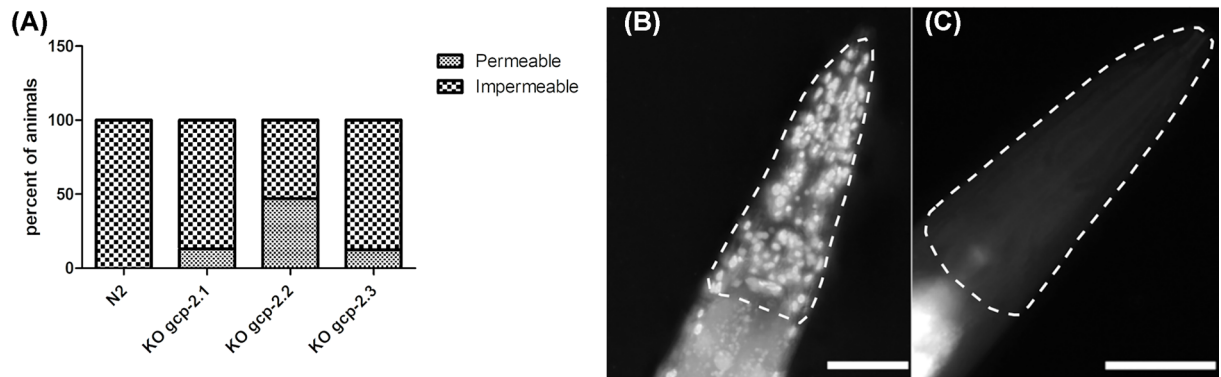
**Figure 11. Hoechst 33258 staining-based permeability assay**
Young adult worms from N2 and the mutant strains *gcp-2.1*, *gcp-2.2*, and *gcp-2.3* were subjected to Hoechst 33258 staining.
Worms displaying a minimum of 15 stained nuclei within the pharynx were designated as permeable. Each independent experiment
involved scoring 17-23 animals from each strain. (**A**) Graph illustrating the percentage distribution of worms with permeable and
impermeable surfaces. (**B**) *Gcp-2.2* mutant worms with Hoechst 33258-stained nuclei in the pharynx (indicated with a broken line).
(**C**) N2 worms impermeable for Hoechst staining (indicated with a broken line). Scale bars represent 50 μm.

the abnormal pharyngeal pumping rate, which could have an impact on the fitness of the worms, knockout of the
gcp-2.1 had only a mild effect on fertility. An alternative consequence of gene knockout might involve alterations in
resistance to hyperosmotic stress. We hypothesized that *gcp-2.2* and *gcp-2.3* mutants would be the most affected,
considering their fragile phenotype and expression localized in the excretory cell, respectively. Surprisingly, it was the
gcp-2.1 mutants that displayed the greatest susceptibility. We speculate that the elimination of *gcp-2.1*, which affects
the nervous system, gut, and muscle, resulted in reduced overall worm fitness. Consequently, any additional stress,
such as hyperosmotic stress in this case, led to a decreased survival time.

Knockout of the *gcp-2.2* gene caused increased fragility of *C. elegans* cuticle. This unique fragile phenotype was ac-
companied by previously undescribed spongy-structure formations in the basal zone of the cuticle which we presume
to be responsible for the observed fragility of worms during the routine manipulation. Changes in cuticle formation
may influence the permeability of the worm's surface to Hoechst staining, a phenomenon observed in 47% of *gcp-2.2*
mutant worms. It is established that deficiencies in collagen or processing enzyme expression can lead to cuticle per-
meability defects [64,65]. We can hypothesize whether GCP2.2 is involved in the correct formation and alignment
of collagens within the basal layer of the cuticle. However, this hypothesis warrants further comprehensive investiga-
tion. We also observed the lowest offspring production of all compared strains. Theoretically, cuticle fragility could
affect offspring production by weakening the vulval cuticle, but we did not detect bursting through the vulva or other
indicators of defective egg laying; therefore, these two phenotypes (lower fecundity and brittle cuticle) are likely un-
related. However, the localization of *gcp-2.2* expression pointing to phasmid neurons (PHA, PHB) does not explain
the described fragility and reduction in offspring production. The reason may be that we were not able to detect all
regulatory elements for epidermal expression and thus did not localize the gene expression in the cuticle.

The expression of *gcp-2.3* was localized to the excretory cell, which maintains osmotic equilibrium in the worm
body [66], and its extended bilateral canals reached from the nose of the worm to the tail region. Many mutations that
affect the excretory system cause a lethal phenotype or mutants with milder abnormalities of these cells often appear
pale, and slightly bloated [66]. Nevertheless, such phenotypes were not noticed in the case of *gcp-2.3* mutant worms
during our study. Knockout of the *gcp-2.3* led to a decrease in fertility similarly as observed in the case of the *gcp-2.1*
mutant worms. We are not able to state that the lower fertility is directly affected by the missing expression of *gcp-2.3*
in the excretory cell. Even though we haven't detected any adverse impacts on resistance to hyperosmotic stress, the
deletion of *gcp-2.3* may have had a detrimental effect on the animals' fitness, potentially resulting in a decline in the
number of offspring.

Concerning the phenotype manifestation of *C. elegans gcp-2* gene knockouts and their particular localizations,
there are noticeable significant similarities with mammalian protease orthologs. Among others, these proteases have
been very well-described and studied in the digestive system [11,67,68] and also in the brain, astrocytes, and Schwann
cells [9]. Similarly, our observations revealed the expression of *C. elegans gcp-2.1* in the intestine and nervous tissue
of adult worms. However, our data from *in silico* modeling and a recent experiment with the recombinant protease
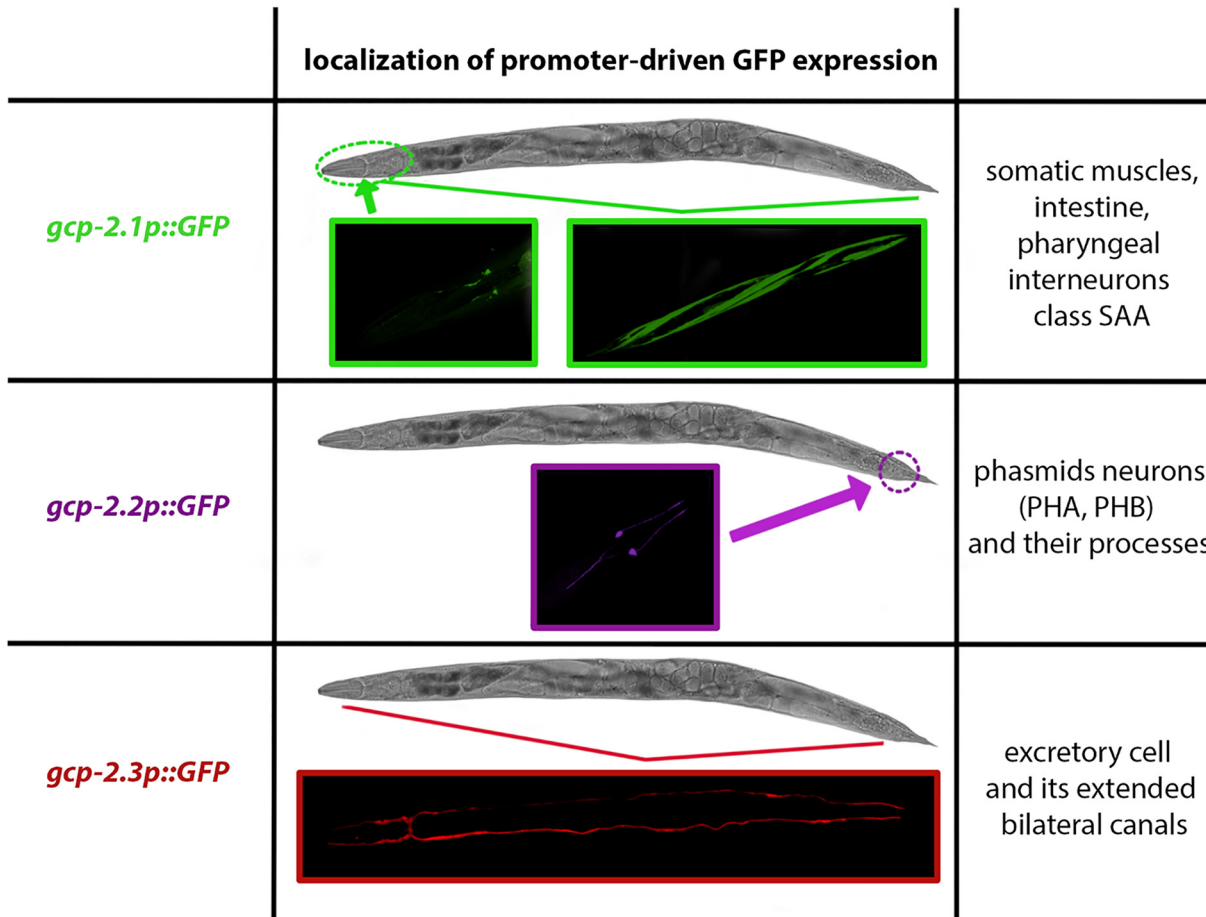
Figure 12. Overview of the promoter-driven GFP expression of three GCP2 paralogs in distinct tissues of *C. elegans*

(not published) do not support the possibility that CeGCP2.1 would copy the physiological function of human GCP2 by removing terminal glutamate.

Using the transcript information from WormBase [27], we analyzed two additional CeGCP2.1 splice variants. A comparison of CeGCP2.1 isoforms revealed that isoform CeGCP2.1b lacks the predicted intracellular domain. Isoform CeGCP2.1c completely lacks the C-terminal domain, which implies the loss or significant changes in the catalytic function, given that in its human homolog GCP2, the C-terminal domain is involved in substrate recognition [69,70]. No biological functions associated with these isoforms of GCP2.1 protein have been described so far, so we can only conclude homology to other isoforms and GCP2-like proteins of *Caenorhabditis* species, NAAG and transferrin receptors.

We have demonstrated that the *gcp-2.2* mutant line almost completely suppressed the expression of the *gcp-2.3* gene; the decreased expression could be explained by the aforementioned proximity of both genes on the chromosome, suggesting the presence of an alternative polycistronic expression of these proteins [18]. Additionally, we have found that knockout of the *gcp-2.1* gene significantly increased the expression of *gcp-2.3*, which is surprising given the differences in phenotypic expression and localization. The question is whether the increased expression of the *gcp-2.3* gene can somehow compensate for the missing *gcp-2.1*. In addition, the knockout of the *gcp-2.3* gene resulted in a decrease in *gcp-2.1* and *gcp-2.2* expression. A possible explanation for this decrease in *gcp-2.1* expression could be impaired function of the excretory system and thus a decrease in the fitness after the *gcp-2.3* gene knockout. However, we do not yet have a satisfactory explanation for these changes in expression.

In summary, we utilized *C. elegans* as a model organism to elucidate the diversification of a single gene. Following a similar approach, we can postulate an indirect comparison regarding the diversification of five independently evolved genes in humans. Through localization and phenotyping studies, we have demonstrated that *gcp-2.3* is not a pseudogene, as erroneously annotated in the WormBase database [27], but instead, an active paralogous gene. Precise

localization via promoter-driven GFP expression unveiled distinct expression patterns for each of the three *C. elegans* genes. Remarkably, the localization of these *C. elegans* proteases belonging to the M28B family closely mirrors their orthologous counterparts in mammals. The knockout of these studied genes resulted in the most pronounced phenotypic consequences ever documented in the animal kingdom. Beyond its impact on the pharyngeal pumping and reproduction of mutant worms, we observed a unique fragility phenotype in the worms, likely originating from abnormalities such as spongy formations in the worm's cuticle.

Data Availability

The data from the present study are available from the Corresponding Author upon request. No mandated data types are included in this manuscript.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

CRediT Author Contribution

Lucie Panska: Data curation, Formal analysis, Investigation, Writing—original draft. **Stepanka Nedvedova:** Data curation, Formal analysis, Investigation, Writing—original draft. **Vojtech Vacek:** Data curation, Formal analysis, Writing—review & editing. **Daniela Krivska:** Formal analysis, Investigation. **Lukas Konecny:** Formal analysis, Writing—review & editing. **Filip Knop:** Formal analysis, Investigation. **Zsofia Kutil:** Formal analysis, Writing—review & editing. **Lubica Skultetyova:** Formal analysis. **Adrian Leontovyc:** Formal analysis. **Lenka Ulrychova:** Formal analysis. **Judy Sakanari:** Methodology, Writing—review & editing. **Masako Asahina:** Methodology, Writing—review & editing. **Cyril Barinka:** Conceptualization, Formal analysis, Investigation, Methodology, Writing—review & editing. **Marie Macurkova:** Conceptualization, Data curation, Formal analysis, Methodology, Writing—review & editing. **Jan Dvorak:** Conceptualization, Data curation, Formal analysis, Supervision, Investigation, Methodology, Writing—original draft, Project administration.

Acknowledgements

Abbreviations

FOLH1, folate hydrolase 1; GCP2, glutamate carboxypeptidase 2; GCP3, glutamate carboxypeptidase 3; HPF, high-pressure freezing; NAAG, N-acetyl-aspartyl-glutamate; NAALADase, N-acetylated-α-linked acidic dipeptidase; NGM, nematode growth medium; PSMA, prostate-specific membrane antigen; TEM, transmission electron microscopy.

References

1 Rawlings, N.D., Barrett, A.J., Thomas, P.D., Huang, X., Bateman, A. and Finn, R.D. (2018) The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* **46**, D624–D632, https://doi.org/10.1093/nar/gkx1134

2 Tykvart, J., Schimer, J., Jančařík, A., Bařinková, J., Navrátil, V., Starková, J. et al. (2015) Design of highly potent urea-based, exosite-binding inhibitors selective for glutamate carboxypeptidase II. *J. Med. Chem.* **58**, 4357–4363, https://doi.org/10.1021/acs.jmedchem.5b00278

3 Hlouchova, K., Navratil, V., Tykvart, J., Sacha, P. and Konvalinka, J. (2012) GCPII variants, paralogs and orthologs. *Curr. Med. Chem.* **19**, 1316–1322, https://doi.org/10.2174/092986712799462676

4 Lambert, L.A. and Mitchell, S.L. (2007) Molecular evolution of the transferrin receptor/glutamate carboxypeptidase II family. *J. Mol. Evol.* **64**, 113–128, https://doi.org/10.1007/s00239-006-0137-4

5 Tykvart, J., Bařinka, C., Svoboda, M., Navrátil, V., Souček, R., Hubálek, M. et al. (2015) Structural and biochemical characterization of a novel aminopeptidase from human intestine. *J. Biol. Chem.* **290**, 11321–11336, https://doi.org/10.1074/jbc.M114.628149

6 Pavlicek, J., Ptacek, J. and Barinka, C. (2012) Glutamate carboxypeptidase ii: an overview of structural studies and their importance for structure-based drug design and deciphering the reaction mechanism of the enzyme. *Curr. Med. Chem.* **19**, 1300–1309, https://doi.org/10.2174/092986712799462667

7 Ptacek, J., Nedvedova, J., Navratil, M., Havlinova, B., Konvalinka, J. and Barinka, C. (2018) The calcium-binding site of human glutamate carboxypeptidase II is critical for dimerization, thermal stability, and enzymatic activity. *Protein Sci.* **27**, 1575–1584, https://doi.org/10.1002/pro.3460

8 Matteucci, F., Mezzenga, E., Caroli, P., Di Iorio, V., Sarnelli, A., Celli, M. et al. (2017) Reduction of 68Ga-PSMA renal uptake with mannitol infusion: preliminary results. *Eur. J. Nucl. Med. Mol. Imaging* **44**, 2189–2194, https://doi.org/10.1007/s00259-017-3791-4

9 Rovenska, M., Hlouchova, K., Sacha, P., Mlcochova, P., Horak, V., Zamecnik, J. et al. (2008) Tissue expression and enzymologic characterization of human prostate specific membrane antigen and its rat and pig orthologs. *Prostate* **68**, 171–82

10 Silver, D.A., Pellicer, I., Fair, W.R., Heston, W.D. and Cordon-Cardo, C. (1997) Prostate-specific membrane antigen expression in normal and malignant human tissues. *Clin. Cancer Res.* **3**, 81–85

11 Visentin, M., Diop-Bove, N., Zhao, R. and Goldman, I.D. (2014) The intestinal absorption of folates. *Annu. Rev. Physiol.* **76**, 251–274, https://doi.org/10.1146/annurev-physiol-020911-153251

12 Rajasekaran, A.K., Anilkumar, G. and Christiansen, J.J. (2005) Is prostate-specific membrane antigen a multifunctional protein? *Am. J. Physiol.-Cell Physiol.* **288**, C975–C981, https://doi.org/10.1152/ajpcell.00506.2004

13 Tsai, G., Stauch-Slusher, B., Sim, L., Hedreen, J.C., Rothstein, J.D., Kuncl, R. et al. (1991) Reductions in acidic amino acids and N-acetylaspartylglutamate in amyotrophic lateral sclerosis CNS. *Brain Res.* **556**, 151–156, https://doi.org/10.1016/0006-8993(91)90560-I

14 Plaitakis, A. (1990) Glutamate dysfunction and selective motor neuron degeneration in amyotrophic lateral sclerosis: a hypothesis. *Ann. Neurol.* **28**, 3–8, https://doi.org/10.1002/ana.410280103

15 Fricker, A.C., Selina Mok, M.H., de la Flor, R., Shah, A.J., Woolley, M., Dawson, L.A. et al. (2009) Effects of N-acetylaspartylglutamate (NAAG) at group II mGluRs and NMDAR. *Neuropharmacology*, https://doi.org/10.1016/j.neuropharm.2009.03.002

16 Šácha, P., Zámečník, J., Bařinka, C., Hlouchová, K., Vícha, A., Mlčochová, P. et al. (2007) Expression of glutamate carboxypeptidase II in human brain. *Neuroscience* **144**, 1361–1372, https://doi.org/10.1016/j.neuroscience.2006.10.022

17 Lu, X.C.M., Tang, Z., Liu, W., Lin, Q. and Slusher, B.S. (2000) N-acetylaspartylglutamate protects against transient focal cerebral ischemia in rats. *Eur. J. Pharmacol.* **408**, 233–239, https://doi.org/10.1016/S0014-2999(00)00762-7

18 Evans, D. and Blumenthal, T. (2000) Trans splicing of polycistronic Caenorhabditis elegans pre-mRNAs: analysis of the SL2 RNA. *Mol. Cell. Biol.* **20**, 6659–6667, https://doi.org/10.1128/MCB.20.18.6659-6667.2000

19 Maurer, T., Eiber, M., Schwaiger, M. and Gschwend, J.E. (2016) Current use of PSMA-PET in prostate cancer management. *Nat. Rev. Urol.* **13**, 226–235, https://doi.org/10.1038/nrurol.2016.26

20 Haffner, M.C., Kronberger, I.E., Ross, J.S., Sheehan, C.E., Zitt, M., Mühlmann, G. et al. (2009) Prostate-specific membrane antigen expression in the neovasculature of gastric and colorectal cancers. *Hum. Pathol.* **40**, 1754–1761, https://doi.org/10.1016/j.humpath.2009.06.003

21 Samplaski, M.K., Heston, W., Elson, P., Magi-Galluzzi, C. and Hansel, D.E. (2011) Folate hydrolase (prostate-specific membrane [corrected] antigen) 1 expression in bladder cancer subtypes and associated tumor neovasculature. *Mod. Pathol.* **24**, 1521–1529, https://doi.org/10.1038/modpathol.2011.112

22 Barinka, C., Rojas, C., Slusher, B. and Pomper, M. (2012) Glutamate carboxypeptidase II in diagnosis and treatment of neurologic disorders and prostate cancer. *Curr. Med. Chem.* **19**, 856–870, https://doi.org/10.2174/092986712799034888

23 Bacich, D.J., Ramadan, E., O'Keefe, D.S., Bukhari, N., Wegorzewska, I., Ojeifo, O. et al. (2002) Deletion of the glutamate carboxypeptidase II gene in mice reveals a second enzyme activity that hydrolyzes N-acetylaspartylglutamate. *J. Neurochem.* **83**, 20–29, https://doi.org/10.1046/j.1471-4159.2002.01117.x

24 Hlouchova, K., Barinka, C., Konvalinka, J. and Lubkowski, J. (2009) Structural insight into the evolutionary and pharmacologic homology of glutamate carboxypeptidases II and III. *FEBS J.* **276**, 4448–4462, https://doi.org/10.1111/j.1742-4658.2009.07152.x

25 Howe, K.L., Bolt, B.J., Shafie, M., Kersey, P. and Berriman, M. (2017) WormBase ParaSite — a comprehensive resource for helminth genomics. *Mol. Biochem. Parasitol.* **215**, 2–10, https://doi.org/10.1016/j.molbiopara.2016.11.005

26 Riddle, D.L., Blumenthal, T., Meyer, B.J. and Priess, J.R. (1997) *C. elegans II*, 2nd ed, Cold Spring Harbor Laboratory Press

27 Harris, T.W., Arnaboldi, V., Cain, S., Chan, J., Chen, W.J., Cho, J. et al. (2020) WormBase: a modern model organism information resource. *Nucleic Acids Res.* **48**, D762–D767

28 Davis, P., Zarowiecki, M., Arnaboldi, V., Becerra, A., Cain, S., Chan, J. et al. (2022) WormBase in 2022—data, processes, and tools for analyzing Caenorhabditis elegans. *Genetics* **220**, https://doi.org/10.1093/genetics/iyac003

29 Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M. and Barton, G.J. (2009) Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191, https://doi.org/10.1093/bioinformatics/btp033

30 Stothard, P. (2000) The sequence manipulation suite: javascript programs for analyzing and formatting protein and DNA sequences. *BioTechniques* **28**, 1102–1104, https://doi.org/10.2144/00286ir01

31 Novakova, Z., Wozniak, K., Jancarik, A., Rais, R., Wu, Y., Pavlicek, J. et al. (2016) Unprecedented binding mode of hydroxamate-based inhibitors of glutamate carboxypeptidase II: structural characterization and biological activity. *J. Med. Chem.* **59**, 4539–4550, https://doi.org/10.1021/acs.jmedchem.5b01806

32 Webb, B. and Sali, A. (2016) Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinform.* **54**, 5.6.1–5.6.37, https://doi.org/10.1002/cpbi.3

33 Sayers, E.W., Bolton, E.E., Brister, J.R., Canese, K., Chan, J., Comeau, D.C. et al. (2022) Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **50**, D20–D26, https://doi.org/10.1093/nar/gkab1112

34 Stiernagle, T. (2006) Maintenance of C. elegans. *WormBook : Online Rev. of C Elegans Biol.* 1–11, https://doi.org/10.1895/wormbook.1.101.1

35 Mitani, S. (2009) Nematode, an experimental animal in the national BioResource project. *Exp. Anim.* **58**, 351–356, https://doi.org/10.1538/expanim.58.351

36 Ahringer, J. (2006) Reverse genetics. *WormBook: the Online Rev. of C Elegans Biol.*, Available from: https://www.ncbi.nlm.nih.gov/books/NBK19711/, https://doi.org/10.1895/wormbook.1.47.1

37 Hoogewijs, D., Houthoofd, K., Matthijssens, F., Vandesompele, J. and Vanfleteren, J.R. (2008) Selection and validation of a set of reliable reference genes for quantitative sod gene expression analysis in C. elegans. *BMC Mol. Biol.* **9**, 9, https://doi.org/10.1186/1471-2199-9-9

38 Fajtová, P., Štefanić, S., Hradilek, M., Dvořák, J., Vondrášek, J., Jílková, A. et al. (2015) Prolyl Oligopeptidase from the Blood Fluke Schistosoma mansoni: From Functional Analysis to Anti-schistosomal Inhibitors. *PLoS Negl.Trop. Dis.* **9**, e0003827, https://doi.org/10.1371/journal.pntd.0003827

39 Maulik, M., Mitra, S., Bult-Ito, A., Taylor, B.E. and Vayndorf, E.M. (2017) Behavioral phenotyping and pathological indicators of Parkinson's disease in C. elegans models. *Front Genet.* **8**, 77, https://doi.org/10.3389/fgene.2017.00077

40 Raizen, D., Song, B.M., Trojanowski, N. and You, Y.J. (2012) Methods for measuring pharyngeal behaviors. *WormBook : Online Rev. C Elegans Biol.* 1–13, https://doi.org/10.1895/wormbook.1.154.1

41 Moribe, H., Yochem, J., Yamada, H., Tabuse, Y., Fujimoto, T. and Mekada, E. (2004) Tetraspanin protein (TSP-15) is required for epidermal integrity in Caenorhabditis elegans. *J. Cell Sci.* **117**, 5209–5220, https://doi.org/10.1242/jcs.01403

42 Evans, T.C. (2006) Transformation and microinjection. *WormBook : Online Rev. of C Elegans Biol.* **10**, 1–15, https://doi.org/10.1895/wormbook.1.108.1

43 Mello, C.C., Kramer, J.M., Stinchcomb, D. and Ambros, V. (1991) Efficient gene transfer in C. elegans: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970, https://doi.org/10.1002/j.1460-2075.1991.tb04966.x

44 Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675, https://doi.org/10.1038/nmeth.2089

45 Tong, Y.-G. and Bürglin, T.R. (2010) Conditions for dye-filling of sensory neurons in Caenorhabditis elegans. *J. Neurosci. Methods* **188**, 58–61, https://doi.org/10.1016/j.jneumeth.2010.02.003

46 Mulcahy, B., Witvliet, D., Holmyard, D., Mitchell, J., Chisholm, A.D., Meirovitch, Y. et al. (2018) A Pipeline for Volume Electron Microscopy of the Caenorhabditis elegans Nervous System. *Front Neural Circuits* **12**, 94, https://doi.org/10.3389/fncir.2018.00094

47 Barinka, C., Hlouchova, K., Rovenska, M., Majer, P., Dauter, M., Hin, N. et al. (2008) Structural basis of interactions between human glutamate carboxypeptidase II and its substrate analogs. *J. Mol. Biol.* **376**, 1438–1450, https://doi.org/10.1016/j.jmb.2007.12.066

48 Klusák, V., Bařinka, C., Plechanovová, A., Mlčochová, P., Konvalinka, J., Rulíšek, L. et al. (2009) Reaction mechanism of glutamate carboxypeptidase II revealed by mutagenesis, X-ray crystallography, and computational methods. *Biochemistry* **48**, 4126–4138, https://doi.org/10.1021/bi900220s

49 Mesters, J.R., Barinka, C., Li, W., Tsukamoto, T., Majer, P., Slusher, B.S. et al. (2006) Structure of glutamate carboxypeptidase II, a drug target in neuronal damage and prostate cancer. *EMBO J.* **25**, 1375–1384, https://doi.org/10.1038/sj.emboj.7600969

50 Barinka, C., Rinnova, M., Sacha, P., Rojas, C., Majer, P., Slusher, B.S. et al. (2002) Substrate specificity, inhibition and enzymological analysis of recombinant human glutamate carboxypeptidase II. *J. Neurochem.* **80**, 477–487, https://doi.org/10.1046/j.0022-3042.2001.00715.x

51 Navrátil, M., Tykvart, J., Schimer, J., Pachl, P., Navrátil, V., Rokob, T.A. et al. (2016) Comparison of human glutamate carboxypeptidases II and III reveals their divergent substrate specificities. *FEBS J.* 2528–2545, https://doi.org/10.1111/febs.13761

52 Ferraris, D.V., Shukla, K. and Tsukamoto, T. (2012) Structure-activity relationships of glutamate carboxypeptidase II (GCPII) inhibitors. *Curr. Med. Chem.* **19**, 1282–1294, https://doi.org/10.2174/092986712799462658

53 Starich, T.A., Herman, R.K., Kari, C.K., Yeh, W.H., Schackwitz, W.S., Schuyler, M.W. et al. (1995) Mutations affecting the chemosensory neurons of Caenorhabditis elegans. *Genetics* **139**, 171–188, https://doi.org/10.1093/genetics/139.1.171

54 Perkins, L.A., Hedgecock, E.M., Thomson, J.N. and Culotti, J.G. (1986) Mutant sensory cilia in the nematode Caenorhabditis elegans. *Dev. Biol.* **117**, 456–487, https://doi.org/10.1016/0012-1606(86)90314-3

55 Gao, Y., Xu, S., Cui, Z., Zhang, M., Lin, Y., Cai, L. et al. (2015) Mice lacking glutamate carboxypeptidase II develop normally, but are less susceptible to traumatic brain injury. *J. Neurochem.* **134**, 340–353, https://doi.org/10.1111/jnc.13123

56 Jeffery, C.J. (2020) *Enzymes, pseudoenzymes, and moonlighting proteins: diversity of function in protein superfamilies*, pp. 4141–4149, Blackwell Publishing Ltd

57 Jang, Y.J., Lin, C.Y., Ma, S. and Erikson, R.L. (2002) Functional studies on the role of the C-terminal domain of mammalian polo-like kinase. *PNAS* **99**, 1984–1989, https://doi.org/10.1073/pnas.042689299

58 Mazumder, S., Ganguly, A. and Ali, N. (2010) The effect of C-terminal domain deletion on the catalytic activity of Leishmania donovani surface proteinase GP63: Role of Ser446 in proteolysis. *Biochimie* **92**, 1876–1885, https://doi.org/10.1016/j.biochi.2010.07.014

59 Boeck, M.E., Huynh, C., Gevirtzman, L., Thompson, O.A., Wang, G., Kasper, D.M. et al. (2016) The time-resolved transcriptome of C. elegans. *Genome Res.* **26**, 1441–1450, https://doi.org/10.1101/gr.202663.115

60 Kaletsky, R., Yao, V., Williams, A., Runnels, A.M., Tadych, A., Zhou, S. et al. (2018) Transcriptome analysis of adult Caenorhabditis elegans cells reveals tissue-specific gene and isoform expression. *PLos Genet.* **14**, e1007559–e, https://doi.org/10.1371/journal.pgen.1007559

61 Pereira, L., Kratsios, P., Serrano-Saiz, E., Sheftel, H., Mayo, A.E., Hall, D.H. et al. (2015) A cellular and regulatory map of the cholinergic nervous system of C. elegans. *Elife* **4**, e12432–e, https://doi.org/10.7554/eLife.12432

62 White, J.G., Southgate, E., Thomson, J.N. and Brenner, S. (1986) The structure of the nervous system of the nematode Caenorhabditis elegans. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **314**, 1–340, https://doi.org/10.1098/rstb.1986.0056

63 Trojanowski, N.F., Raizen, D.M. and Fang-Yen, C. (2016) Pharyngeal pumping in Caenorhabditis elegans depends on tonic and phasic signaling from the nervous system. *Sci. Rep.* **6**, 22940, https://doi.org/10.1038/srep22940

64 Sandhu, A., Badal, D., Sheokand, R., Tyagi, S. and Singh, V. (2021) Specific collagens maintain the cuticle permeability barrier in Caenorhabditis elegans. *Genetics* **217**, iyaa047, https://doi.org/10.1093/genetics/iyaa047

65 Schultz, R.D., Bennett, E.E., Ellis, E.A. and Gumienny, T.L. (2014) Regulation of extracellular matrix organization by BMP signaling in Caenorhabditis elegans. *PloS ONE* **9**, e101929, https://doi.org/10.1371/journal.pone.0101929

66 Sundaram, M.V. and Buechner, M. (2016) The Caenorhabditis elegans excretory system: a model for tubulogenesis, cell fate specification, and plasticity. *Genetics* **203**, 35–63, https://doi.org/10.1534/genetics.116.189357

67 Rais, R., Jiang, W., Zhai, H., Wozniak, K.M., Stathis, M., Hollinger, K.R. et al. (2016) FOLH1/GCPII is elevated in IBD patients, and its inhibition ameliorates murine IBD abnormalities. *JCI Insight* **1**, e88634, https://doi.org/10.1172/jci.insight.88634

68 Pangalos, M.N., Neefs, J.-M., Somers, M., Verhasselt, P., Bekkers, M., van der Helm, L. et al. (1999) Isolation and expression of novel human glutamate carboxypeptidases with N-acetylated α-linked acidic dipeptidase and dipeptidyl peptidase IV activity. *J. Biol. Chem.* **274**, 8470–8483, https://doi.org/10.1074/jbc.274.13.8470

69 Barinka, C., Starkova, J., Konvalinka, J. and Lubkowski, J. (2007) A high-resolution structure of ligand-free human glutamate carboxypeptidase II. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **63**, 150–153, https://doi.org/10.1107/S174430910700379X

70 Mesters, J.R. and Hilgenfeld, R. (2008) *Glutamate Carboxypeptidase II. 258*, pp. 1620–1627, John Wiley & Sons, Ltd, Chichester