Investigating cytoplasmic polyadenylation and its role in gene regulation and physiology in *Caenorhabditis elegans*

by

Zuzanna Ewa Mackiewicz

Doctoral dissertation
Laboratory of RNA Biology
International Institute of Molecular and Cell Biology in Warsaw
Warsaw PhD School in Natural and BioMedical Sciences

Supervisor: Prof. Andrzej Dziembowski
Auxiliary Supervisor: Dr. Vladyslava Liudkovska

Warsaw, January 2025

I would like to acknowledge people who contributed to the creation of this PhD thesis:

First and foremost, ***Prof. Andrzej Dziembowski****, for believing in me at the very beginning of my PhD journey, supporting my ideas, and giving me the freedom to pursue my research projects.*

Dr. Vladyslava Liudkovska *for sparking my passion for C. elegans research, sharing her knowledge and experience, and for our many inspiring discussions.*

Dr. Paweł Krawczyk *for sharing his bioinformatics expertise, which was essential for understanding and analyzing my data.*

All members of the ***Laboratory of RNA Biology*** *for creating a fantastic, inspiring environment to grow as a scientist.*

And finally, my ***Parents*** *for cheering me on through both my successes and moments of doubt.*

Thank you all!

Abstract

Gene expression is a multi-step process during which genetic information encoded in DNA is used as a template for protein production. For mRNA molecules involved in this process, the poly(A) tail is particularly crucial, as it influences their stability and the efficiency of translation. The poly(A) tail can be added both in the nucleus and in the cytoplasm. Cytoplasmic polyadenylation is catalyzed by non-canonical poly(A) polymerases belonging to the TENT family, which have been increasingly recognized for their role in gene expression regulation. Among them, the relatively recently discovered TENT5 proteins stand out. Previous studies in mouse models have demonstrated that TENT5 proteins play significant roles in various physiological processes, such as the immune response, bone formation, and gametogenesis. Moreover, mutations in the *TENT5C* gene are among the most common in multiple myeloma patients, highlighting the importance of these proteins in maintaining organismal homeostasis.

To deepen our understanding of the mechanism underlying TENT5 proteins' function, our research group analyzed the role of TENT-5 (PQN-44) – the only homolog of this family in the nematode *Caenorhabditis elegans*. *C. elegans* is a widely used model organism in molecular biology, valued for its simplicity, short life cycle, and ease of genetic manipulation. Our studies revealed that the TENT-5 poly(A) polymerase in *C. elegans* plays a key role in regulating the innate immune response by stabilizing transcripts involved in this process through the extension of their poly(A) tails. We observed that most TENT-5 substrates encode proteins secreted through the endoplasmic reticulum (ER). Additionally, we demonstrated that TENT-5 partially localizes to the ER, explaining the observed substrate preference.

The aim of my doctoral project was to expand on these findings and gain a more comprehensive understanding of cytoplasmic polyadenylation. First, I examined the differences in polyadenylation between the two nematode sexes. My analyses revealed significant differences in gene expression and poly(A) tail lengths between males and hermaphrodites. I also observed that TENT-5 is abundantly expressed in male-specific tissues, regulating components of the seminal fluid.

The next step was to investigate the potential redundancy between TENT-5 and other poly(A) polymerases: GLD-2 and GLD-4. My results excluded functional overlap between these proteins, showing that they regulate distinct groups of transcripts.

I also focused on elucidating the mechanisms directing TENT-5 to the ER. I examined the influence of three potential regulators – LARP-5, ATX-2, and C34F6.10 – on poly(A) tail profiles. Among these, I demonstrated that LARP-5 is the most likely factor regulating TENT-5 activity.

Finally, I studied the genes whose mRNAs are the most prominent substrates of TENT-5, the nematode specific NSPC genes. We discovered that these proteins localize exclusively

to a single cell – the excretory gland cell. My analyses revealed that this cell is primarily responsible for the production of NSPC proteins, which are partially involved in regulating the insulin signaling pathway in *C. elegans*.

Concluding, the results presented in my doctoral thesis provide new insights into the mechanisms of cytoplasmic polyadenylation and the role of TENT-5 in gene expression regulation. Moreover, my findings form a great foundation for further studies in this area.

Streszczenie

Ekspresja genów to wieloetapowy proces, w wyniku którego informacja genetyczna zakodowana w postaci DNA jest wykorzystywana jako matryca do produkcji odpowiednich białek. Dla pośredniczących w tym procesie cząsteczek mRNA szczególnie istotnym elementem jest ogon poli(A), który wpływa na ich stabilność oraz efektywność translacji. Ogon poli(A) może być dodawany zarówno w jądrze, jak i w cytoplazmie. Proces cytoplazmatycznej poliadenylacji jest katalizowany przez niekanoniczne poli(A) polimerazy z rodziny TENT, którym przypisuje się coraz większy udział w regulacji ekspresji genów. Spośród nich warto wyróżnić stosunkowo niedawno odkryte białka z rodziny TENT5. Dotychczasowe badania na modelach mysich wykazały, że białka TENT5 odgrywają istotną rolę w wielu procesach fizjologicznych, takich jak odpowiedź immunologiczna, formowanie kości czy gametogeneza. Mutacje w genie *TENT5C* są także jednymi z częstszych zmian genetycznych obserwowanych u pacjentów ze szpiczakiem mnogim, co podkreśla istotność tych białek dla prawidłowego funkcjonowania organizmów.

W celu pogłębienia wiedzy na temat mechanizmu działania białek TENT5, w naszej grupie badawczej podjęto się analizy funkcji TENT-5 (PQN-44) – jedynego homologa tej rodziny białek u nicieni *Caenorhabditis elegans*. *C. elegans* to popularny model w biologii molekularnej, ceniony za niewielką złożoność organizmu, krótki cykl życiowy i łatwość manipulacji genetycznych. Nasze badania wykazały, że poli(A) polimeraza TENT-5 u *C. elegans* odgrywa kluczową rolę w regulacji wrodzonej odpowiedzi immunologicznej, stabilizując transkrypty uczestniczące w tym procesie poprzez wydłużanie ich ogonów poli(A). Zauważyliśmy, że większość substratów TENT-5 koduje białka wydzielane przez retikulum endoplazmatyczne (ER). Dodatkowo udało się wykazać, że TENT-5 częściowo lokalizuje się w ER, co uzasadnia zaobserwowaną specyficzność substratową.

Mój projekt doktorski miał na celu poszerzenie tych badań i bardziej kompleksowe zrozumienie cytoplazmatycznej poliadenylacji. W pierwszej kolejności zbadałam różnice w poliadenylacji w zależności od płci nicieni. Moje analizy ujawniły istotne różnice w ekspresji genów oraz długościach ogonów poli(A) pomiędzy samcami a hermafrodytami. Zaobserwowałam również, że białko TENT-5 znajduje się w dużych ilościach w tkankach specyficznych dla samców, gdzie odpowiedzialne jest za regulację genów kodujących składniki nasienia.

Kolejnym krokiem było zbadanie potencjalnej redundancji TENT-5 z innymi poli(A) polimerazami: GLD-2 i GLD-4. Wyniki moich badań wykluczyły związek między tymi białkami, ukazując, że regulują one odmienne grupy transkryptów.

Skupiłam się także na zrozumieniu mechanizmów kierujących TENT-5 do ER. Przebadałam wpływ trzech potencjalnych regulatorów – LARP-5, ATX-2 i C34F6.10 – na profile ogonów poli(A). Wykazałam, że spośród tych białek LARP-5 jest najbardziej prawdopodobnym czynnikiem regulującym aktywność TENT-5.

Wreszcie, ważnym aspektem moich badań było przyjrzenie się bliżej genom, których mRNA reprezentują główne substraty TENT-5, czyli specyficznej dla nicieni rodzinie genów NSPC. Odkryliśmy, że białka te lokalizują się wyłącznie w jednej komórce – komórce gruczołu wydalniczego. Moje analizy wykazały, że komórka ta odpowiada przede wszystkim za produkcję NSPC, które częściowo uczestniczą w regulacji ścieżki insulinowej u *C. elegans*.

Podsumowując, wyniki przedstawione w mojej pracy doktorskiej dostarczają nowych informacji na temat mechanizmów cytoplazmatycznej poliadenylacji oraz funkcji TENT-5 w regulacji ekspresji genów, a uzyskane dane stanowią podstawę dla dalszych badań w tym obszarze.

Table of contents

1. Introduction

The 2024 Nobel Prize in Physiology or Medicine was awarded for the groundbreaking discovery of microRNAs (miRNAs) and their importance in gene regulation[1]. The laureates, Victor Ambros and Gary Ruvkun, utilized the *Caenorhabditis elegans* model organism to investigate why cells with identical DNA perform such distinct functions within an organism. Their research led to the identification of miRNAs – short RNA molecules that bind to complementary regions of mRNAs to control their expression. This fine-tuning is crucial for cellular function, and its dysregulation is implicated in severe pathologies[2,3].

This year's Nobel Prize underscores the importance of understanding post-transcriptional gene regulation. Remarkably, it is the fourth time that *C. elegans* research has been honored, further affirming the value of this model in unraveling crucial biological processes in metazoans. Together, this highlights the relevance of my PhD work, which explores the importance of post-transcriptional polyadenylation in gene regulation using *C. elegans*.

1.1. *Caenorhabditis elegans* as a model organism in molecular biology

Caenorhabditis elegans is a small, 1-mm-long nematode that has transitioned from a free-living soil worm to one of the most studied model organisms in molecular biology. It was introduced to the scientific community in 1963 by Sydney Brenner, whose work on the genetic regulation of cell division and organ development was awarded a Nobel prize in 2002[4]. The utility of *C. elegans* in biological research arises from several factors[5].

One key feature of *C. elegans* is its optimal level of complexity. Hermaphrodites consist of 959 somatic cells and 302 neurons, while males have 1033 somatic cells and 385 neurons[6]. This simplicity allows for whole-organism studies, providing valuable insights into complex biological systems. As a result, *C. elegans* was the first animal to have its entire genome sequenced[7], its cell lineage mapped[8–10], and its neuronal connectome fully charted[11,12]. Moreover, despite its distinct morphology, *C. elegans* shares remarkable molecular and genetic conservation with higher organisms, with approximately 60–80% of its proteins having human homologs[13].

C. elegans exist as self-fertile hermaphrodites (XX) and males (X0), with males comprising only about 0.05% of the population. Hermaphrodites produce both sperm and oocytes, enabling them to lay up to 300 eggs during their lifetime. Progeny then develop from an egg, through four larval stages (L1 to L4) into adulthood in about 3-4 days, making it straightforward to generate large, genetically identical populations for research. Males can mate with hermaphrodites, increasing progeny production and maintaining genetic diversity[5]. Crosses with males are often used to merge different genetic backgrounds, allowing the integration of multiple transgenes into a single animal[14].

Researchers can also utilize various tools to manipulate gene expression and generate transgenic *C. elegans* strains[15]. Key methods include RNA interference (RNAi) and CRISPR/Cas9-based approaches for precise genome editing. RNAi enables specific gene suppression via complementary double-stranded RNA, which can be introduced into the worms through injection, soaking, or feeding on RNA-expressing bacteria[16,17]. CRISPR/Cas9, on the other hand, utilizes the Cas9 protein, which can be guided to a specific site and subsequently cut a DNA strand, allowing for precise editing, including gene deletions, insertions, or substitutions[15,18].

Another significant advantage of *C. elegans* is its transparency, which makes it an excellent model for microscopy. The nematode's clear body allows for detailed observation of cellular and tissue morphology and enables real-time monitoring of dynamic changes in living organisms[19].

These numerous advantages make *C. elegans* an indispensable tool for diverse biological studies. Even while focusing on higher organisms, *C. elegans* serves as a complementary model that can simplify and accelerate preliminary experiments or illustrate the conservation of biological processes across species. The studies in this PhD thesis, investigating post-transcriptional gene regulation, are excellent examples of how incorporating *C. elegans* enhances understanding and provides novel insights into fundamental biological mechanisms.

1.2. Polyadenylation – a key process of gene regulation

In eukaryotes, gene expression begins in the nucleus, where DNA-encoded genetic information is transcribed into precursor mRNA (pre-mRNA) by RNA polymerase II. Subsequently, the freshly produced pre-mRNA molecules undergo three critical processing steps: 5′-end capping, splicing, and 3′-end cleavage and polyadenylation, which can already be initiated during transcription. Mature mRNAs are then transported to the cytoplasm to serve as a template for translation by ribosomes, enabling the production of appropriate proteins[20].

The poly(A) tail is a key determinant of mRNA fate, both in the nucleus and later in the cytoplasm. It is essential for transcript stability, their proper transport to the cytoplasm, and efficient translation[21,22]. Polyadenylation occurs through the non-templated addition of nucleotides (mostly adenosines) to the 3′-end of mRNAs by a specialized protein complex, which is also responsible for the cleavage of the 3′-end of the nascent transcript. In humans, the initial processing of the mRNA 3′-end in the nucleus includes the recruitment of around 85 proteins[23]. Among them are the canonical poly (A) polymerases (PAPs) encoded by three different genes: *PAPOLA, PAPOLB,* and *PAPOLG.* PAPs are composed of three distinct domains: a catalytic N-terminus, an RNA-binding region, and a regulatory C-terminal domain containing nuclear localization signals[24].

Once processed mRNAs reach the cytoplasm, their poly(A) tails start to undergo further modifications. They are gradually shortened during translation by deadenylating complexes PAN2-PAN3 and CCR4-NOT, which, by the removal of poly(A) tails, mark transcripts

for degradation. Degradation can proceed through the major pathway, which is initiated by mRNA cap removal at the 5′-end, or the minor pathway, which employs direct 3′ to 5′ exoribonucleolytic degradation[25,26]. However, cytoplasmic re-adenylation, mediated by non-canonical poly(A) polymerases (ncPAPs), introduces an additional regulatory layer, balancing adenylation and deadenylation in the cell[27,28]. These enzymes belong to the terminal nucleotidyltransferase (TENT) family and, unlike canonical PAPs, lack an intrinsic RNA recognition motif and often rely on RNA-binding proteins for substrate specificity[29]. Moreover, some TENT members can also incorporate non-adenosine nucleotides into the poly(A) tails, adding regulatory diversity: adenosine addition usually stabilizes mRNAs, uridylation marks them for degradation, and mixed adenosine/guanosine tailing increases resistance to deadenylation[28]. In vertebrates, the TENT family includes eight ncPAPs (TENT2, TENT4A-B, TENT5A-D, and TENT6/MTPAP) and three terminal uridyltransferases (TUTases) (TENT1, TENT3A/TUT4, and TENT3B/TUT7). In *C. elegans*, there are three ncPAPs (GLD-2, GLD-4, TENT-5/PQN-44), five TUTases (USIP-1, PUP-2, PUP-3, CDE-1, and MUT-2)[27], and several other predicted TENT proteins (GLDR-2, F43H9.3, C53A5.16, and F43E2.1)[30].

Cytoplasmic polyadenylation represents an essential post-transcriptional modification, vital for accurate gene expression regulation. Moreover, recent studies by our group reveal that members of the TENT5 family polyadenylate mRNA from the externally introduced Moderna SARS-CoV-2 vaccine[31]. These findings highlight the broader implications of cytoplasmic polyadenylation mechanisms, offering potential advancements in biomedical research and therapeutic development.

1.2.1. Cytoplasmic polyadenylation in *C. elegans*

As mentioned, cytoplasmic polyadenylation in *C. elegans* is mediated by three non-canonical poly(A) polymerases: GLD-2, GLD-4, and TENT-5[27]. Among them, the best studied is germline development defective-2 (GLD-2), which was discovered in *C. elegans*[32]. It plays a crucial role in germline physiology by regulating numerous mRNAs through its poly(A) elongating activity[33]. Studies of *gld-2* mutants reveal GLD-2 involvement in processes such as the transition from mitotic to meiotic cell cycles, meiotic progression, and gamete production. Surprisingly, GLD-2 itself exhibits minimal polyadenylation activity due to the absence of an RNA-binding domain, a common feature among TENT family members[34]. For activation, GLD-2 requires molecular partners that can vary depending on its specific roles. To date, two proteins, GLD-3 and RNP-8, have been shown to interact with GLD-2. By binding with the appropriate partner, GLD-2 can control gamete specification. Initially, together with GLD-3, it promotes spermatogenesis, but later switches to RNP-8-mediated activation of oogenesis[35]. Among many potential GLD-2 targets, the most significant and best-studied is *gld-1* mRNA, which encodes a repressor of mitosis-promoting

mRNAs. By regulating *gld-1,* GLD-2 orchestrates the essential process of mitosis-to-meiosis transition in the germline[36].

The function of GLD-4 is less understood, but it also appears to be mainly engaged in the regulation of germline development[37]. Its activity and recruitment of mRNA targets depend on GLS-1, which is a direct interactor of GLD-3, linking these two GLD-mediated regulatory pathways. Both GLD-4 and GLS-1 are localized to germline-specific P granules, which are ribonucleoprotein complexes important for RNA regulation and germ cell identity. Similar to GLD-2/GLD-3, GLD-4/GLS-1 acts on *gld-1* mRNA to prevent its degradation and promote translation. However, the effects of GLD-4 on poly(A) tail length and gene expression are less pronounced than those of GLD-2, which correlates with the relatively mild phenotypes observed in *gld-4* mutants. While *gld-2* mutants are sterile[32], *gld-4* mutants exhibit only partial reductions in hermaphrodite fertility and maintain normal male fertility[37]. Instead, GLD-4 is suggested to play a role in promoting translation, as deletion of either GLD-4 or its co-factor GSL-1 reduces polysome formation[38]. Despite these insights, the precise mechanisms underlying GLD-4 activity and its potential functions outside the germline remain to be established.

As presented, while GLD-2 and GLD-4 are relevant for germline development, our global understanding of non-canonical polyadenylation and how these three ncPAPs interplay to maintain the balance of poly(A) tail metabolism remains incomplete.

1.2.2. Family of TENT5 poly(A) polymerases

In mammals, the TENT5 family includes four proteins: TENT5A, TENTB, TENT5C, and TENT5D, which display tissue- and organ-specific expression patterns[39]. While all members function as non-canonical poly(A) polymerases that regulate mRNA stability and expression, their physiological roles are diverse. Despite the relatively recent discovery of TENT5 proteins, their functions in various organisms have already been extensively explored, with much of this research conducted by members of our group[40–44].

For instance, TENT5A is crucial for proper bone formation[42]. It is expressed in osteoblasts and osteocytes and predominantly polyadenylates mRNAs encoding collagens and other proteins involved in bone development. As a result, the absence of TENT5A in mice leads to smaller body size, abnormal posture, and fragile bones, mirroring the osteogenesis imperfecta bone disease often observed in humans with TENT5A mutations[45]. Moreover, TENT5A, together with TENT5C, is expressed in murine macrophages, where it enhances the innate immune response by elongating the poly(A) tails of transcripts involved in the defense process[46]. Interestingly, while single knockouts of *Tent5a* or *Tent5c* in mice are viable, their combined deletion leads to lethality, hinting at their functional redundancy.

TENT5 family members also influence gametogenesis in mice[43,47]. It has been shown that TENT5B and TENT5C act redundantly to polyadenylate mRNAs necessary for proper oocyte formation, while TENT5C and TENT5D are expressed in spermatocytes and spermatozoa,

affecting different aspects of spermatogenesis. Consequently, double knockouts of *Tent5b* and *Tent5c* cause female infertility, whereas *Tent5c* and *Tent5d* deficiencies result in male infertility[43].

Among TENT5 proteins, TENT5C is the most extensively studied so far. Notably, it is one of the most frequently mutated genes in multiple myeloma (MM) and acts as a growth suppressor in MM cells by polyadenylating mRNAs encoding endoplasmic reticulum (ER)-targeted proteins, inducing ER stress and apoptosis[40]. Additionally, TENT5C mediates humoral immune responses by polyadenylating mRNAs of immunoglobulins (Ig) in B lymphocytes, with its knockout reducing Ig protein levels and antibody production[41]. Moreover, *Tent5c* disruption causes microcytic anemia in mice by impairing polyadenylation of globin mRNAs critical for red blood cell maturation[44].

Despite the functional diversity of TENT5 proteins, their molecular targets share similar characteristics. The majority of identified TENT5 targets, such as collagens, immunoglobulins, and immune effectors possess ER-targeting signal peptides, indicating that TENT5 activity preferentially targets mRNAs encoding ER-secreted proteins. However, the exact mechanisms underlying this specificity remain unclear. Recent findings suggest that TENT5 proteins might be recruited to the ER[48], implying that their localization may partly dictate their substrate specificity.

1.2.3. Methods used for studying polyadenylation

There are multiple established methods for measuring poly(A) tail lengths, that provide critical insights into this complex aspect of gene regulation. These approaches are generally classified into traditional poly(A) length assays and high-throughput sequencing-based techniques[49].

Initially, poly(A) tail analyses relied on low-throughput methods such as the RNase H/oligo(dT) assay or the polymerase chain reaction (PCR) amplification of the 3′-ends. The first method involves the hybridization of oligo(dT) to the poly(A) tail and subsequent digestion of the resulting RNA-DNA duplexes by RNase H[50]. The information about the poly(A) length can then be retrieved by comparing RNase H-treated and untreated samples. PCR-based techniques, such as the Rapid Amplification of cDNA End-Poly(A) Tail assay (RACE-PAT), involve hybridization with oligo(dT) anchor, reverse transcription, and PCR amplification using anchor- and gene-specific primers[51]. However, PCR-based approaches often struggle with amplifying long homopolymers and introduce biases, which lowers the accuracy of poly(A) length estimations and limits their usefulness for detecting subtle poly(A) tail changes. While many improved variants of these methods exist, they remain constrained to single-transcript analyses and are now primarily used for validation purposes.

The development of Next Generation Sequencing (NGS) and increasing interest in poly(A) metabolism accelerated the development of high-throughput ways to study poly(A) tail lengths on a broader scale. Techniques like PAL-Seq[52], TAIL-Seq[53], and PAT-Seq[54] are based

on Illumina sequencing and involve ligation of 3′-end adaptor to RNA, partial RNA fragmentation with RNase T1, 3′-end enrichment, and ligation of a 5′-end adaptor, followed by reverse transcription, PCR amplification, and sequencing. Although extending poly(A) studies to a larger scale, these methods are still hindered by PCR-related biases and are not adequate for investigating longer poly(A) tails.

Therefore, among recent innovations, Oxford Nanopore Direct RNA Sequencing (DRS) stands out as the most suitable approach for poly(A) tail profiling. Importantly, unlike previous methods, Nanopore DRS is PCR-free, allowing unbiased sequencing of full-length RNA molecules. Library preparation includes subsequent ligation of two adapters to polyadenylated RNAs – first, a 3′-end adapter and then a sequencing adapter attached to a motor protein that facilitates the passage of RNAs through protein pores embedded in a sequencing flow cell. The RNA molecule traversing the pore induces current changes that can be later translated into nucleotide sequences using computational tools[55]. Moreover, this method provides additional data on differential expression, isoforms, and RNA modifications, making it a versatile tool that has already been successfully used to study poly(A) tail dynamics across various organisms. However, as Nanopore DRS is a relatively new method, its implementation was not possible in earlier studies on cytoplasmic polyadenylation, significantly limiting their scope.

Due to its numerous advantages, Nanopore DRS has been implemented in our group and in this PhD work as a standard method for poly(A) tail profiling and the identification of poly(A) polymerase targets.

2. Background of the study and preliminary results

2.1. Function of TENT-5 poly(A) polymerase in *C. elegans* (Manuscript 1)

When I began my PhD, one of the ongoing projects in the group focused on studying TENT-5 (PQN-44) – the single homolog of the TENT5 family in *C. elegans.* It has already been confirmed that, in agreement with the high sequence similarity of nematode TENT-5 to mammalian TENT5 members, it also possesses poly(A) polymerase activity, which enhances mRNA expression through poly(A) tail elongation. Moreover, expression analysis revealed that TENT-5 is active throughout all developmental stages and tissues, underscoring its essential role in nematode physiology.

To explore its functional significance, transcriptomic and proteomic analyses of *tent-5* mutants were performed, and they uncovered a notable downregulation of innate immune effectors, later supported by increased susceptibility of *tent-5* mutants to various bacterial pathogens. Furthermore, the function of TENT-5 as a regulator of immunity is evolutionarily conserved, as demonstrated by my colleagues for the mammalian orthologs, TENT5A and TENT5C, in the macrophages.

To understand the mechanism of TENT-5-mediated regulation of pathogen resistance, the Nanopore Direct RNA Sequencing (DRS) was utilized to investigate the polyadenylation dynamics in *tent-5* mutants. Among the 96 transcripts exhibiting significant poly(A) tail shortening, many encoded immune response proteins. This shortening correlated with reduced mRNA stability, supporting the hypothesis that TENT-5 activity positively modulates transcript stability via polyadenylation. Additionally, apart from defense response genes, TENT-5 strongly regulated transcripts encoding NSPC proteins, members of a nematode-specific family with previously unknown functions. Notably, both immune effectors and NSPCs contain signal peptides directing them toward the secretory pathway through the endoplasmic reticulum (ER). We demonstrated that in the intestine TENT-5 partially localizes in the ER, suggesting that TENT-5 substrate specificity is driven, at least to some extent, by its spatial proximity to mRNAs encoding secreted proteins. However, the exact mechanism of this regulation was left undiscovered.

Since I joined the project in its concluding stages, my contribution was primarily during the revision of the publication. One of my tasks involved performing lifespan analyses of *tent-5* mutants grown on UV- or heat-killed *E. coli.* These experiments demonstrated that the shortened lifespan observed with live, mildly pathogenic *E. coli* resulted from a compromised immune response in the absence of TENT-5. Additionally, I contributed to colony-forming unit (CFU) analyses, which investigated bacterial accumulation in the intestines of wild-type and mutant worms. The significantly higher bacterial load in *tent-5* mutants further confirmed the critical role of TENT-5 in responding to pathogen infection. Through genetic crossing, I also generated a strain co-expressing GFP-tagged TENT-5

and CemOrange2 fused to the ER marker protein TRAM-1. This strain was subsequently used in microscopy studies to demonstrate the ER enrichment of TENT-5, shedding light on its possible target recognition mechanism.

The results of this project provide the first characterization of TENT-5 activity in *C. elegans*, offering valuable insights into both nematode physiology and the regulation of poly(A) tail metabolism. Moreover, this study underscores the conserved function of TENT-5 as a regulator of innate immunity across species, highlighting the value of multi-organismal studies in understanding the mechanisms and significance of specific gene regulation pathways.

This study has been published in *Science Advances* and is enclosed in the appendix to this thesis as **Manuscript 1**.

Reference – Manuscript 1:

Liudkovska V., Krawczyk P.S., Mroczek S., Bilska A., Gumińska N., Cysewski D., **Mackiewicz Z.**, Ewbank J.J., Drabikowski K., Dziembowski A. TENT5 cytoplasmic non-canonical poly(A) polymerases regulate the innate immune response in animals. *Sci. Adv.* **8**, eadd9468 (2022). DOI:10.1126/sciadv.add9468

3. Research objectives

Building on the findings described in the previous chapter, my PhD project sought to investigate TENT-5-mediated regulation in *C. elegans* within a broader biological context.

Firstly, complementing our study on hermaphrodites, I aimed to examine poly(A) tail metabolism in males to gain a more comprehensive understanding of cytoplasmic polyadenylation and its potential dependence on sex.

Moreover, to explore how poly(A) polymerases interact to maintain the balance of poly(A) tail metabolism, I planned to investigate potential redundancy between the three poly(A) polymerases in *C. elegans*: TENT-5, GLD-2, and GLD-4.

Further, I aimed to uncover the factors driving the ER localization of TENT-5 and identify other proteins that might be involved in shaping its substrate specificity.

Finally, I wanted to investigate the cause of the pronounced TENT-5-mediated polyadenylation of NSPC family transcripts. Additionally, preliminary data indicated that NSPCs are expressed specifically in the excretory gland cell, whose function remained unknown, prompting me to extend the study's objectives to explore the role of this cell in *C. elegans* physiology.

4. Results

4.1. Function of TENT-5 in *C. elegans* males – sexual dimorphism in poly(A) tail metabolism (Manuscript 2)

While familiarizing myself with the *C. elegans* model, I noticed that most studies in the field are limited to experiments on mixed populations, which are predominantly composed of hermaphrodites. *C. elegans* populations consist primarily of self-fertile hermaphrodites, with males constituting only about 0.05% of the population[5]. Consequently, males are often overlooked in nematode research. However, it has been known for a long time, that sex has a crucial influence on organismal physiology and genetics, determining for example disease susceptibility or treatment responses[56]. Therefore, limiting studies to one sex may reduce the reliability of findings, a critical issue in biomedical research. Intriguingly, I observed that poly(A) tail metabolism, despite being one of the most important parts of gene regulation machinery, remains underexplored in the context of sexual dimorphism. To address this gap, I used *C. elegans* as a model to investigate how polyadenylation differs between hermaphrodites and males at the organismal level.

To achieve this, I employed Nanopore Direct RNA Sequencing (DRS), the most reliable method for poly(A) tail length estimation[49], to simultaneously analyze differences in poly(A) tail metabolism and gene expression between hermaphrodites and males. The sex-dependent differences in gene expression are well-documented across various species, including nematodes. However, prior studies on *C. elegans* males relied only on single worms[57] or male-enriched *him* mutants[58], compromising the reliability of results. The advantage of my approach was using a hand-picked population of wild-type males, which resulted in the creation of an unbiased and valuable dataset for future research. My differential expression analysis confirmed many known sex-related differences, with significant overlap with previous findings[57,58]. Notably, I identified a group of genes uniquely expressed in males, most of which are still uncharacterized. These genes predominantly encode proteins localized in two male-specific tissues – the seminal vesicle and vas deferens, which produce seminal fluid that is transferred along with the sperm to the hermaphrodite uterus during mating. The secretory function of these male-specific tissues aligns with their pronounced enrichment in ER, as well as the observation that most male-specific genes are predicted to encode secreted proteins. Given that this characteristic resembles the known properties of TENT-5 targets, I hypothesized that these male-specific transcripts might also be regulated by TENT-5.

Another indication that TENT-5 might be important for male-specific gene regulation emerged from the analysis of poly(A) tail length differences between *C. elegans* hermaphrodites and males. I found that males generally have transcripts with longer poly(A) tails than hermaphrodites, particularly for genes associated with the ER function. This observation supported my hypothesis that TENT-5 might play a crucial role in regulating male-specific transcripts. To test this, I examined TENT-5 localization in males and observed

strong expression in spermatids, seminal vesicle, and vas deferens, consistent with its potential activity in these tissues. To validate this, I conducted additional Nanopore DRS analysis on *tent-5* mutant males and compared their gene expression and poly(A) tail profiles to those of wild-type males. I noticed that TENT-5 deletion leads to the downregulation of defense response genes not only in hermaphrodites but also in males, proving its universal role in immunity. Moreover, in line with my hypothesis, *tent-5* mutant males exhibited significant shortening of the poly(A) tails, particularly for male-specific transcripts. Despite this, I did not observe any reproduction-related phenotypes, such as altered mating behavior, decreased fertility, or sperm morphology, between wild-type and *tent-5* mutant males.

My results underscore the importance of studying protein functions in both sexes, since, as I show, they can differ substantially. Interestingly, TENT-5 targets display minimal overlap between hermaphrodites and males, with only two genes – *heh-1* and *far-2* – commonly regulated in both sexes. I propose that TENT-5 prioritizes the regulation of male-specific proteins, which are components of seminal fluid, to support the successful mating of males. As a result, its activity toward immune effectors appears to be less pronounced compared to hermaphrodites. A potential sex-independent role of TENT-5 may involve the regulation of cholesterol metabolism, as two of its common targets are associated with cholesterol and fatty acid transport. A more detailed discussion of this possible function is provided in subsequent chapters of this thesis.

That part of my PhD was particularly challenging due to the limited understanding of male *C. elegans* physiology and genetics. Most of the TENT-5 targets in males remain unannotated, making it difficult to predict their function. The same applies to seminal fluid, the role of which in ensuring mating success has been only superficially addressed in prior studies. Despite this, I conducted the first analysis of sex-dependent polyadenylation profiles in *C. elegans*, thereby producing a valuable resource for future research. Additionally, I advanced our understanding of TENT-5 functionality by demonstrating its sex-specific roles in nematodes.

This study has been published as a pre-print in *bioRxiv* and is enclosed in the appendix to this thesis as **Manuscript 2**.

Reference – Manuscript 2:

Mackiewicz Z., Liudkovska V., Dziembowski A. TENT-5 regulates the expression of male-specific genes in *Caenorhabditis elegans*. *bioRxiv* (2024). DOI: https://doi.org/10.1101/2024.06.18.599341

4.2. Relationship between three non-canonical poly(A) polymerases: TENT-5, GLD-2, and GLD-4 (unpublished results)

As mentioned earlier, in *C. elegans,* there are three non-canonical poly(A) polymerases – GLD-2, GLD-4, and TENT-5. To date, each of these enzymes has been studied independently, with GLD-2 and GLD-4 shown to play roles in germline development[37], and TENT-5 in regulating innate immunity later in the nematode's lifespan[46]. While partial synergy between

GLD-2 and GLD-4 has been previously speculated[38], it has not been thoroughly explored. Despite the clear germline-related activity of GLD-2 and GLD-4, their expression is also observed in somatic tissues (based on the CeNGEN database[59] and personal communications), suggesting the possibility of their expanded activity outside the germline. Furthermore, TENT-5 is expressed in the germline, although at levels significantly lower than GLD-2 and GLD-4[60]. Its expression is markedly lower than GLD-2 and moderately lower than GLD-4[60], raising the possibility of germline-related roles for TENT-5. Therefore, we aimed to investigate the potential interplay and relationship between these three ncPAPs in whole-worm samples.

First, I analyzed the poly(A) profiles of *gld-2* and *gld-4* mutants using Nanopore DRS, a method unavailable during earlier studies on GLD-2 and GLD-4. My analysis revealed that poly(A) tail length differences were most pronounced in *gld-2* mutant worms (Figure 1A), even though, due to the sterility of homozygotes, the heterozygous population had to be used for sequencing. This finding proves the central role of GLD-2 in poly(A) tail regulation, which is also manifested in the strongest phenotypes of its mutants, such as delayed entry into meiosis and lack of functional gamete production[38]. The global poly(A) distribution in *gld-2* mutants did not show a median length change compared to the wild type (Figure 1B). However, a significant change was observed after narrowing the analysis down to the germline-enriched transcripts (Figure 1C). Moreover, examining the global poly(A) distribution, it is evident that GLD-2 deletion results in poly(A) shortening for some transcripts and elongation for others. This effect could be explained by the active production of newly transcribed mRNAs with tails that have not yet undergone deadenylation. Comparing my results with previously identified GLD-2 targets (*gld-1, zfp-3, lip-1, oma-2*, *him-5*, *cpb-3*, *gla-3*, and *puf-8*)[33], I found that only *gld-1, oma-2,* and *cpb-3* showed significant shortening in the *gld-2* mutant (Figure 1A). This suggests that previous studies[33,38], which used Illumina RNA sequencing or RT-qPCR to identify downregulated transcripts under GLD-2 deficiency, may have misinterpreted some as direct GLD-2 targets due to secondary transcriptional effects unrelated to poly(A) regulation. Nevertheless, my findings confirm that *gld-1* is a strong GLD-2 target, with its poly(A) tail shortened by 21 nucleotides in the mutant (Figure 1D). Moreover, functional enrichment analysis of other GLD-2 targets further supports its primary role in germline development (Figure 1E).

Similarly, I examined poly(A) tail dynamics in *gld-4* mutant (Figure 1F). The observed differences in poly(A) tail lengths were rather small, confirming that GLD-4 mediates the addition of shorter adenosine stretches than GLD-2[38]. Surprisingly, although GLD-4 was predicted to have the smallest impact on poly(A) tail metabolism[38], only for this poly(A) polymerase did I observe a significant global poly(A) tail shortening (Figure 1B). Functional enrichment analysis revealed that GLD-4 seems to polyadenylate mainly mRNAs encoding ribosomal proteins, likely enhancing their stability and ensuring efficient translation (Figure 1G). These findings align with previous predictions of GLD-4's role in promoting translation and polysome formation[38], as evidenced by the enrichment of most GLD-4 targets

in the polysome fraction (Figure 1H). Additionally, the significant poly(A) shortening observed for CEY family members, which are proteins involved in polysome formation[61], further supports its association with this process.

Interestingly, I observed that TENT-5 targets, similarly to GLD-4-regulated transcripts, are significantly enriched in the polysome fraction (Figure 1H). To investigate the potential overlap between mRNAs regulated by GLD-4 and TENT-5, I generated the *gld-4/tent-5* double mutant worms and performed Nanopore DRS (Figure 1I). As expected, the data showed significant poly(A) tail shortening for both GLD-4 and TENT-5 targets. However, I did not observe any transcripts behaving differently depending on the presence of the second poly(A) polymerase, thereby excluding their functional redundancy.

To further elaborate on redundancy among the three polymerases, I compared the transcript pools with significantly shortened poly(A) tails in *tent-5, gld-2,* and *gld-4* mutants (Figure 1J). The overlaps were minimal, even between GLD-2 and GLD-4, both associated with germline development. Among the 115 shared GLD-2 and GLD-4 targets, most were sperm-enriched and functionally connected to the pseudopodium, the structural part of the sperm responsible for its mobility[62] (Figure 1K). Therefore, I assume that GLD-2 and GLD-4 might co-regulate sperm cell specification, but their other functions are rather distinct. Surprisingly, based on my Nanopore results, even *gld-1* mRNA is not a common target of both poly(A) polymerases, as indicated before[37] (Figure 1D). Similarly, there was negligible overlap between targets of TENT-5 and GLD-4 or TENT-5 and GLD-2, again indicating different roles of the three ncPAPs. Slightly larger overlaps were observed for genes downregulated upon deletion of each polymerase (Figure 1J). Most of the genes downregulated in *gld-4* mutants and many in *tent-5* mutants were also downregulated upon GLD-2 depletion, indicating that while their direct substrates might be distinct, their broader regulatory effects may converge.

In conclusion, my findings indicate that redundancy among the three *C. elegans* poly(A) polymerases is highly unlikely. However, as shown in earlier chapters, even a single polymerase's activity can vary depending on factors like sex. This explains why distinct polymerases could have largely non-overlapping functions. Nevertheless, some further experiments could be performed to provide a more comprehensive understanding of the activity of poly(A) polymerases. For instance, the degron-based approach could be used for precisely deleting all three poly(A) polymerases. Moreover, given TENT-5 enrichment and activity in germline-located spermatids and male-specific tissues, I assume redundancy of TENT-5 with either GLD-2 or GLD-4 might be different in males. While such investigations lie beyond the scope of this thesis, I believe my work has clarified some aspects of cytoplasmic polyadenylation in *C. elegans* and provides a foundation for future research in this field.

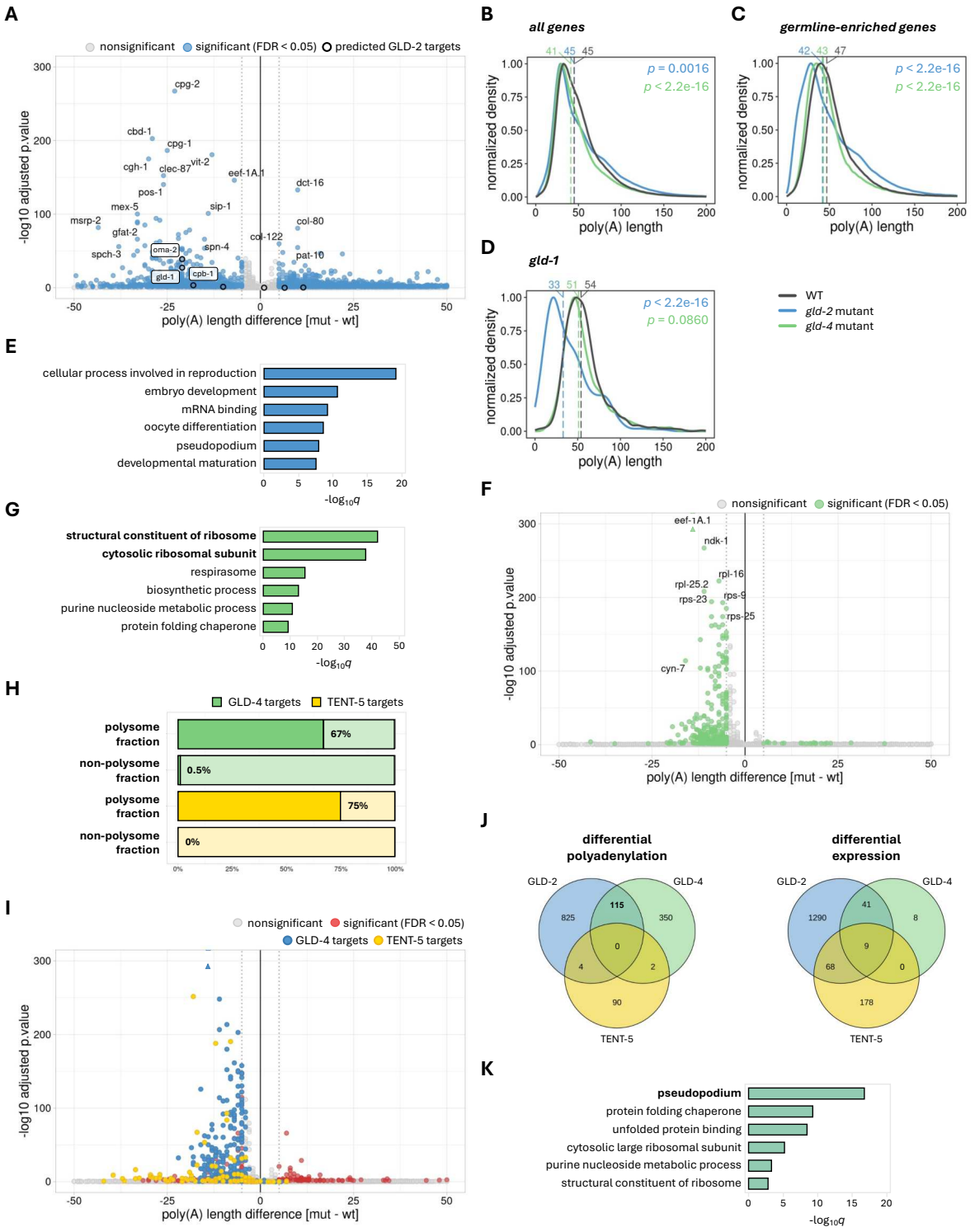


Figure 1. Relationship between three non-canonical poly(A) polymerases: TENT-5, GLD-2, and GLD-4.
(A) Volcano plot displaying differential polyadenylation between wild-type and *gld-2* mutant worms. Transcripts with significantly changed median poly(A) tail length (FDR < 0.05) by a minimum of 5 nucleotides (dotted lines) are marked with blue dots. Predicted GLD-2 targets from Nousch M. *et al.* 2017[33] are marked with black borderlines.
(B) Density plot showing global differences in the poly(A) tail distribution between wild-type (dark gray), *gld-2* mutant (blue), and *gld-4* mutant worms (green). Vertical dashed lines represent the median poly(A) tail length for each condition (in nucleotides). The plot was generated for all identified transcripts and normalized to 1.
(C) Density plot showing differences in the poly(A) tail distribution for germline-enriched transcripts between wild-type (dark gray), *gld-2* mutant (blue), and *gld-4* mutant worms (green). Vertical dashed lines represent

the median poly(A) tail length for each condition (in nucleotides). The plot was generated for germline-enriched transcripts described in Reinke V. *et al.* 2004[63] and normalized to 1.
(D) Density plot showing differences in the poly(A) tail distribution for the *gld-1* mRNA between wild-type (dark gray), *gld-2* mutant (blue), and *gld-4* mutant worms (green). Vertical dashed lines represent the median poly(A) tail length for each condition (in nucleotides). The plot was normalized to 1.
(E) Top GO terms for genes with poly(A) tails significantly shortened in *gld-2* mutant ordered by adjusted *p*-value.
(F) Volcano plot displaying differential polyadenylation between wild-type and *gld-4* mutant worms. Transcripts with significantly changed median poly(A) tail length (FDR < 0.05) by a minimum of 5 nucleotides (dotted lines) are marked with green dots. Triangles represent data points outside the *y*-axis limit.
(G) Top GO terms for genes with poly(A) tails significantly shortened in *gld-4* mutant worms ordered by adjusted *p*-value. Terms related to ribosomes are shown in bold.
(H) Percentages of GLD-4 and TENT-5 targets enriched in polysome and non-polysome fractions based on the dataset from Nousch M. *et al.* 2014[38]. The list of TENT-5 targets comes from Liudkovska V. *et al.* 2022[46].
(I) Volcano plot displaying differential polyadenylation between wild-type and *gld-4/tent-5* double mutant worms. Transcripts with significantly changed median poly(A) tail length (FDR < 0.05) by a minimum of 5 nucleotides (dotted lines) are marked with red dots. Triangles represent data points outside the *y*-axis limit. Additionally, GLD-4 and TENT-5 targets (mRNAs with significantly shorter poly(A) tail in *gld-4* or *tent-5* mutant worms by a minimum of 5 nucleotides; FDR < 0.05) are marked with blue and yellow dots, respectively. The list of TENT-5 targets comes from Liudkovska V. *et al.* 2022[46].
(J) Venn diagram showing overlaps between transcripts with significantly shorter poly(A) tail by a minimum of 5 nucleotides (left) or significantly downregulated (right) for GLD-2 (blue), GLD-4 (green), and TENT-5 (yellow). Data for TENT-5-regulated genes comes from Liudkovska V. *et al.* 2022[46].
(K) Top GO terms for 115 genes with poly(A) tails significantly shortened in both *gld-2* and *gld-4* mutant worms ordered by adjusted *p*-value. The term related to spermatogenesis is shown in bold.

4.3. Possible mechanism of TENT-5 target recognition (unpublished results)

Although comprehensive studies on the TENT5 family of proteins have been conducted across multiple species, the exact mechanism of its substrate recognition remains to be determined. Since TENTs lack an RNA-recognition domain, in many cases they depend on interactions with other RNA-binding proteins to recruit them to their substrates. For instance, in *C. elegans*, the GLD-2 poly(A) polymerase interacts with GLD-3 or RNP-8 to polyadenylate specific transcript pools, influencing the decision between sperm and oocyte cell fates[64]. In our research group, we aim to elucidate the target-recognition mechanism for TENT5 proteins. Previous studies have identified potential TENT5 interactors, including FNDC3A/B (fibronectin type-III domain-containing protein 3A/B)[48], which are particularly intriguing due to their ER-transmembrane localization. FNDC3A/B proteins possess a disordered N-terminal domain, nine fibronectin type-III domains, and a C-terminal transmembrane domain and are hypothesized to activate TENT5 by localizing it to the ER.

To validate these findings, my colleagues employed TurboID proximity labeling in B lymphocytes expressing a single TENT5C protein. This approach identified 29 potential TENT5C interactors, with FNDC3A/B, LARP4/5, and ATXN2 being particularly interesting. LARP4 and LARP5 are known to bind mRNAs and protect their poly(A) tails from deadenylation[65], while ATXN2 has been implicated in recruiting the TENT2 polymerase to its substrates[66]. The interaction between TENT5C and FNDC3A/B was additionally confirmed in other TENT5-expressing models established in our laboratory, including mouse testes and bone marrow-derived macrophages (BMDMs). Moreover, Nanopore sequencing following FNDC3A/B silencing in both B lymphocytes and BMDMs revealed a clear

shortening of poly(A) tails of the TENT5-regulated mRNAs, supporting the hypothesis of FNDC3A/B involvement in TENT5-mediated polyadenylation.

A similar TurboID experiment was also performed for erythroblasts, where TENT5C primarily polyadenylates globin mRNAs, playing a critical role in red blood cell development. As globin mRNAs are not translated on the ER, FNDC3A/B-mediated localization of TENT5C to the ER seemed unnecessary in this context. Consequently, FNDC3A/B were not detected as TENT5 interactors in erythroblasts. Instead, LARP proteins, mentioned above, emerged as the primary interactors. In particular, LARP5 deletion resulted in significant shortening of globin poly(A) tails, which were not affected further by additional TENT5C deletion, suggesting the critical role of LARP5 in regulating TENT5 activity[44].

Based on these findings, I decided to investigate whether the mechanism of TENT5 target recognition, mediated by either FNDC3A/B, LARP4/5, or ATXN2, is conserved across species. For that purpose, I identified the *C. elegans* homologs of these interactors: LARP-5, ATX-2, and C34F6.10, a putative FNDC3A/B homolog[30]. However, none of these proteins have been studied in the context of polyadenylation regulation in worms. To evaluate their roles, I performed RNAi silencing of *larp-5*, *atx-2*, and *C34F6.10* and conducted poly(A) profiling using Nanopore DRS (Figure 2A). This revealed significant poly(A) tail shortening for 125, 48, and 14 transcripts, respectively, upon depletion of these proteins, highlighting LARP-5 as the most influential regulator of polyadenylation dynamics in *C. elegans* (Figure 2A and 2B).

Further comparison of transcripts affected by silencing of each protein with known TENT-5 targets showed that *larp-5*, *atx-2*, and *C34F6.10* silencing resulted in the shortening of poly(A) tails for 76%, 28%, and 31% of TENT-5 targets, respectively (Figure 2A). Notably, the remaining 24% of TENT-5 targets were affected by *larp-5* silencing but lacked statistical significance, strengthening the link between LARP-5 and TENT-5. Moreover, since LARP-5 also influenced genes unrelated to TENT-5, its activity is likely not limited to TENT-5 regulation. Importantly, the absence of overlap between LARP-5-affected genes and those regulated by GLD-2 or GLD-4 (Figure 2C) excludes its involvement in parallel polyadenylation pathways. Instead, the observed shortening for genes other than those TENT-5-regulated may result from reduced poly(A) tail protection by LARP-5, leading to increased deadenylation.

In contrast, ATX-2 exhibited a much weaker influence on TENT-5 substrate mRNAs (Figure 2A and 2B). However, as ATX-2 has been implicated in promoting germline proliferation and oocyte fate in nematodes and TENT2 substrate recognition in vertebrates, it is more likely associated with GLD-2 functionality. Despite such indications, my analyses showed minimal overlap between GLD-2/GLD-4 targets and ATX-2-affected transcripts (Figure 2C), suggesting that the role of ATX-2 in germline development is unrelated to polyadenylation regulation.

Unexpectedly, silencing of the FNDC3A/B homolog C34F6.10 triggered minimal effects on poly(A) tails (Figure 2A). This result might be explained by less effective silencing, as confirmed by RT-qPCR (Figure 2D). To address this, I designed and ordered a transgenic

strain with a C34F6.10 knockout and repeated Nanopore sequencing (Figure 2E). Surprisingly, the knockout strain confirmed this minimal impact of C34F6.10, identifying only 18 potential mRNA substrates, nine of which were consistent across experiments. Additionally, the C34F6.10 knockout did not alter the expression of even a single gene, suggesting its limited role in gene regulation.

To further investigate the potential interaction between C34F6.10 and TENT-5, I visualized TENT-5 localization in *C. elegans* following C34F6.10 deletion. I did not observe any changes in TENT-5 distribution within cells or in its overall expression levels (Figure 2F). Subsequently, I generated and sequenced mutant worms lacking both TENT-5 and C34F6.10 to assess their combined impact on polyadenylation (Figure 2G). Interestingly, TENT5-mediated polyadenylation appeared partially dependent on C34F6.10 activity, as there was only a 38% overlap between TENT5-regulated genes in the presence or absence of C34F6.10. Notably, most genes with shorter poly(A) tails in the C34F6.10 mutant were expressed in the reproductive tissues of *C. elegans* (Figure 2H). Additionally, as most of them are also regulated by LARP-5, I hypothesize that TENT-5 activity for this subset of transcripts may be co-regulated by both LARP-5 and C34F6.10.

In conclusion, my findings identify LARP-5 as the primary TENT-5 cooperator in *C. elegans*. This is surprising, as in mice, LARPs predominantly regulate TENT-5 in erythroblasts, where TENT-5 exceptionally polyadenylates mRNAs encoding cytoplasmic proteins – hemoglobin subunits. In both *C. elegans* hermaphrodites and males, most TENT-5 targets are ER-translated, directing attention rather towards FNDC3A/B-mediated regulation. However, it remains plausible that C34F6.10 is not a true FNDC3A/B functional homolog. Further research into its localization, activity, and interactions is required to clarify its role. Alternatively, the C34F6.10 relevance to TENT-5 activity may be tissue-specific, which can often be overlooked in whole-worm experiments. It is also possible that apart from the basal LARPs-mediated regulation of TENT5, the additional FNDC3A/B interaction is unique to species with more complex polyadenylation machinery. To explore this, members of our group are conducting studies using a mouse model. The results obtained from these and the *C. elegans* experiments are planned to be published soon.

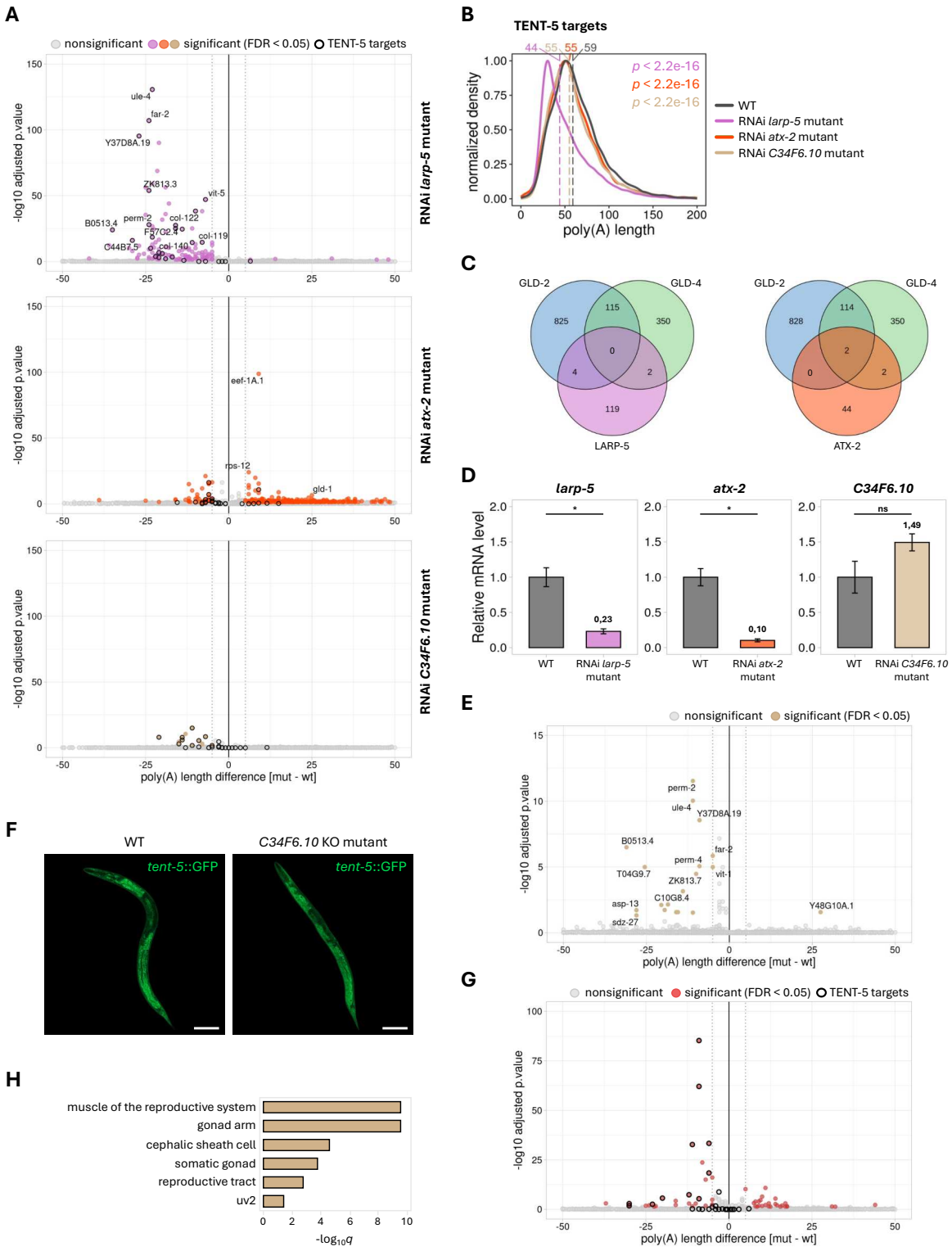


Figure 2. Possible mechanism of TENT-5 target recognition.
(A) Volcano plot displaying differential polyadenylation between wild-type and RNAi *larp-5*, *atx-2*, or *C34F6.10* mutant worms. Transcripts with significantly changed median poly(A) tail length (FDR < 0.05) by a minimum of 5 nucleotides (dotted lines) are marked with purple, orange, and beige dots, respectively. TENT-5 targets are marked with black borderlines.
(B) Density plot showing differences in the poly(A) tail distribution for TENT-5 targets between wild-type (dark gray) and *larp-5* (purple), *atx-2* (orange), and *C34F6.10* (beige) RNAi-treated worms. Vertical dashed lines represent the median poly(A) tail length for each condition (in nucleotides). The plot was generated for all identified TENT-5 targets and normalized to 1.

(C) Venn diagram showing overlaps between GLD-2 (blue) and GLD-4 (green) targets with LARP-5 (left; purple) or ATX-2 (right; orange) targets (defined as mRNAs with significantly shorter poly(A) tail by a minimum of 5 nucleotides).
(D) RT-qPCR showing the efficiency of *larp-5*, *atx-2*, or *C34F6.10* RNAi silencing. Relative *nspc* mRNA levels were normalized to *rps-23*. Bar plots represent mean values with SD. ns => not significant; * => p-value < 0.05 (two-tailed t-test).
(E) Volcano plot displaying differential polyadenylation between wild-type and *C34F6.10* KO mutant worms. Transcripts with significantly changed median poly(A) tail length (FDR < 0.05) by a minimum of 5 nucleotides (dotted lines) are marked with beige dots.
(F) Fluorescence microscopy images of TENT-5-GFP localization in wild-type and *C34F6.10* KO mutant worms. Scale bars => 100 μm.
(G) Volcano plot displaying differential polyadenylation between *C34F6.10* and *C34F6.10/tent-5* mutant worms (TENT-5-driven polyadenylation in the absence of C34F6.10). mRNAs with significantly changed median poly(A) tail length (FDR < 0.05) by a minimum of 5 nucleotides (dotted lines) are marked with red dots. TENT-5 targets in the presence of C34F6.10 are marked with black borderlines.
(H) Top tissue enrichment terms for genes with poly(A) tails significantly shortened in *C34F6.10* KO mutant worms ordered by adjusted *p*-value.

4.4. Function of nematode-specific NSPC proteins – most prominent TENT-5 targets (Manuscript 3)

To better understand the functional significance of the TENT-5 protein in *C. elegans*, it is particularly important to explore the function of its targets. As a poly(A) polymerase, TENT-5 extends the poly(A) tails of various mRNAs, enhancing their stability and expression. As I have already shown, most TENT-5 targets encode small, secreted proteins, which primarily act as immune effectors in hermaphrodites and as components of seminal fluid in males. However, in hermaphrodites, the most prominent group of TENT-5-regulated mRNAs belongs to the NSPC (Nematode-Specific Peptides, group C) family, a group of 18 highly homologous genes that have not been thoroughly characterized[67]. To extend the exploration of TENT-5, we decided to investigate the function of NSPCs.

We began by studying NSPC localization using strains expressing *mCherry*-tagged NSPCs and found that they are exclusively expressed in a single nematode cell – the excretory gland cell. Remarkably, the distinctive poly(A) tail length differences for NSPC transcripts detected in RNA extracted from the entire worm underscore the significance of this regulation and suggest that the excretory gland cell might be a crucial site of TENT-5 activity. However, like NSPC proteins, the role of the excretory gland cell in nematode physiology remains poorly understood[68]. As part of my PhD research, I decided to investigate this unexplored aspect of *C. elegans* physiology and find reasons for such a profound TENT-5 polyadenylation of NSPC-encoding mRNAs.

To achieve this, I optimized a technique for specifically ablating the excretory gland cell. I implemented an optogenetic approach[69,70], expressing the *miniSOG* protein in the cell under the *nspc-10* promoter sequence. Upon blue light illumination, *miniSOG* generates reactive oxygen species, leading to targeted cell destruction. To scale up experiments, I innovatively used an LED advertising board, enabling subsequent high-throughput analyses of ablated worms. Despite the successful ablation, no visible phenotypes were observed in the worms. Therefore, I performed RNA sequencing to uncover potential indications of the excretory gland

cell's function. Initial results suggested its role in nematode reproduction as many sperm- and oocyte-related genes were dysregulated after ablation. However, further investigation revealed that these patterns were artifacts caused by age variation among worms, rather than a true phenotype. Worms' developmental stage is influenced by multiple factors, such as temperature, bacterial concentration, or worms' density, that are often hard to keep constant throughout the experiment[71]. Interestingly, this problem is rarely addressed in the literature, although it significantly impacts the accuracy of data interpretation. For example, my RNA-seq results show that even slight age-related differences overshadowed the effect of entire cell destruction. To mitigate these confounding factors, I performed three replicates of the ablation experiment and averaged the data. Intriguingly, the refined analysis showed that aside from NSPC downregulation, no significant transcriptomic changes occurred following excretory gland cell ablation. This finding suggests that the primary function of this cell is the production of NSPC proteins.

To further investigate the role of NSPCs, I successfully deleted all 18 members of the NSPC family using CRISPR/Cas9. Surprisingly, worms lacking NSPCs also exhibited no visible phenotypes. RNA sequencing analysis revealed that the lack of NSPCs results in a rather weak but statistically significant dysregulation of defense response genes, aligning with the connection of NSPC to the immune regulator TENT-5. Importantly, deleting NSPCs without TENT-5 resulted in minimal additional expression changes, confirming that NSPC-mediated activity is heavily dependent on TENT-5. Additionally, I observed a profound upregulation of PMK family genes in *tent-5* mutants, which diminished when both TENT-5 and NSPCs were deleted. The PMK family is known to regulate defense responses in *C. elegans*[72], suggesting a potential interplay between TENT-5, NSPCs, and PMKs. However, unraveling this complex mechanism lies beyond the scope of this thesis.

Following these RNA-seq results, I studied the survival of *nspc* mutants after exposure to pathogenic *P. aeruginosa* bacteria. Contrary to expectations, NSPC deletion did not alter the worms' lifespan, suggesting that immune defense is not their primary role. Shifting my focus, I noted that the excretory gland cell, the site of NSPC expression, is exceptionally enriched in cholesterol, as are male-specific tissues where TENT-5 activity is prominent. This led me to hypothesize a potential role for TENT-5 and NSPCs in the regulation of cholesterol metabolism. Lifespan experiments on plates with varying cholesterol concentrations, under both pathogenic and non-pathogenic conditions, revealed only minor survival changes. However, the results suggest that TENT-5 may facilitate cholesterol uptake and processing, while NSPCs appear to have an opposing effect. Notably, *tent-5* mutants exhibited improved survival upon *P. aeruginosa* infection at higher cholesterol concentrations, proving the importance of cholesterol for defense responses. This aligns with its known role in regulating the p38/PMK-1 MAPK pathway[73], potentially connecting it to TENT-5 activity.

While NSPCs may influence cholesterol-related processes, this is unlikely to be their primary function. A literature search revealed frequent dysregulation of NSPCs in studies

on the DAF-2/DAF-16 insulin signaling pathway. Furthermore, NSPC structures and sequences resemble those of insulins, which act as direct DAF-2 ligands. Inhibition of DAF-2 triggers nuclear translocation of the transcription factor DAF-16, activating genes involved in lifespan regulation, immunity, and metabolism[74]. I visualized DAF-16 localization in the presence and absence of TENT-5 or NSPCs but found no differences. However, RNA sequencing of *daf-16* and *nspc/daf-16* mutants revealed that NSPCs act upstream of DAF-16, as in their absence DAF-16 deletion results in a significantly smaller transcriptome change. Notably, NSPC activity does not seem to be limited to the insulin pathway regulation, as only part of genes upregulated in *nspc* mutant worms featured the opposite trend when DAF-16 was absent. The rest, among which were defense response genes, were consequently upregulated upon NSPC deletion, with or without DAF-16, which is also consistent with a lack of changes in their survival upon *P. aeruginosa* infection.

Based on these findings, I propose that NSPC proteins function as neuropeptides that modulate various signaling pathways, including the DAF-2/DAF-16 insulin pathway. However, their redundancy with other small peptides likely obscures observable physiological phenotypes.

This study has been published as a pre-print in *bioRxiv* and is enclosed in the appendix to this thesis as **Manuscript 3**.

Reference – Manuscript 3:

Mackiewicz Z., Liudkovska V., Dziembowski A. Knockout of all nematode-specific NSPC genes expressed exclusively in the excretory gland cell results in transcriptomic signatures indicating an affected insulin signaling. *bioRxiv* (2024). DOI: 10.1101/2024.12.06.627198

5. Tools for high-throughput data analysis

The biggest advantage of high-throughput analyses, such as RNA sequencing, is that they provide an enormous amount of information and allow researchers to capture even very complex relationships between genes. However, processing through thousands of rows of data across numerous experimental conditions can often pose a challenge. Throughout my PhD research, I have primarily relied on gene expression changes and differences in poly(A) tail lengths, obtained through Illumina and Nanopore DRS sequencing. In total, I performed sequencing for 35 different conditions, among which I was trying to find some genetic connections. As the dataset grew, data analysis became almost impossible. To address this challenge, I developed **five user-friendly custom data analysis tools** tailored to facilitate the processing, visualization, and interpretation of my data.

As my PhD research focused on studying proteins involved in post-transcriptional gene regulation, it often required investigating transcriptome changes driven by their activity. **The first tool** was designed to facilitate RNA sequencing data analysis, providing robust and clear visualizations of gene expression changes or poly(A) tail length alterations across multiple conditions. Users can focus on individual genes or groups of genes and observe how their expression or poly(A) tails change in different mutants. For instance, by providing the tool with a list of NSPC-regulated transcripts, I discovered that not all of them behaved the same way when DAF-16 was additionally deleted, indicating a potential functional connection as described in Manuscript 3. Similarly, by inserting known TENT-5 targets and visualizing them on Nanopore DRS datasets after silencing potential TENT-5 interactors, I identified LARP-5 as an important regulator of TENT-5 activity.

The second tool builds on the first but focuses on screening RNA sequencing results for changes within specific gene groups. In my research, I mostly touched on aspects of nematode genetics that had never been studied before and frequently encountered genes with no functional annotations, making it challenging to indicate the roles of studied proteins or tissues. Using this tool, I could detect similarities in gene expression trends with existing datasets. For example, by screening through groups of genes enriched in *C. elegans* cells and tissues, I observed that among NSPC-regulated genes, many are enriched in the epidermis and intestine, suggesting the involvement of NSPCs in defense responses triggered in these tissues upon pathogen exposure. In the same way, I could also observe that the poly(A) polymerase GLD-4 preferentially targets germline-enriched genes.

The third application focuses on analyzing a different kind of data. Throughout my PhD, I frequently performed quantitative reverse transcription polymerase chain reactions (RT-qPCR) to validate RNA-seq results or to confirm successful ablation of the excretory gland cell or RNAi silencing. To facilitate the process of data analysis following the RT-qPCR reactions, I designed an application that features a built-in script to analyze RT-qPCR results using the $2^{-\Delta\Delta Ct}$ method. This tool requires only a table with RT-qPCR results as an input file

and automatically performs all the calculations. Later, users can select the reference gene and control condition and explore expression changes for all tested genes. This solution minimizes potential errors during analysis, resulting in more reliable and consistent results.

During my PhD, I frequently investigated various experimental conditions, hypothesizing that the functionality of many of them might be interrelated. As a result, I often needed to analyze the overlap between differentially expressed genes across multiple experiments. To accelerate this process, I developed **the fourth tool** to identify genes overlapping between two gene groups efficiently. The application allows for a robust selection of common genes that can be used for further analyses. For instance, by obtaining the group of mutual GLD-2 and GLD-4 targets and subsequently performing functional enrichment analysis, I could quickly observe their shared role in spermatogenesis.

Lastly, to address the diversity of gene identifiers across datasets, I created **the fifth tool** that simplifies translation between gene names, transcript names, and WormBase IDs. As a result, this application ensures the compatibility of any gene list with various available tools and datasets.

All the tools described were essential during my PhD, allowing me to process multiple datasets more easily and identify important correlations between conditions. Moreover, I believe these applications represent a simple but valuable contribution to the field, providing universal solutions for researchers facing similar challenges. Their development also illustrates the complexity and versatility of tasks encountered during my PhD.

A detailed description of the apps' functionality is provided in the **User manual for data analysis tools** enclosed in this thesis as **Appendix 4**. Tools can also be accessed online and on my GitHub profile with links provided below.

URL: *http://zmackiewicz-rstudio.iimcb.gov.pl:3838/*
GitHub: *https://github.com/zuzanna-mackiewicz/PhD-data-analysis*

6. Summary and future perspectives

The work presented in this PhD thesis focuses on an important aspect of post-transcriptional regulation of gene expression. Early in my PhD, together with other group members, I studied the function of the cytoplasmic poly(A) polymerase TENT-5 in *C. elegans* and highlighted its role in nematode innate immunity. Later, I pursued several projects investigating various aspects of nematode physiology affected by cytoplasmic polyadenylation.

First, I comprehensively analyzed sex-dependent polyadenylation profiles in *C. elegans* and showed that poly(A) tail metabolism varies between sexes. Furthermore, the activity of a single protein can have different functions depending on the animal's sex. Notably, my work sheds light on the previously overlooked role of TENT-5 in regulating transcripts encoding seminal fluid components in *C. elegans* males. It is also the first study to describe sex dimorphism at this level of post-transcriptional gene regulation.

Subsequently, I explored the potential redundancy of TENT-5 with two other poly(A) polymerases – GLD-2 and GLD-4. Through poly(A) profiling using Nanopore DRS, I demonstrated that the roles of these three polymerases do not overlap in *C. elegans*. Moreover, I showed that GLD-4 is responsible for the bulk polyadenylation of mRNAs encoding ribosomal proteins, leading to a global difference in poly(A) tail distributions – an effect not observed for other non-canonical poly(A) polymerases (ncPAPs).

I also investigated the mechanisms underlying TENT-5 target recognition in *C. elegans*. Although previous studies suggested that TENT-5 might be targeted to the endoplasmic reticulum and activated by FNDC3A/B, my data indicate that this process is likely not conserved in nematodes. Instead, my results highlight LARP-5 as the primary regulator of TENT-5-mediated polyadenylation.

Finally, I studied a previously uncharacterized family of nematode-specific NSPC proteins, which are the most prominent targets of TENT-5 and are expressed exclusively in the excretory gland cell. Surprisingly, I found that the function of the excretory gland cell is restricted to the production of NSPCs. I generated a transgenic strain lacking all 18 members of the NSPC family and proposed their involvement in regulating the DAF-2/DAF-16 insulin signaling pathway.

In addition to the biological discoveries, I also designed several R Shiny applications to accelerate large-scale data analysis. Among these, the RNA sequencing analysis tool has proven especially useful, enabling quick comparisons and the detection of complex relationships between multiple experimental conditions. Currently, this application is functional only for the data presented in this thesis. However, I plan to develop it further and release it as an open-source tool, allowing others to use it for a wide range of purposes.

Despite the extensive effort put into my PhD research, many of the initial research questions remain open. For example, how does the massive polyadenylation of seminal fluid components in *C. elegans* males not manifest at the physiological level? Why does the deletion of the entire

cell not have a more dramatic impact on nematode physiology? What is the role and underlying mechanism of NSPC-mediated regulation of the insulin signaling pathway? These are just a few examples of the questions that, unfortunately, could not be addressed within the scope of this thesis.

The absence of obvious phenotypes in *tent-5* mutant males, as well as in worms lacking either the excretory gland cell or its specific NSPC proteins, may reflect their greater relevance under environmental conditions that are more challenging than those in the laboratory. For example, in the natural environment of *C. elegans*, males could face more difficulty in finding mating partners[75]. In such contexts, TENT-5-mediated polyadenylation of seminal fluid component mRNAs might play a critical role in ensuring successful reproduction during these rare opportunities. Similarly, the excretory gland cell's function might be suppressed in laboratory settings. It is plausible that the extensive production of NSPCs by this cell becomes particularly important under conditions where nutrient availability fluctuates, causing significant dysregulation of insulins – typical DAF-2 ligands[76]. Similar studies should be conducted on wild isolates freshly obtained from their natural habitat to test these hypotheses.

Furthermore, future studies could explore the role of the excretory gland cell and NSPCs, as well as the functions of GLD-2 and GLD-4, also in males. These studies could further enhance our understanding of post-transcriptional polyadenylation and its impact on nematode physiology.

In conclusion, the greatest challenge I faced during my PhD project was working on completely unexplored topics in *C. elegans* biology. There are almost no other studies on NSPCs and the excretory gland cell and very few that focus on male physiology and genetics. Most of the genes that appeared in my results have not yet been studied or linked to any biological functions. Therefore, it was extremely difficult to anticipate their potential roles. Nevertheless, despite the demanding nature of this research, it was also incredibly exciting to explore something completely unknown. Overall, I believe the work presented in this PhD thesis provides valuable new perspectives on non-canonical polyadenylation. In particular, the numerous high-throughput datasets generated in this work can lay the foundation for future research in these still poorly explored areas.

7. Methods

Methods for experiments described in chapters 2.1, 4.1, and 4.4 are described in respective publications. Methods for unpublished results are listed below.

C. elegans culture and growth conditions

The following *C. elegans* strains were used: wild type N2 Bristol and RB1181 (*gld-2(ok1117) I*) obtained from the Caenorhabditis Genetics Center (CGC); *tent-5(tm3504) I* obtained from the National Bioresource Project of Japan (NBRP); EV62 (*gld-4(ef9) I*) obtained from Prof. Christian R. Eckmann; PHX9447 (*C34F6.10(syb9447) X*) obtained from SunyBiotech; and ADZ21 (*tent-5(rtt6[tent-5::gfp::3xflag]) I*), previously generated in our laboratory[46]. Strains ADZ85 (*gld-4(ef9) I; tent-5(tm3504) I*), ADZ135 (*C34F6.10(syb9447) X; tent-5(tm3504) I*), and ADZ136 (*C34F6.10(syb9447) X; tent-5(rtt6[tent-5::gfp::3xflag]) I*) were prepared by crossing between respective strains. All strains were cultured at 20°C on nematode growth medium (NGM) plates seeded with *E. coli* HB101 as a food source.

RNA interference (RNAi)

RNA interference (RNAi) was used to silence *larp-5*, *atx-2*, and *C34F6.10* genes in *C. elegans*. NGM plates for RNAi were supplemented with 1 mM IPTG (Sigma-Aldrich), 25 µg/ml carbenicillin (Sigma-Aldrich), and 25 U/ml nystatin (Sigma-Aldrich). *E. coli* HT115 bacteria transformed with RNAi plasmids (The Ahringer *C. elegans* RNAi feeding library[77]) were prepared by overnight culture, centrifuged, concentrated 15x times, and seeded onto RNAi plates, which were left at room temperature overnight. Age-synchronized L4 worms (approximately 30 per 100-mm plate) were transferred to the RNAi plates, and their progeny were used for subsequent analyses. Silencing efficiency was verified using RT-qPCR.

RNA isolation

Populations of age-synchronized L4 worms or F1 progeny after RNAi silencing (approximately in the L4 stage) were collected and washed three times with 50 mM NaCl. Worm pellets were resuspended in 1 ml TRI Reagent (Sigma-Aldrich), vortexed for 15 minutes at room temperature, and stored at -80°C until further processing. RNA extraction was performed according to the manufacturer's protocol for TRIzol-based isolation, with additional purification using KAPA Pure magnetic beads (RNA to beads ratio: 1:3 v/v). All RNA samples isolated for the project described in chapter 4.2. were additionally cap-enriched to match the conditions used previously for studying TENT-5[46]. The cap-enriched mRNA was prepared from 80 µg of total RNA with homemade GST-eIF4E^{K119A} protein and glutathione sepharose 4B (GE Healthcare), as described previously[41]. The rest of the samples were processed as total RNA. RNA quality and integrity were confirmed using the Agilent TapeStation system. RNA

samples were prepared in three independent biological replicates for downstream RT-qPCR, and two replicates were used for Nanopore Direct RNA Sequencing.

RT-qPCR

Expression levels of *larp-5*, *atx-2*, and *C34F6.10* after RNAi silencing were evaluated using RT-qPCR. Total RNA was treated with TURBO DNase (Thermo Fisher Scientific) to eliminate genomic DNA. cDNA synthesis was conducted following the manufacturer's protocol with the use of SuperScript III (Thermo Fisher Scientific) and a mix of oligo(dT)20 and random primers. RT-qPCR reactions were carried out using Platinum SYBR Green Mix (Thermo Fisher Scientific) on a QuantStudio 5 PCR system (Thermo Fisher Scientific). Relative gene expression was normalized to *rps-23* and quantified using the $2^{-\Delta\Delta C(t)}$ method. Statistical significance was determined by two-tailed unpaired Student's t-tests.

Nanopore Direct RNA sequencing and data analysis

Libraries were generated using the Direct RNA Sequencing Kit (SQK-RNA002, Oxford Nanopore Technologies), using 1-4 µg of total or cap-enriched RNA with added in vitro transcribed poly(A) standards (2 ng). Sequencing was conducted on a MinION device and followed by basecalling by Guppy v6.0.0. Reads were aligned to the WBCel235 reference using MiniMap v2.17[78] (options -k 14 -ax map-ont –secondary=no) and processed with samtools v1.9[79]. Poly(A) tail lengths were estimated using Nanopolish v0.13.2 as described before[41,46]. Distribution comparisons across conditions were assessed by the Wilcoxon test, with p-values adjusted via the Benjamini-Hochberg method. Differential expression analysis was performed using DESeq2 Bioconductor package v1.28[80]. Gene Ontology (GO) and tissue enrichment analyses were performed using WormBase Enrichment Suite[81].

TENT-5 localization

The localization of TENT-5 in wild-type and *C34F6.10* mutant worms was visualized in strains ADZ21 (*tent-5(rtt6[tent-5::gfp::3xflag]) I*) and ADZ136 (*C34F6.10(syb9447) X; tent-5(rtt6[tent-5::gfp::3xflag]) I*) using the Zeiss LSM800 confocal microscope with a 40×/oil immersion lens. Synchronized L4 worms were immobilized with 25 µM levamisole on freshly prepared 2% agarose pads and imaged immediately.

8. References

1. MLA style: The Nobel Prize in Physiology or Medicine 2024. NobelPrize.org. Nobel Prize Outreach AB 2024. *https://www.nobelprize.org/prizes/medicine/2024/summary/*.
2. Lee, R. C., Feinbaum, R. L. & Ambrost, V. The C. elegans Heterochronic Gene lin-4 Encodes Small RNAs with Antisense Complementarity to lin-14. *Cell* **75**, 843–854 (1993).
3. Wightman, B., Ha, L. & Ruvkun, G. Posttranscriptional Regulation of the Heterochronic Gene lin-14 by lin-4 Mediates Temporal Pattern Formation in C. elegans. **75**, 855–862 (1993).
4. Brenner, S. The Genetics of Caenorhabditis Elegans. *Genetics* **77**, 71–94 (1974).
5. Corsi, A. K., Wightman, B. & Chalfie, M. A transparent window into biology: A primer on Caenorhabditis elegans. *Genetics* **200**, 387–407 (2015).
6. Lints, R. & Hall, D. H. Male introduction. *WormAtlas* (2009) doi:10.3908/wormatlas.2.1.
7. The C. elegans Sequencing Consortium. Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. *Science* **282**, 2012–2018 (1998).
8. Sulston, J. E. & Horvitz, H. R. Post-embryonic Cell Lineages of the Nematode, Caenorhabditis elegans. *Dev Biol* **56**, 110–156 (1977).
9. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. The Embryonic Cell Lineage of the Nematode Caenorhabditis elegans. *Dev Biol* **100**, 64–119 (1983).
10. Kimble, J. & Hirsh, D. The Postembryonic Cell Lineages of the Hermaphrodite and Male Gonads in Caenorhabditis elegans. *Dev Biol* **70**, 396–417 (1979).
11. Cook, S. J. *et al.* Whole-animal connectomes of both Caenorhabditis elegans sexes. *Nature* **571**, 63–71 (2019).
12. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The Structure Of The Nervous System Of The Nematode Caenorhabditis Elegans. *Phil. Trans. R. Soc. Lond.* **314**, 1–340 (1986).
13. Kaletta, T. & Hengartner, M. O. Finding function in novel targets: C. elegans as a model organism. *Nat Rev Drug Discov* **5**, 387–399 (2006).
14. Fay, D. Genetic mapping and manipulation. *WormBook* (2006) doi:10.1895/wormbook.1.90.1.
15. Nance, J. & Frøkjær-Jensen, C. The caenorhabditis elegans transgenic toolbox. *Genetics* **212**, 959–990 (2019).
16. Ahringer, J. Reverse genetics. *WormBook* (2006) doi:10.1895/wormbook.1.47.1.
17. Conte, D., MacNei, L. T., Walhout, A. J. M. & Mello, C. C. RNA Interference in Caenorhabditis elegans. *Curr Protoc Mol Biol* **2015**, 26.3.1-26.3.30 (2015).
18. Farboud, B., Severson, A. F. & Meyer, B. J. Strategies for efficient genome editing using CRISPR-Cas9. *Genetics* **211**, 431–457 (2019).
19. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green Fluorescent Protein as a Marker for Gene Experssion. *Science* **263**, 802–805 (1994).

20. Proudfoot, N. J., Furger, A., Dye Sir, M. J. & Dunn, W. Integrating mRNA Processing with Transcription. *Cell* **108**, 501–512 (2002).

21. Jalkanen, A. L., Coleman, S. J. & Wilusz, J. Determinants and implications of mRNA poly(A) tail size - Does this protein make my tail look big? *Semin Cell Dev Biol* **34**, 24–32 (2014).

22. Passmore, L. A. & Coller, J. Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. *Nat Rev Mol Cell Biol* **23**, 93–106 (2022).

23. Shi, Y. *et al.* Molecular Architecture of the Human Pre-mRNA 3′ Processing Complex. *Mol Cell* **33**, 365–376 (2009).

24. Yang, Q., Nausch, L. W. M., Martin, G., Keller, W. & Doublié, S. Crystal structure of human poly(A) polymerase gamma reveals a conserved catalytic core for canonical poly(A) polymerases. *J Mol Biol* **426**, 43–50 (2014).

25. Yan, Y. Bin. Deadenylation: Enzymes, regulation, and functional implications. *Wiley Interdiscip Rev RNA* **5**, 421–443 (2014).

26. Zhang, X., Virtanen, A. & Kleiman, F. E. To polyadenylate or to deadenylate: That is the question. *Cell Cycle* **9**, 4437–4449 (2010).

27. Liudkovska, V. & Dziembowski, A. Functions and mechanisms of RNA tailing by metazoan terminal nucleotidyltransferases. *Wiley Interdiscip Rev RNA* **12**, e1622 (2021).

28. Yu, S. & Kim, V. N. A tale of non-canonical tails: gene regulation by post-transcriptional RNA tailing. *Nat Rev Mol Cell Biol* **21**, 542–556 (2020).

29. Martin, G. & Keller, W. RNA-specific ribonucleotidyl transferases. *RNA* **13**, 1834–1849 (2007).

30. WormBase. https://wormbase.org/

31. Krawczyk, P. S. *et al.* SARS-CoV-2 mRNA vaccine is re-adenylated in vivo, enhancing antigen production and immune response. *bioRxiv* (2022) doi:10.1101/2022.12.01.518149.

32. Kadyk, L. C. & Kimble, J. Genetic regulation of entry into meiosis in Caenorhabditis elegans. *Development* **125**, 1803–1813 (1998).

33. Nousch, M., Minasaki, R. & Eckmann, C. R. Polyadenylation is the key aspect of GLD-2 function. *RNA* **23**, 1180–1187 (2017).

34. Wang, L., Eckmann, C. R., Kadyk, L. C., Wickens, M. & Kimble, J. A regulaotry cytoplasmic poly(A) polymerase in Caenorhabditis elegans. *Nature* **419**, 312–316 (2002).

35. Kim, K. W. *et al.* Antagonism between GLD-2 Binding Partners Controls Gamete Sex. *Dev Cell* **16**, 723–733 (2009).

36. Suh, N. *et al.* The GLD-2 poly(A) polymerase activates gld-1 mRNA in the Caenorhabditis elegans germ line. *PNAS* **103**, 15108–15112 (2006).

37. Schmid, M., Küchler, B. & Eckmann, C. R. Two conserved regulatory cytoplasmic poly(A) polymerases, GLD-4 and GLD-2, regulate meiotic progression in C. elegans. *Genes Dev* **23**, 824–836 (2009).

38. Nousch, M., Yeroslaviz, A., Habermann, B. & Eckmann, C. R. The cytoplasmic poly(A) polymerases GLD-2 and GLD-4 promote general gene expression via distinct mechanisms. *Nucleic Acids Res* **42**, 11622–11633 (2014).

39. Kuchta, K. *et al.* FAM46 proteins are novel eukaryotic non-canonical poly(A) polymerases. *Nucleic Acids Res* **44**, 3534–3548 (2016).

40. Mroczek, S. *et al.* The non-canonical poly(A) polymerase FAM46C acts as an onco-suppressor in multiple myeloma. *Nat Commun* **8** (2017).

41. Bilska, A. *et al.* Immunoglobulin expression and the humoral immune response is regulated by the non-canonical poly(A) polymerase TENT5C. *Nat Commun* **11** (2020).

42. Gewartowska, O. *et al.* Cytoplasmic polyadenylation by TENT5A is required for proper bone formation. *Cell Rep* **35** (2021).

43. Brouze, M. *et al.* TENT5-mediated polyadenylation of mRNAs encoding secreted proteins is essential for gametogenesis in mice. *Nat Commun* **15** (2024).

44. Mazur, M. *et al.* Efficient globin production during terminal erythropoiesis depends on the synergistic action of TENT5C poly(A) polymerase and LARP4/5. *bioRxiv* (2024) doi:10.1101/2024.11.14.623596.

45. Doyard, M. *et al.* FAM46A mutations are responsible for autosomal recessive osteogenesis imperfecta. *J Med Genet* **55**, 278–284 (2018).

46. Liudkovska, V. *et al.* TENT5 cytoplasmic noncanonical poly(A) polymerases regulate the innate immune response in animals. *Sci. Adv* **8**, eadd9468 (2022).

47. Zheng, C. *et al.* Non-canonical RNA polyadenylation polymerase FAM46C is essential for fastening sperm head and flagellum in mice. *Biol Reprod* **100**, 1673–1685 (2019).

48. Fucci, C. *et al.* The Interaction of the Tumor Suppressor FAM46C with p62 and FNDC3 Proteins Integrates Protein and Secretory Homeostasis. *Cell Rep* **32** (2020).

49. Brouze, A., Krawczyk, P. S., Dziembowski, A. & Mroczek, S. Measuring the tail: Methods for poly(A) tail profiling. *Wiley Interdiscip Rev RNA* **14** (2023).

50. Sippel, A. E., Stavrianopoulost, J. G., Schutz, G. & Feigelson, P. Translational Properties of Rabbit Globin mRNA after Specific Removal of Poly(A) with Ribonuclease H. *Proc. Nat. Acad. Sci. USA* **71**, 4635–4639 (1974).

51. Sallés, F. J. & Strickland, S. Analysis of Poly(A) Tail Lengths by PCR: The PAT Assay. *Methods in Molecular Biology* **118**, 441–448 (1999).

52. Subtelny, A. O., Eichhorn, S. W., Chen, G. R., Sive, H. & Bartel, D. P. Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature* **508**, 66–71 (2014).

53. Chang, H., Lim, J., Ha, M. & Kim, V. N. TAIL-seq: Genome-wide determination of poly(A) tail length and 3' end modifications. *Mol Cell* **53**, 1044–1052 (2014).

54. Harrison, P. F. *et al.* PAT-seq: A method to study the integration of 3′-UTR dynamics with gene expression in the eukaryotic transcriptome. *RNA* **21**, 1502–1510 (2015).

55. Garalde, D. R. *et al.* Highly parallel direct RN A sequencing on an array of nanopores. *Nat Methods* **15**, 201–206 (2018).

56. Klein, S. L. & Flanagan, K. L. Sex differences in immune responses. *Nat Rev Immunol* **16**, 626–638 (2016).

57. Ebbing, A. *et al.* Spatial Transcriptomics of C. elegans Males and Hermaphrodites Identifies Sex-Specific Differences in Gene Expression Patterns. *Dev Cell* **47**, 801–813 (2018).

58. Kim, B., Suo, B. & Emmons, S. W. Gene Function Prediction Based on Developmental Transcriptomes of the Two Sexes in C. elegans. *Cell Rep* **17**, 917–928 (2016).

59. CeNGEN. https://cengen.shinyapps.io/CengenApp/

60. Cao, W. *et al.* A nucleic acid binding protein map of germline regulation in Caenorhabditis elegans. *Nat Commun* **15** (2024).

61. Arnold, A. *et al.* Functional characterization of C. elegans Y-box-binding proteins reveals tissue-specific functions and a critical role in the formation of polysomes. *Nucleic Acids Res* **42**, 13353–13369 (2014).

62. Flora, Y. & Bohnert, K. A. SPIN-4/Spinster supports sperm activation in C. elegans via sphingosine-1-phosphate transport. *Dev Biol* **504**, 137–148 (2023).

63. Reinke, V., Gil, I. S., Ward, S. & Kazmer, K. Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. *Development* **131**, 311–323 (2004).

64. Kim, K. W. *et al.* Antagonism between GLD-2 Binding Partners Controls Gamete Sex. *Dev Cell* **16**, 723–733 (2009).

65. Mattijssen, S., Kozlov, G., Fonseca, B. D., Gehring, K. & Maraia, R. J. LARP1 and LARP4: up close with PABP for mRNA 3' poly(A) protection and stabilization. *RNA Biol* **18**, 259–274 (2021).

66. Inagaki, H., Hosoda, N., Tsuiji, H. & Hoshino, S. I. Direct evidence that ataxin-2 is a translational activator mediating cytoplasmic polyadenylation. *Journal of Biological Chemistry* **295**, 15810–18525 (2020).

67. Thomas, J. H. Concerted evolution of two novel protein families in Caenorhabditis species. *Genetics* **172**, 2269–2281 (2006).

68. Sundaram, M. V. & Buechner, M. The Caenorhabditis elegans excretory system: A model for tubulogenesis, cell fate specification, and plasticity. *Genetics* **203**, 35–63 (2016).

69. Qi, Y. B., Garren, E. J., Shu, X., Tsien, R. Y. & Jin, Y. Photo-inducible cell ablation in Caenorhabditis elegans using the genetically encoded singlet oxygen generating protein miniSOG. *Proc Natl Acad Sci U S A* **109**, 7499–7504 (2012).

70. Xu, S. & Chisholm, A. D. Highly efficient optogenetic cell ablation in C. Elegans using membrane-targeted miniSOG. *Sci Rep* **6** (2016).

71. Bulteau, R. & Francesconi, M. Real age prediction from the transcriptome with RAPToR. *Nat Methods* **19**, 969–975 (2022).

72. Kim, D. H. *et al.* A Conserved p38 MAP Kinase Pathway in Caenorhabditis elegans Innate Immunity. *Science* **297**, 623–626 (2002).

73. Otarigho, B. & Aballay, A. Cholesterol Regulates Innate Immunity via Nuclear Hormone Receptor NHR-8. *iScience* **23** (2020).

74. Murphy, C. T. & Patrick J Hu. Insulin/insulin-like growth factor signaling in C. elegans. *WormBook* 1–43 (2013) doi:10.1895/wormbook.1.164.1.

75. Anderson, J. L., Morran, L. T. & Phillips, P. C. Outcrossing and the maintenance of males within C. elegans populations. *Journal of Heredity* **101**, S62–S74 (2010).

76. Zhu, R. & Chin-Sang, I. D. C. elegans insulin-like peptides. *Mol Cell Endocrinol* **585** (2024).

77. Kamath, R. S. & Ahringer, J. Genome-wide RNAi screening in Caenorhabditis elegans. *Methods* **30**, 313–321 (2003).

78. Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).

79. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, 1–4 (2021).

80. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15** (2014).

81. Angeles-Albores, D., Raymond, R. Y., Chan, J. & Sternberg, P. W. Tissue enrichment analysis for C. elegans genomics. *BMC Bioinformatics* **17** (2016).

Appendix 1

Manuscript 1

Liudkovska V., Krawczyk P.S., Mroczek S., Bilska A., Gumińska N., Cysewski D., **Mackiewicz Z.**, Ewbank J.J., Drabikowski K., Dziembowski A. TENT5 cytoplasmic non-canonical poly(A) polymerases regulate the innate immune response in animals. *Sci. Adv.* **8**, eadd9468 (2022). DOI:10.1126/sciadv.add9468

Warsaw, 7.01.2025

Zuzanna Mackiewicz
Laboratory of RNA Biology
International Institute of Molecular and Cell Biology
4 Ks. Trojdena
02-109 Warsaw

PHD CANDIDATE'S
CONTRIBUTION STATEMENT

Article's title: TENT5 cytoplasmic noncanonical poly(A) polymerases regulate the innate immune response in animals.

Authors: Vladyslava Liudkovska, Paweł S. Krawczyk, Aleksandra Brouze, Natalia Gumińska, Tomasz Wegierski, Dominik Cysewski, Zuzanna Mackiewicz, Jonathan J. Ewbank, Krzysztof Drabikowski, Seweryn Mroczek, Andrzej Dziembowski

Journal: Science Advances

Date of publishing: 16th November 2022

DOI: 10.1126/sciadv.add9468

I hereby declare that my contribution to the article titled *"TENT5 cytoplasmic noncanonical poly(A) polymerases regulate the innate immune response in animals"* was as following:

- performing lifespan analyses of *tent-5* mutants grown on UV- or heat-killed bacteria
- preparing worms for colony-forming unit (CFU) analyses
- preparing *C. elegans* strains for studying ER-localization of TENT-5

.....
PhD candidate signature

.....
Corresponding author signature

MOLECULAR BIOLOGY

TENT5 cytoplasmic noncanonical poly(A) polymerases regulate the innate immune response in animals

Vladyslava Liudkovska[1,2]†, Paweł S. Krawczyk[1,3]‡, Aleksandra Brouze[1,2]‡, Natalia Gumińska[1], Tomasz Wegierski[1], Dominik Cysewski[3], Zuzanna Mackiewicz[1], Jonathan J. Ewbank[4]§, Krzysztof Drabikowski[3], Seweryn Mroczek[1,2], Andrzej Dziembowski[1,2,3]*

Innate immunity is the first line of host defense against pathogens. Here, through global transcriptome and proteome analyses, we uncover that newly described cytoplasmic poly(A) polymerase TENT-5 (terminal nucleotidyltransferase 5) enhances the expression of secreted innate immunity effector proteins in *Caenorhabditis elegans*. Direct RNA sequencing revealed that multiple mRNAs with signal peptide–encoding sequences have shorter poly(A) tails in *tent-5*–deficient worms. Those mRNAs are translated at the endoplasmic reticulum where a fraction of TENT-5 is present, implying that they represent its direct substrates. Loss of *tent-5* makes worms more susceptible to bacterial infection. Notably, the role of TENT-5 in innate immunity is evolutionarily conserved. Its orthologs, TENT5A and TENT5C, are expressed in macrophages and induced during their activation. Analysis of macrophages devoid of TENT5A/C revealed their role in the regulation of secreted proteins involved in defense response. In summary, our study reveals cytoplasmic polyadenylation to be a previously unknown component of the posttranscriptional regulation of innate immunity in animals.

INTRODUCTION

Innate immunity is an evolutionarily ancient mechanism that provides general host protection against pathogens (*1*). In mammals, innate immunity functions alongside adaptive immunity and also plays a key role in its activation (*2*, *3*). On the other hand, many organisms, including the bacterivore nematode *Caenorhabditis elegans*, lack adaptive immunity and defend themselves solely with innate immune mechanisms (*4*, *5*). Some core components of innate immunity are conserved, but the particular means of protection used by different animals vary. In mammals, the innate system depends on physical and anatomical barriers (e.g., the barrier epithelial cells, mucus, tears, earwax, and stomach acid), the humoral component (i.e., cytokines, chemokines, and defensins secreted by innate immune cells), and several types of phagocytic cells, among them are macrophages, which recognize and destroy pathogens (*6*). *C. elegans* has no specialized immune cells and relies on its barrier tissues, epidermal and intestinal cells, for defense (*4*). An efficient immune response is provided by the high secretion capacity of the particular host cells, as they release a variety of antimicrobial peptides and enzymes that can directly attack pathogens (*4*, *6*). In all organisms studied, the innate immune response is regulated at both the transcriptional and posttranscriptional levels (*4*, *7*). The signaling pathways and transcriptional factors that control innate immunity in worms have been studied in detail, but less is known about the posttranscriptional mechanisms involved in this process.

In eukaryotes, most mRNAs are polyadenylated by the canonical polyadenylate [poly(A)] polymerase during mRNA 3′ end processing in the nucleus (*8*, *9*). The poly(A) tail is essential for mRNA stability, export to the cytoplasm, translation, and turnover (*10*). Poly(A) tails are gradually shortened in the cytoplasm by deadenylases, and their reduction to less than 20 nucleotides (nt) leads to mRNA degradation (*11*). However, in some cases, a poly(A) tail can be extended in the cytoplasm by noncanonical poly(A) polymerases (ncPAPs). These enzymes belong to the family of terminal nucleotidyltransferases (TENTs) and are implicated in a range of physiological processes (*12*–*14*). Cytoplasmic polyadenylation has been mostly studied in the context of the activation of dormant deadenylated mRNAs during gametogenesis (*15*–*17*) and in neuronal processes (*18*, *19*), in which GLD-2 (Germ Line Development 2)/TENT2 polyadenylates certain mRNAs in response to cellular signals. The recent discovery of the TENT5 family of cytoplasmic ncPAPs and the characterization of their functions expanded the repertoire of physiological processes that are affected by this type of posttranscriptional regulation (*20*–*24*). Mammalian genomes encode four TENT5 proteins (TENT5A to TENT5D, also known as FAM46A to FAM46D), all of which are active ncPAPs (*21*, *25*). Functional analysis revealed that TENT5C is a multiple myeloma growth suppressor (*21*, *22*). In multiple myeloma cell lines, TENT5C polyadenylates and stabilizes numerous mRNAs that encode secreted proteins (*21*). TENT5C also plays a crucial role in the regulation of immunoglobulin expression and the humoral immune response in mice through polyadenylation of mRNAs that encode immunoglobulins (*23*, *24*). TENT5A polyadenylates mRNAs encoding collagens and is thus required for the proper bone formation (*26*). However, TENT5 proteins are differentially expressed in mammalian tissues and organs with potential redundancy, which makes study of their functions difficult.

Here, we characterized the only TENT5 family member in *C. elegans*, F55A12.9/PQN-44, which we renamed TENT-5. Transcriptomic and proteomic analysis, along with functional studies, revealed that TENT-5 is an innate immune response regulator. Poly(A) tail profiling by direct RNA sequencing (RNA-seq) showed that TENT-5

[1]Laboratory of RNA Biology, International Institute of Molecular and Cell Biology, Trojdena 4, 02-109 Warsaw, Poland. [2]Faculty of Biology, University of Warsaw, Pawinskiego 5a, 02-106 Warsaw, Poland. [3]Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawinskiego 5a, 02-106 Warsaw, Poland. [4]Aix Marseille University, INSERM, CNRS, CIML, Turing Centre for Living Systems, Marseille, France.
*Corresponding author. Email: adziembowski@iimcb.gov.pl
†Present address: Laboratory of Stem Cell RNA Metabolism, The International Institute of Molecular Mechanisms and Machines, Polish Academy of Sciences, Smetany 2, 00-783 Warsaw, Poland.
‡These authors contributed equally to this work.
§Present address: European Research Infrastructure on Highly Pathogenic Agents (ERINHA), 98, rue du Trône, Boite 4, Brussels 1050, Belgium.

polyadenylates and stabilizes mRNAs that encode defense proteins. *tent-5* deficiency led to an impaired innate immune response in worms. The role of TENT5 proteins in innate immunity is evolutionarily conserved because murine macrophages devoid of TENT5A and TENT5C also exhibited defects in polyadenylation of mRNAs that encode proteins with a role in innate immunity. Together, we identified *C. elegans* TENT-5 and its mammalian orthologs TENT5A and TENT5C as previously unknown players that regulate innate host defenses.

RESULTS

TENT-5 is a noncanonical cytoplasmic poly(A) polymerase in worms

In *C. elegans*, there is only one homolog of mammalian TENT5 ncPAPs, PQN-44 (WBGene00004131). Multiple sequence alignment indicated that PQN-44 exhibits high conservation with human TENT5 proteins (fig. S1A) and clusters with the vertebrate TENT5 homologs in phylogenetic analysis (Fig. 1A). Similar to other TENTs, PQN-44 comprises nucleotidyltransferase and poly(A) polymerase–associated domains (fig. S1A). The nucleotidyltransferase domain of TENT-5 contains a well-conserved triad of aspartates/glutamates ([DE]h[DE]h and h[DE]h, with h being a hydrophobic residue) in its catalytic center (fig. S1A) (*12*). These observations suggested that PQN-44 could be a previously unidentified ncPAP in worms; hence, we renamed it TENT-5. To determine whether TENT-5 is an active ncPAP with the potential to polyadenylate substrate mRNAs and enhance their expression, we performed an RNA-tethering assay (*27*), in which TENT-5 is brought to a *Renilla* luciferase (*RL*) reporter mRNA through an artificial interaction (Fig. 1B). 293T cells were cotransfected with the pRL-5Box plasmid, which carries an *RL* containing five boxB sites in its 3′ untranslated region (3′UTR), and construct expressing either wild-type (TENT-5^{WT}) or catalytically inactive [TENT-$5^{D151A, D153A}$ (isoform a)] protein harboring the N-terminal λN boxB-binding domain that ensures interaction with boxB sites. Next, we performed direct full-length RNA-seq (DRS) to measure poly(A) tail length at a genome-wide scale. DRS revealed that tethering of the wild-type TENT-5 led to increased polyadenylation of *RL* mRNA [median poly(A) length of 180 nt], compared to the catalytically inactive TENT-5 [median poly(A) length of 106 nt] (Fig. 1C and fig. S1, B and C). At the same time, the global distribution of transcripts' poly(A) tail lengths was not affected by overexpression of TENT-$5^{WT/MUT}$ (Fig. 1D). Furthermore, the extension of the *RL* poly(A) tail upon tethering of the TENT-5^{WT} resulted in the enhanced production of the reporter protein (fig. S1D), indicating that TENT-5–mediated polyadenylation positively regulates gene expression.

To determine the pattern of TENT-5 expression and its cellular localization, we generated a *tent-5::gfp* knockin strain by CRISPR-Cas9 (fig. S1, E, F, and H). Confocal microscopy revealed that *tent-5* was expressed during all developmental stages, including in embryos and adults (Fig. 1E). Expression was observed in most tissues, including the pharynx, head and tail neurons, seam and hypodermal cells, and intestine (Fig. 1, E and F, and fig. S1G). Notably, TENT-5–green fluorescent protein (GFP) was predominantly localized in the cytoplasm in all cells (Fig. 1, E and F, and fig. S1G). Together, on the basis of the high sequence similarity of TENT-5 with its mammalian counterparts, its polyadenylation activity, and its cytoplasmic localization, we propose that TENT-5 is a novel cytoplasmic ncPAP in worms.

To characterize the role of TENT-5, we used a mutant *tent-5(tm3504)* strain that harbors a deletion that introduces a premature stop codon in *tent-5* (fig. S1H) and produces a reduced level of mRNA predicted to encode a nonfunctional protein (fig. S1I). Using CRISPR-Cas9, we also generated a *tent-5(rtt5)* mutant lacking the entire coding sequence (fig. S1, H and I). Both *tm3504* and *rtt5* are expected to be null alleles. Homozygous mutants carrying these alleles did not, however, display any visible developmental or morphological abnormalities when grown under standard conditions, as revealed by routine observation, measurement of brood size (fig. S1J), and worm body parameters and locomotion analyses (fig. S1, K to N). This observation is in agreement with results from mammals, as *Tent5c* knockout (KO) mice do not display any gross phenotypes (*21*, *23*). We conclude that *tent-5* encodes a poly(A) polymerase that is not essential for worm development under standard conditions.

Loss of *tent-5* leads to the down-regulation of genes that encode innate immune effectors

To gain insight into TENT-5's molecular function, we analyzed the whole transcriptome of age-synchronized L4 worms devoid of *tent-5*. Using RNA-seq, we identified 745 genes differentially expressed in *tent-5(tm3504)* mutant compared to wild-type worms [false discovery rate (FDR) < 0.05] (Fig. 2A and data S1). The expression level of 308 genes was down-regulated and that of 96 genes was up-regulated at least 1.5-fold in mutant animals. The higher number of down-regulated than up-regulated genes in *tent-5(tm3504)* mutant worms is in agreement with our initial hypothesis that polyadenylation by TENT-5 may stabilize mRNAs, and potential direct targets of TENT-5 could be found among the genes which expression is lower in the mutant animals. Clustered Gene Ontology (GO) analysis revealed the association of the down-regulated genes with immune and defense responses, response to biotic stimulus, peptidase and hydrolase activities, lytic vacuole, and extracellular space (Fig. 2B). A substantial number of down-regulated genes encoded various antimicrobial and cytoprotective immune effectors that are constitutively expressed and additionally induced during infection (Fig. 2C). Notably, many of these genes belong to genomic clusters, such as *F55G11.2*, *F55G11.4*, *F55G11.8*, *K02E11.4*, *K02E11.5*, and *K02E11.6*, or to gene families, often linked to innate immunity, including infection response genes (*irg-3*, *irg-5*, and *irg-7*) (*28*), genes that encode caenopores (SPP; 10 genes) (*29*), proteins containing C-type lectin-like domains (CLEC; 10 genes) (*30*), CUB domain containing proteins (7 genes) (*28*), proteins containing the transthyretin-related domain (12 genes) (*31*), collagens (COL; 14 genes) (*31*, *32*), members of the Nematode Specific Peptide family, group C (NSPCs; 10 genes) (*33*), aspartyl proteases (ASP; 10 genes) (*34*), and lysozymes (*lys-7*, *lys-8*, *ilys-5*, and *lys-2*) (*35*). Tissue enrichment analysis (TEA) of down-regulated genes indicated, among other terms, their association with the intestine (Fig. 2D and fig. S2A). In worms, the intestine has a prominent role in innate immunity, as intestinal cells secrete antimicrobial peptides and digestive enzymes into the gut lumen (*4*, *35*, *36*).

Although GO term analysis of the up-regulated genes also revealed their connection with the immune response and intestine, only a few of them encoded the immune effector proteins (*gst-13*, *asp-10*, *clec-187*, *clec-265*, *clec-42*, *dod-22*, *fipr-22*, *lys-4*, *ftn-2*, *hrg-3*, *math-38*, *K08D8.4*, *ZK228.4*, *C37A5.3*, and *F31F7.1*) (data S1). Among the genes up-regulated in mutant worms were also three genes belonging to a single operon (CEOP4272) and encoding mitogen-activated protein kinase homologs essential for innate immune response—*pmk-1*, *pmk-2*, and *pmk-3* (*4*). This observation is important as it may suggest the compensatory effect of signaling pathways in *tent-5(tm3504)*

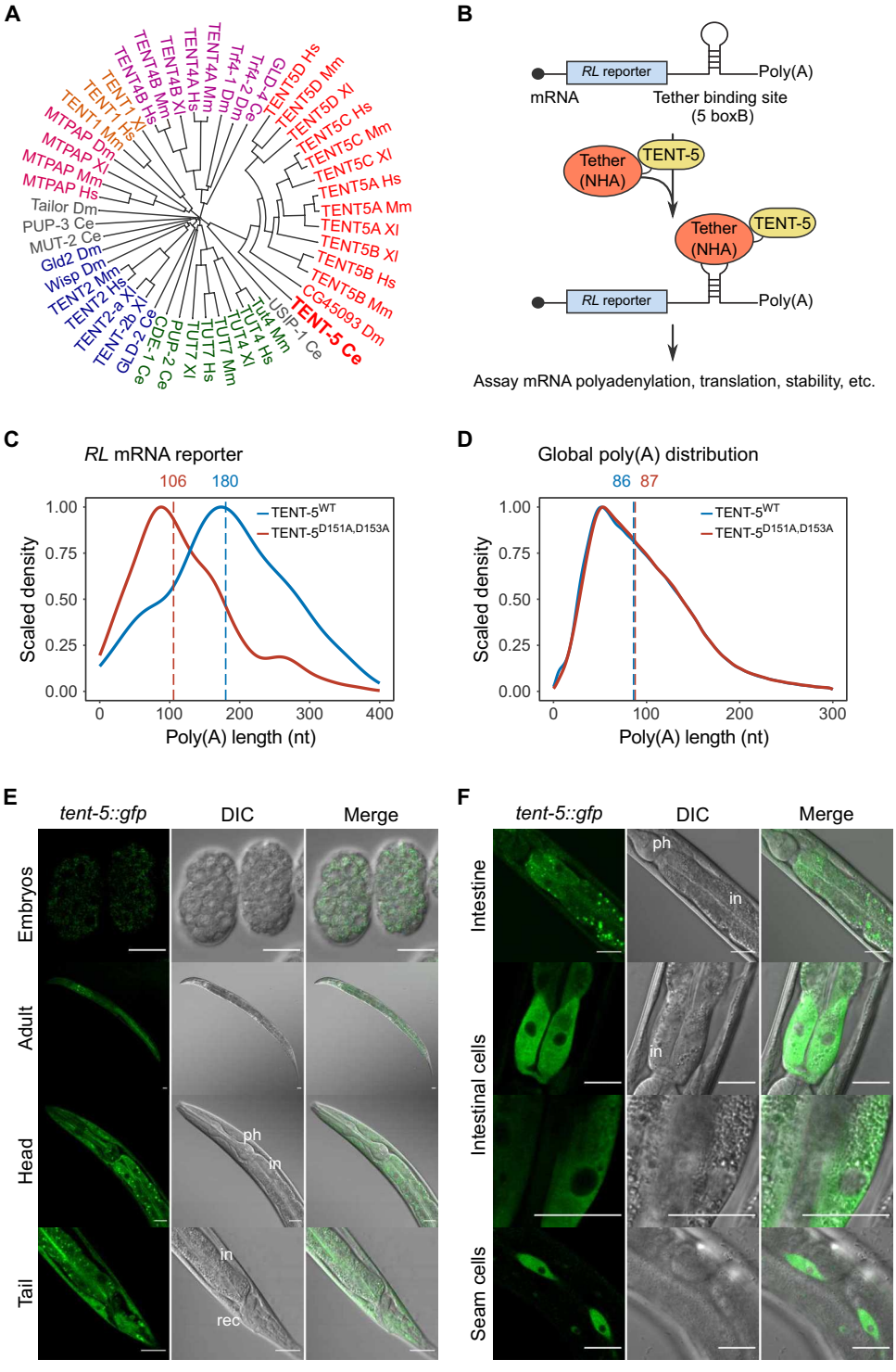


Fig. 1. TENT-5 is a noncanonical cytoplasmic poly(A) polymerase in worms. (**A**) Phylogenetic relationship among 48 TENTs from five model organisms (Hs, *Homo sapiens*; Mm, *Mus musculus*; Xl, *Xenopus laevis*; Dm, *Drosophila melanogaster*; Ce, *C. elegans*). (**B**) Schematic illustration of RNA tethering assay, modified from (*27*). A protein of interest (TENT-5$^{WT/D151A,D153A}$) binds to a *RL* reporter mRNA through an artificial protein-RNA interaction (tether). For tethering assays described in this work, 293T cells were cotransfected with a construct that expresses *RL*, containing a tether binding site (five boxB sites) in its 3′UTR, and a construct expressing either wild-type or catalytically inactive TENT-5 harboring the N-terminal λN boxB-binding domain and an HA-tag (NHA). (**C** and **D**) DRS-based poly(A) length profiling of mRNA following tethering assay. Tethering of wild-type TENT-5 led to increased polyadenylation of an *RL* reporter mRNA compared to a tethering with a catalytically inactive TENT-5^{D151A,D153A} mutant. Shown are density distribution plots for *RL* mRNA (C) and all other transcripts detected in 293T cells (D) scaled to 1. Vertical dashed lines represent median poly(A) lengths (in nucleotides). (**E** and **F**) Representative fluorescence and differential interference contrast (DIC) microscopy images of TENT-5–GFP expression in embryos and adult *tent-5::gfp* knockin worms (*tent-5(rtt6[tent-5::gfp::3xflag] I)*. ph, pharynx; in, intestine; rec, rectum. Scale bars, 20 μm.

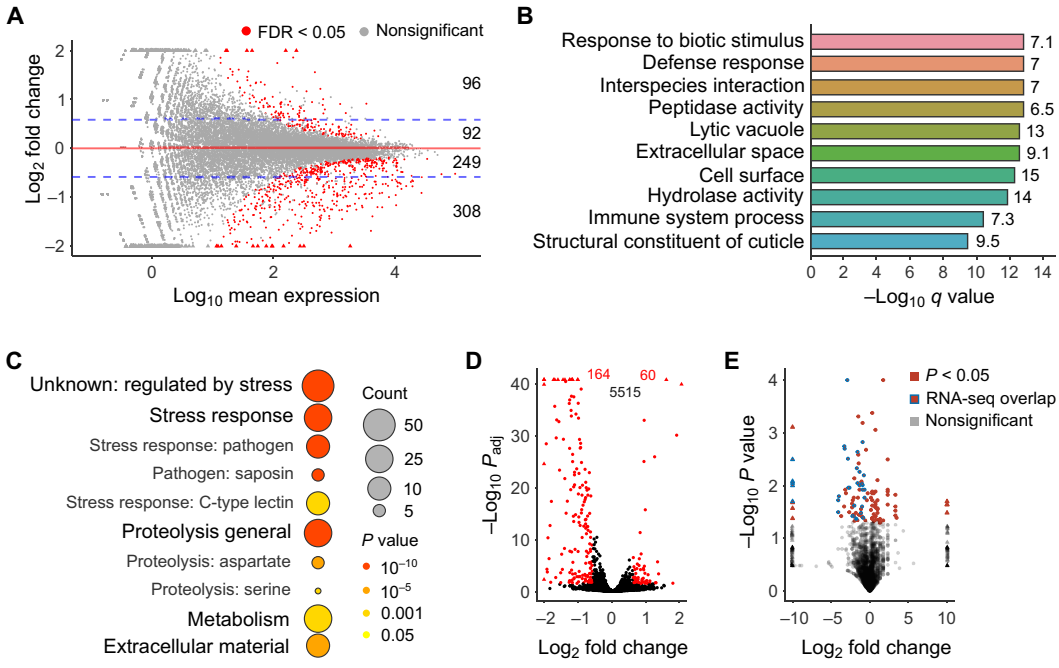


Fig. 2. Loss of *tent-5* leads to the down-regulation of genes that encode innate immune effectors. (**A**) MA plot illustrates differential gene expression in *tent-5(tm3504)* mutant compared to the wild-type worms. Statistically significant values (FDR < 0.05) are shown in red. Blue dashed lines mark the threshold of the $\log_2(1.5)$ fold change. Numbers on the right indicate the number of genes differentially expressed for each fold change group. Triangles, here and (D) and (E), represent data points outside the axes limits. (**B**) Overrepresented functional GO terms of genes down-regulated [$\log_2$FC ($\log_2$ fold change) < $-\log_2(1.5)$] in *tent-5(tm3504)* mutant in comparison to the wild type. Numbers on the right indicate the enrichment fold change for each term. (**C**) WormCat visualization of categories enriched in genes down-regulated 1.5-fold or more in *tent-5(tm3504)* worms. Results are presented as scaled heatmap bubble charts, where the color signifies *P* values and the size specifies the number of genes in each category. (**D**) In the RNA-seq dataset, 5739 genes are associated with the "intestine" GO term cellular compartment (WBbt:0005772). From genes down-regulated or up-regulated in mutant worms at least 1.5-fold (A), 164 and 60 genes are associated with this GO term, respectively. Statistically significant values (FDR < 0.05) are shown in red and nonsignificant in black ($n = 5515$). (**E**) Volcano plot showing proteins, which abundance was significantly changed in mutant (red shapes, $n = 114$). Red shapes with blue borders represent a common part with RNA-seq results ($n = 33$).

mutant to counteract the lower expression level of numerous immune effectors or may point out some nondirect processes connected to the *tent-5* loss. There are no up-regulated genes statistically significantly enriched for the "extracellular region" GO term.

To check whether deregulation of mRNA levels in mutant worms correlated with changes in protein abundance, we performed a large-scale semiquantitative proteomic analysis comparing extracts from L4-stage wild-type and *tent-5(tm3504)* mutant animals. Among proteins with significantly changed abundance ($P < 0.05$), we identified 67 proteins with decreased levels in *tent-5*–deficient worms (Fig. 2E and data S2). Overall, the results of the proteomic analysis were in agreement with the transcriptomic data (Fig. 2E). Notably, some proteins from the defense-associated families mentioned above were less abundant in mutant worms (e.g., three CLEC, three COL, two SPP, two ASP, and two LYS; data S2), and this was mirrored in a more general enrichment for proteins with described or putative roles in innate immunity, as well as protein catabolic processes (fig. S2B). Furthermore, proteins with a decreased level in mutant worms were associated with the intestine (fig. S2C), again in agreement with the transcriptomic analysis.

Consistent with the role of TENT-5 ncPAP in the immune response, the expression of *tent-5* was induced approximately twofold upon the infection with *Staphylococcus aureus*, as revealed by reverse transcription quantitative polymerase chain reaction (RT-qPCR) analysis (Fig. 3A and fig. S3A), in line with previously published datasets (*37*, *38*). To explore further the function of TENT-5 in the immune response, we checked whether immune- and stress-related genes were deregulated in *tent-5(tm3504)* mutants in comparison to wild-type worms after 8 hours of infection with *S. aureus*. RNA-seq results showed that 296 genes were down-regulated 1.5-fold or more (FDR < 0.05) (Fig. 3B and data S3). More than half of all genes for which expression was down-regulated in *tent-5(tm3504)* mutant worms grown on *Escherichia coli* HB101 were also down-regulated in wild-type worms upon infection (Fig. 3C). In agreement with this, results of the GO analysis and TEA were comparable between both datasets and showed that most down-regulated genes were potentially involved in innate immunity (fig. S3, B and C). TENT-5 was necessary for the proper expression of immune genes in both infected and noninfected animals (fig. S3D). Furthermore, the genes down-regulated in *tent-5*–deficient worms are known to be regulated by a variety of pathways (*28*, *36*, *37*, *39*–*43*), indicating that TENT-5's action is not limited to a single signal transduction pathway. Note that some of the 62 genes whose expression was up-regulated in *tent-5*–deficient worms at least 1.5-fold were associated with a defense response (*pmk-1*, *pmk-2*, *pmk-3*, *lys-10*, *ilys-2*, *zip-10*, *clec-60*, *clec-86*, *fbxa-60*, *fbxa-105*, *C10C5.2*, *K08D8.4*, *M04C3.2*, and *C49G7.10*) (data S3). However, their overlap with genes up-regulated in mutant worms grown on *E. coli* HB101 was quite limited (14%). Again, this might suggest that up-regulation of gene expression may be a mainly secondary effect.

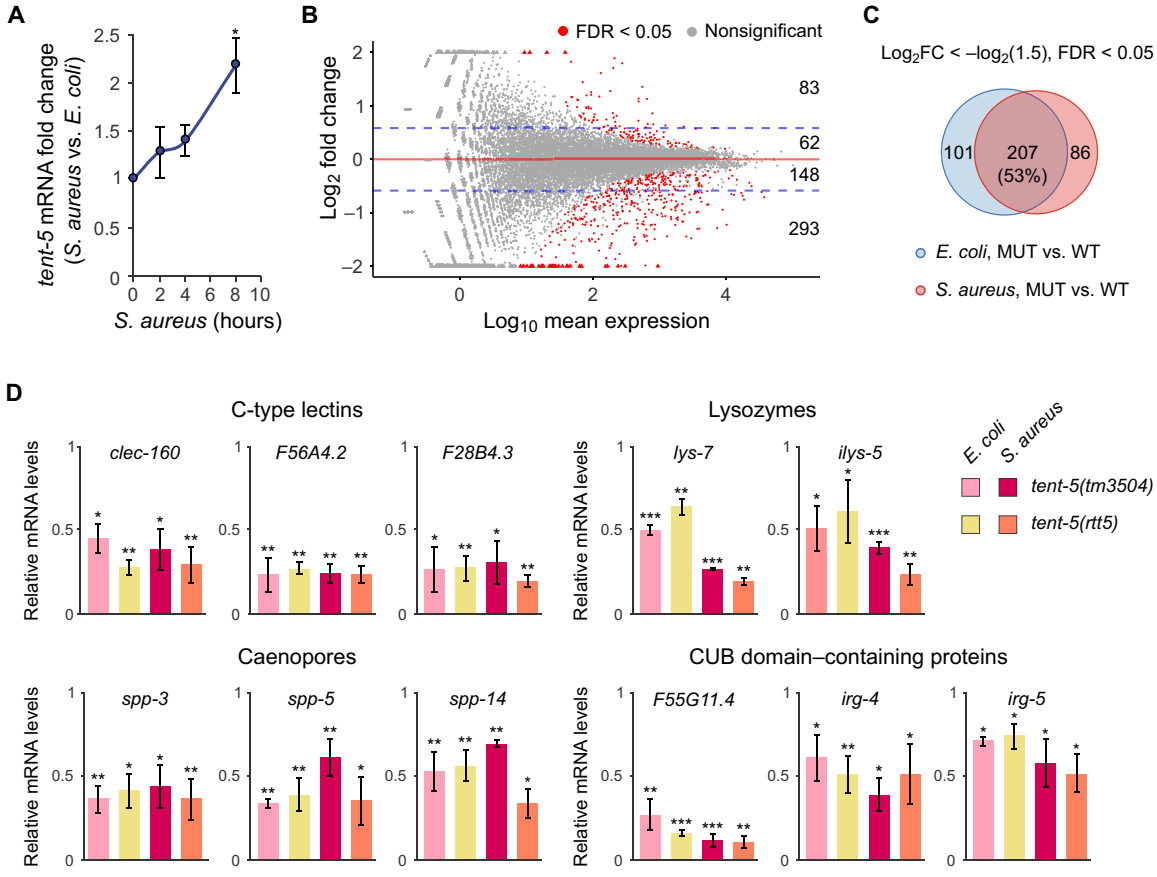


Fig. 3. TENT-5 deficiency leads to the down-regulation of genes that encode innate immune effectors. (**A**) RT-qPCR illustrating the induction of *tent-5* expression in the wild-type worms challenged with *S. aureus* at the time point. Relative *tent-5* mRNA level was normalized to *act-1* and uninfected animals. Data represent means ± SD of three biological replicates; $*P \leq 0.05$ (one-tailed *t* test). (**B**) MA plot showing differential expression of genes in *tent-5(tm3504)* mutant compared to wild type upon infection of L4 worms with *S. aureus* for 8 hours. Plot description as in Fig. 2A. (**C**) Venn diagrams demonstrating the overlap between genes differentially expressed in tent-5–deficient worms grown on nonpathogenic *E. coli* and upon infection by *S. aureus* (FDR < 0.05). MUT, mutant; WT, wild type. (**D**) RT-qPCR analysis of relative levels of mRNA expression in *tent-5(tm3504)* and CRISPR-Cas9–generated *tent-5(rtt5)* mutant grown on *E. coli* or exposed to *S. aureus* for 8 hours. Relative abundance of mRNAs was normalized to *act-1* and wild type. Data represent means ± SD of three biological replicates; $*P \leq 0.05$, $**P \leq 0.01$, and $***P \leq 0.001$ (two-tailed *t* test).

To validate the observed molecular phenotypes of the *tent-5(tm3504)* strain, we grouped down-regulated genes into several functional categories or families and verified their expression in infected and noninfected *tent-5(rtt5)* mutant animals using the RT-qPCR analysis. The results for *tent-5(rtt5)* recapitulated those of *tent-5(tm3504)* strain, with the null mutant exhibiting a down-regulation of expression for all transcripts tested (Fig. 3D and fig. S3E). Thus, we conclude that the observed gene expression changes were due to TENT-5 deficiency. Together, the results of the global transcriptomic and proteomic analyses indicate that TENT-5 is involved in the innate immune response by controlling the expression of immune genes regulated by several signaling pathways.

TENT-5 is required for the host defense against various bacterial strains

The results of the global transcriptomic analysis strongly suggested that TENT-5 could be directly involved in the regulation of genes influencing the interaction between *C. elegans* and bacteria, including many encoding secreted proteins with antimicrobial activity. When *tent-5* mutant worms were grown on *E. coli* HB101, they exhibited a slightly shorter survival than the wild type when grown at both 20°C (Fig. 4A) and 25°C (Fig. 4B). *E. coli* is not a natural food source for *C. elegans* and, under standard culture conditions, is mildly pathogenic to adult worms (*35*, *44*, *45*). To check the possibility that mutants have a shorter life span due to the pathogenesis of *E. coli*, we tested the survival of worms fed on either ultraviolet (UV)– or heat-killed *E. coli* in comparison to the intact bacteria. *tent-5*–deficient mutants exhibited survival curves more comparable to the wild-type worms when cultured on dead bacteria (fig. S4, A to D), indicating that the survival defect of mutant worms could result from their compromised immunity. We then monitored the survival of wild-type and *tent-5*–deficient worms infected with *S. aureus*, *Pseudomonas aeruginosa*, *Serratia marcescens*, and *Photorhabdus luminescens*. These bacterial pathogens colonize the worm's intestinal lumen, causing intestinal distention and worms' death within 3 to 10 days (*46*–*48*). The *tent-5*–deficient mutants were significantly more susceptible to the infection than the wild type (Fig. 4, D and E, and fig. S4, E to G). To investigate whether there are differences in the accumulation of *E. coli* HB101 and *P. aeruginosa* PAO1 in the intestines of wild-type and mutant worms, we performed colony-forming unit (CFU) assays (*49*, *50*). We observed that *tent-5*–deficient worms had higher bacterial load in their guts compared to

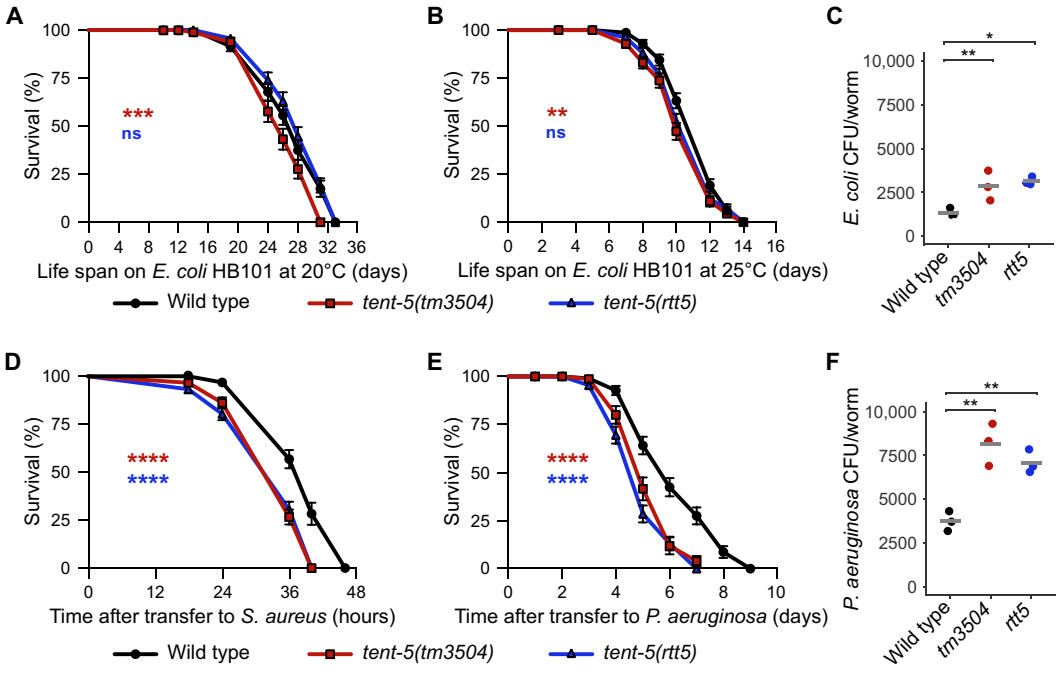


Fig. 4. TENT-5 is required for the host defense against various bacterial strains. Survival of wild-type, *tent-5(tm3504)*, and *tent-5(rtt5)* mutant worms grown on non-pathogenic *E. coli* HB101 at 20°C (**A**) and 25°C (**B**) or infected with *S. aureus* (**D**) and *P. aeruginosa* PAO1 (**E**). For survival on *P. aeruginosa* and *E. coli*, nematode growth medium (NGM) plates contained FUdR. Experiments are representative of three independent trials; $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$; ns, not significant (log-rank test). Survival data can be found in table S1. Intestinal accumulation of *E. coli* HB101 (**C**) and *P. aeruginosa* (**F**) was measured on the 5th and 4th days of adulthood, respectively. The number of viable internal bacteria in worms' gut is shown in CFU per animal. Points show values for experimental biological replicates calculated as a mean of at least two technical replicas. Line denotes the average value for biological replicates. $^{*}P \leq 0.05$ and $^{**}P \leq 0.01$ (two-tailed *t* test).

the wild type (Fig. 4, C and F). The observed effects indicate that TENT-5 is important for the defense against a range of Gram-positive and Gram-negative pathogens, which use different infection strategies and induce different but overlapping host gene expression responses (*31*, *34*, *46*, *51*). Notably, the localization of TENT-5–GFP in the cytoplasm of the intestinal cells (Fig. 1F) is in agreement with its role in the regulation of innate immune genes, as many immune genes down-regulated in mutant worms are expressed in this tissue (fig. S2, A and C, and data S1 and S2). Together with the observation that *tent-5* mutants are less resistant to infection with several pathogenic bacteria, our results support the hypothesis that TENT-5 plays an important role during host defense and also suggest that TENT-5 may contribute to the regulation of longevity or general fitness in *C. elegans*.

Transcripts down-regulated in the *tent-5*–deficient mutants are direct targets for polyadenylation by TENT-5

To elucidate the mechanism by which TENT-5 regulates innate immunity in worms, we first explored the possibility that TENT-5 might regulate one or more transcription factors responsible for the expression of immune genes (*4*). However, we did not identify any transcription factors among the genes down-regulated in *tent-5(tm3504)* animals or transcriptional repressors among the up-regulated ones (data S1). As a cytoplasmic poly(A) polymerase, TENT-5 would be predicted to have a stabilizing effect on mRNAs by extending their poly(A) tails. We thus hypothesized that transcripts down-regulated in *tent-5* mutants might be direct targets for polyadenylation by TENT-5. To identify mRNAs that may be subjected to TENT-5 polyadenylation, we performed genome-wide poly(A) tail profiling in *tent-5(tm3504)* and wild-type age-synchronized L4 worms using Oxford Nanopore DRS (fig. S5A), which we have previously implemented to find substrates of TENT5C and TENT5A in B cells and osteoblasts, respectively (*23*, *26*). We obtained in total ~12 million transcriptome-wide reads from two wild-type and two mutant biological replicas, of which 70% passed quality control and were used for the analysis of the poly(A) tail lengths (data S4) and differential expression of genes (data S5). The expression levels of identified transcripts were in strong agreement with Illumina RNA-seq results, providing additional confidence in data quality (fig. S5B). Moreover, GO analysis of genes for which expression levels were down-regulated at least 1.5-fold in *tent-5(tm3504)* mutant worms (FDR < 0.05) (data S5) revealed their association with the immune response (fig. S5C), again recapitulating observations obtained by standard RNA-seq (Fig. 2B and data S5).

The distribution of the global poly(A) tail lengths was consistent with the previous reports (Fig. 5A) (*52*, *53*). It has been shown that in adult worms, highly expressed housekeeping transcripts show negative correlation between their poly(A) tail length and mRNA levels (*52*, *53*). Accordingly, we found that genes with highly expressed mRNA tended to have slightly shorter median poly(A) tail length than the ones with medium and low expression levels (fig. S5D), with this trend being particularly evident for the highly expressed transcripts encoding ribosomal proteins (fig. S5E). However, such a trend does not undermine the crucial role of poly(A) tail length in regulating mRNA stability. It rather supports the notion that different functional groups of mRNAs might have various determinates of their stability with varying effects of poly(A) tail length compared to other features, such as ribosome association or sequence motives.

We found 96 mRNAs that had a median poly(A) tail length that was at least 5 nt shorter in the *tent-5*–deficient worms (FDR < 0.05) compared to the wild-type animals (Fig. 5B and data S4). At the same time, we observed no changes in the mRNA polyadenylation between mutant and wild type for the transcripts encoding components of the translational apparatus (Fig. 5C) and other transcripts, which are routinely used as references in high-throughput studies (fig. S5F) (*54*). Almost 70% of these mRNAs were also down-regulated at least 1.5-fold in *tent-5(tm3504)*, according to both RNA-seq and DRS (fig. S5G), implying that they represent the direct substrates for polyadenylation by TENT-5 in wild-type animals. Next, we analyzed the relationship between the median poly(A) tail length

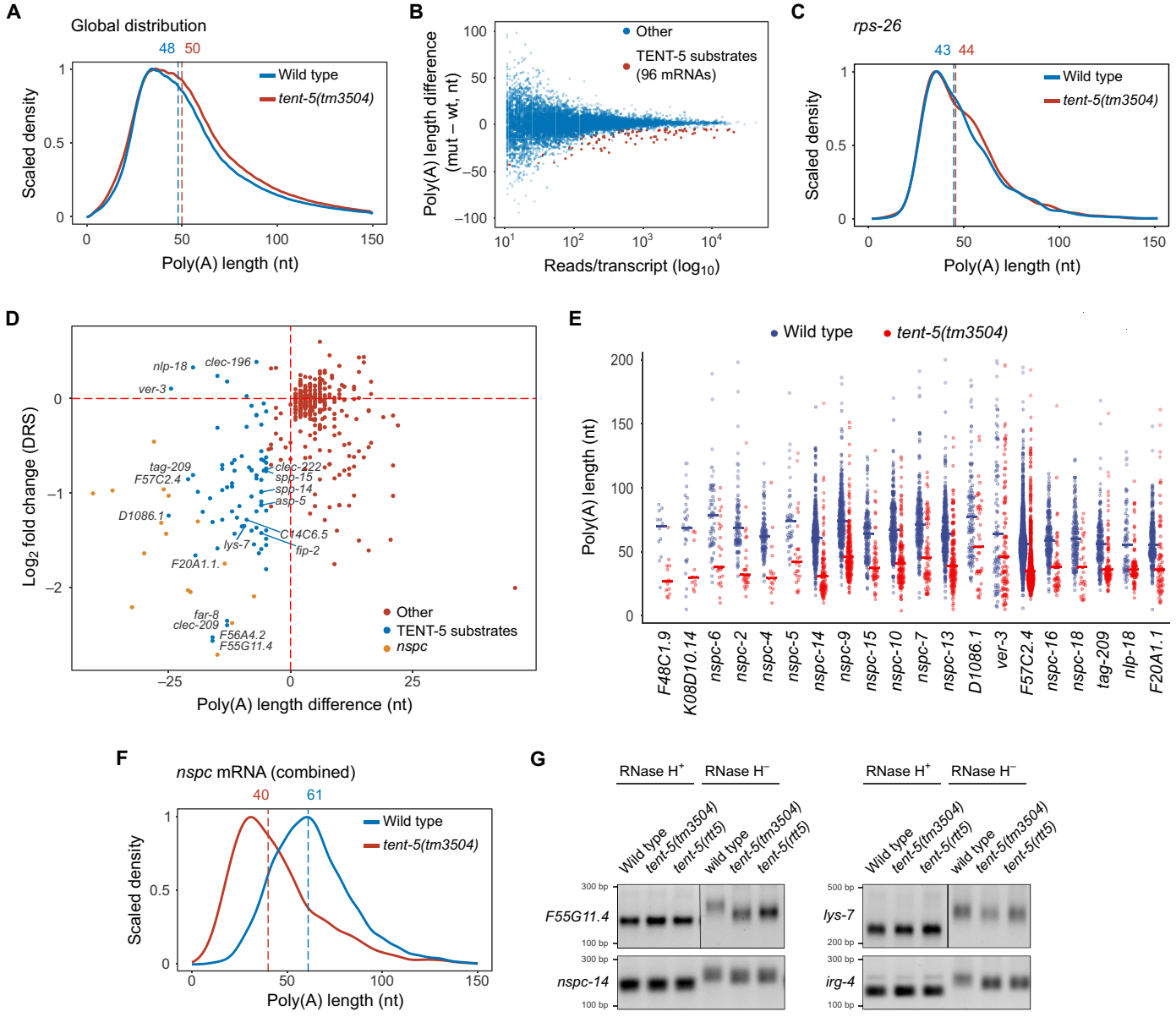


Fig. 5. Transcripts down-regulated in the *tent-5*–deficient mutants are the direct targets for polyadenylation by TENT-5. (**A**) DRS-based global poly(A) tail length profiling of mRNA isolated from *tent-5(tm3504)* and wild type. Density distribution plots are shown for all identified transcripts, scaled to 1. Vertical dashed lines represent median poly(A) lengths (in nucleotides) for each condition. (**B**) Plot showing expression levels and difference in the median poly(A) tail length between *tent-5* mutant and wild-type worms. TENT-5–regulated transcripts ($n = 96$) are indicated in red (FDR < 0.05). (**C**) *rps-26* mRNA that encodes ribosomal protein shows no change in the median poly(A) tail length between mutant and wild-type animals. Plot description as in (A). (**D**) The relationship between the median poly(A) tail length and the expression levels of respective mRNAs in *tent-5*–defective worms. Blue and orange dots indicate TENT-5 substrate mRNAs ($n = 96$). For readability, only IDs of the top 20 substrates and mentioned immune effectors are indicated. (**E**) Bee swarm plots of poly(A) tail length distributions for the top 20 TENT-5 molecular substrates ordered by the median poly(A) tail length fold change. Horizontal lines indicate the median poly(A) tail length for each transcript. (**F**) mRNAs that encode NSPC proteins show significant shortening of the median poly(A) tail length in mutant worms. Plot description as in (A). (**G**) Poly(A) tail PCR tests (PAT) of selected TENT-5–regulated transcripts. Before sample preparation for PAT, samples were treated with RNase H in the presence of oligo(dT)$_{25}$ (RNase H$^+$). PCR products were analyzed on 2% agarose gels.

and the level of respective mRNAs in *tent-5*–defective worms and observed that transcripts with shorter poly(A) tails exhibited a significant decrease in their levels, including for multiple members of the *nspc* gene family (Fig. 5D). The observed correlation between the poly(A) tail length and mRNA level strongly suggests that polyadenylation by TENT-5 increases the stability of target transcripts.

Next, we addressed the impact of polyadenylation by TENT-5 on mRNAs that encode innate immune effectors. We observed that the median poly(A) tail length of mRNAs that encode proteins that contain CUB domains (e.g., *F55G11.4*), CLECs (e.g., *F56A4.2* and *clec-196*), lysozymes (e.g., *lys-7*), and other proteins involved in host defense (e.g., *F48C1.9*, *C14C6.5*, *asp-12*, and *spp-14*) were significantly decreased in the *tent-5(tm3504)* mutant (fig. S5H). We also noticed that 11 mRNAs that encode collagens had 5- to 9-nt shorter median poly(A) tails length in mutant worms (fig. S5I), suggesting that, similarly to mouse *Tent5a* KO (*26*), deficiency of *tent-5* in worms may cause some defects in an extracellular matrix (ECM) formation. Regulation of collagen genes and ECM has been shown to affect a worm's defense (*55*–*58*). Notably, at the top of the list of transcripts with the most significantly shortened poly(A) tails in mutant worms (Fig. 5E and data S4) were *nspc* genes, referred to above. The median poly(A) tail length of all *nspc* mRNAs was decreased from 61 to 40 nt in *tent-5(tm3504)* worms (Fig. 5F). As mentioned above, we also identified *nspc* among genes whose expression levels were most significantly down-regulated in *tent-5*–deficient worms, both by RNA-seq (fig. S5J and data S1) and DRS (Fig. 5D and data S5). Genes from the *nspc* family are rapidly evolving and encode 18 almost identical short (~100 amino acids) proteins that have been proposed to encode antimicrobial peptides (*33*). Next, we examined the poly(A) tail length of four representative transcripts, including *nspc-14*, using poly(A) PCR tests (PAT) that revealed that amplicons from both *tent-5*–deficient mutants were shorter compared to those from wild-type worms (Fig. 5G and fig. S5K). Digestion with ribonuclease H (RNase H) proved that the observed differences resulted from the change in the poly(A) tail length (Fig. 5G). Together, our results lead us to propose that TENT-5 has a stabilizing effect on a specific subset of mRNAs by extending their poly(A) tails.

TENT-5 regulates mRNAs that encode secreted proteins

To determine what distinguishes TENT-5 ncPAP substrates from other mRNAs, we performed a comprehensive analysis of their sequences. First, we observed that TENT-5 substrate mRNAs were shorter (Fig. 6, A and B) and had shorter 3′UTRs (Fig. 6C) compared to all other transcripts identified by DRS. At the same time, there was no difference in 5′UTR length between the TENT-5–regulated and nonregulated mRNAs (fig. S6A). We observed similar effects for the mRNAs of genes whose expression levels were down-regulated in *tent-5(tm3504)* mutant worms compared to wild type according to RNA-seq (fig. S6, B to E). Next, we sought to find linear sequence motifs that were specifically enriched in coding or 3′UTR sequences of mRNAs, whose poly(A) tails were shorter in *tent-5*–deficient worms. We found no specific motifs across these transcripts, although note that the DREME (Discriminative Regular Expression Motif Elicitation) motif discovery tool (*59*) we used only enables the identification of relatively short motifs (up to 8 nt).

However, we noticed that many of the genes for which expression was significantly down-regulated in mutant worms were annotated as encoding proteins of the extracellular space (Fig. 2B and fig. S5C). In *C. elegans*, roughly 19% of all proteins are predicted to have an N-terminal signal peptide that targets them toward the secretory pathway through the endoplasmic reticulum (ER) (*60*). Notably, we observed that as much as 71 and 80% of genes down-regulated in *tent-5*–deficient worms at least 1.5-fold (RNA-seq and DRS, respectively) and 84% of all TENT-5 substrates (DRS) encoded proteins with a predicted signal peptide (Fig. 6D, fig. S6F, and data S4), strongly indicating that TENT-5 is responsible for the polyadenylation of mRNAs that encode ER-targeted secreted proteins. Transcripts that encode secreted antimicrobial proteins, such as lysozymes, caenopores (SPP), and proteins containing CLEC or CUB domains, and some enzymes involved in extracellular macromolecule digestion were remarkably overrepresented among the genes that were down-regulated in *tent-5(tm3504)* mutants (Fig. 3D and fig. S3E) and carried shorter poly(A) tails (fig. S5H). As we observed such a prominent enrichment of mRNAs that encode secreted proteins among TENT-5 substrates and as in worms, the expression of roughly one-third of all genes encoding secreted proteins are induced during infection (*31*, *60*); this could explain the higher susceptibility of *tent-5* mutant to the infection.

Given that TENT-5 regulates levels of mRNAs that encode secreted proteins, we aimed to test whether it is enriched at the ER in the intestinal cells, which have large rough ER (*61*). Although adult *C. elegans* intestines are known to contain multiple autofluorescent gut granules (*62*), we confirmed that granules observed in the *tent-5 (rtt6[tent-5::gfp::3xflag] I)* knockin strain (Fig. 1F) do not result from the aggregation of TENT-5-GFP. The chromatography analysis of total protein extracts from the *tent-5(rtt6)* worms revealed that most of the protein is present in the fraction corresponding to the TENT-5-GFP molecular weight (fig. S6G). However, a small amount of TENT-5 was also found in macromolecular complexes or aggregates. The expression levels of TENT-5 (fig. S1E) and animal viability (fig. S6H) of *tent-5::gfp* strain were unaltered in comparison to the wild-type animals, pointing to the expression of functional protein. Nevertheless, to avoid the misinterpretation of the microscopic results due to the possible presence of aggregates, we excluded the strongest signals observed in a GFP channel from the subsequent colocalization analysis. To test whether TENT-5 localizes to the ER, we crossed *tent-5::gfp* worms (Fig. 1F and fig. S1, E to H) with a transgenic strain expressing the ER marker protein TRAM-1 (translocating chain-associating membrane protein transporter) fused to a CemOrange2 fluorescent protein designed to be exclusively expressed in the intestine (*vkEx2664 [nhx-2p::CemOrange2::tram-1, myo-2p::gfp]*) (*63*). The microscopic analysis revealed partial colocalization of CemOrange2–TRAM-1 and TENT-5–GFP (Pearson correlation coefficient: 0.4 ± 0.11 SD, $n = 18$) (Fig. 6E), indicating that a fraction of TENT-5 localizes to the ER. This observation is in agreement with the recently published data showing that TENT5C partially localizes to ER in mammalian cells (*23*, *64*, *65*). To analyze TENT-5 association with ER further, we performed a subcellular protein fractionation of whole worm extracts prepared from wild-type, *tent-5::gfp*, and *vgln-1::mKate2 (vgln-1(rtt9[vgln-1::mKate2:::3xmyc] II)* strains. VGLN-1 (ViGiLN homolog) is localized in the cytoplasm and nucleus (*66*), and its human ortholog, HDLBP (high density lipoprotein binding protein), has been also shown to be associated with the ER (*67*). Subcellular protein fractionation confirmed that TENT-5–GFP resides in the cytosol but is also present at membranes, including ER (Fig. 6F and fig. S6, I and J). Our results allow speculating that TENT-5 might be targeted to at least some of its substrates by colocalization with ER-translated short mRNAs. However, the exact mechanism of the TENT-5 substrate specificity remains to be elucidated.

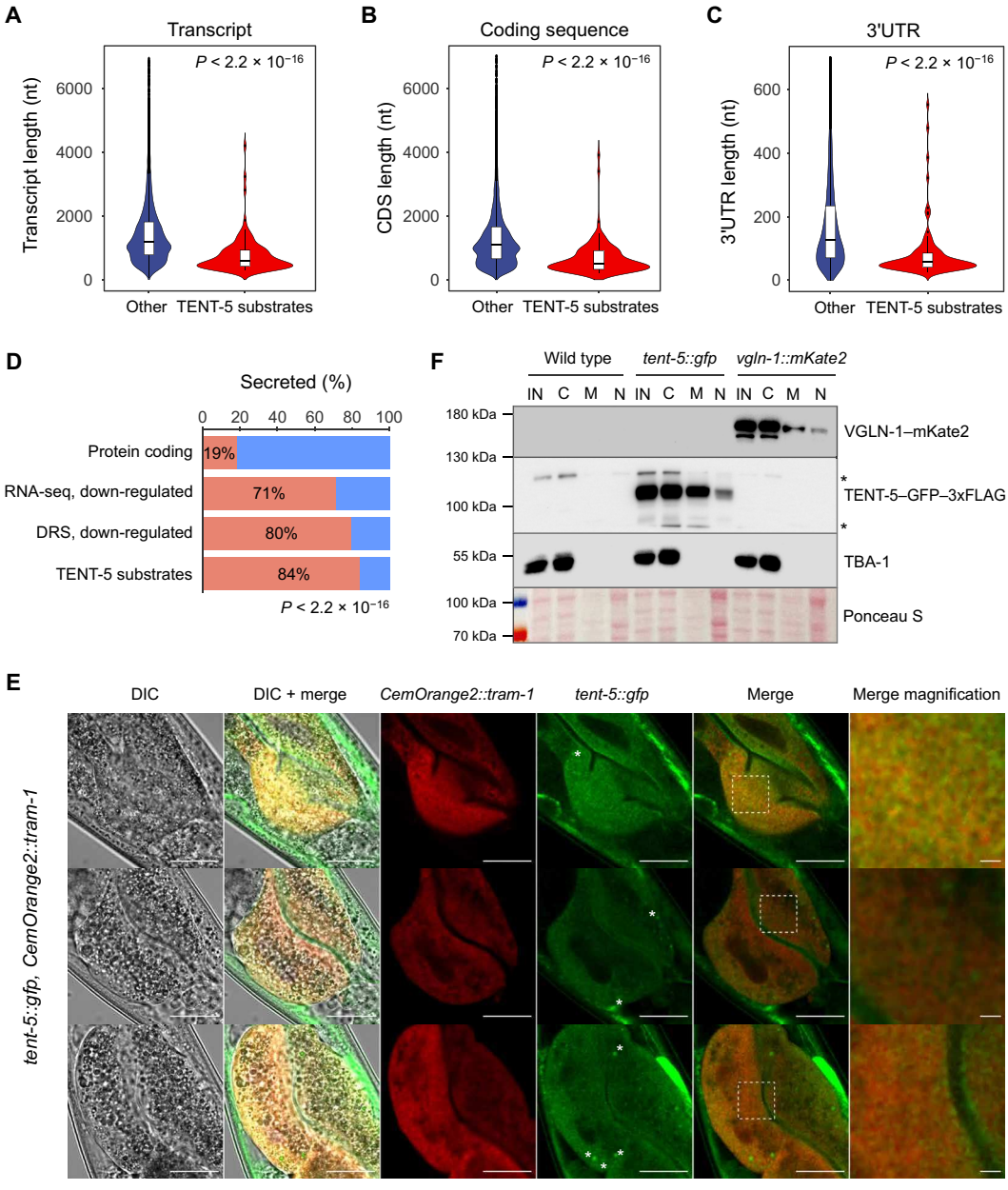


Fig. 6. TENT-5 regulates mRNAs that encode secreted proteins. Violin plots showing length distribution of transcripts (**A**), coding sequence (**B**), and 3′UTRs (**C**) of TENT-5 substrates ($n = 96$) compared to other transcripts identified by DRS ($n = 16{,}568$) (*P* value, Wilcoxon test). (**D**) Fractions of genes that encode secreted proteins (red) as defined in (*60*) across indicated datasets: expressed protein-coding genes (WS270; RNA-seq, base mean > 0), RNA-seq and DRS *tent-5(tm3504)* down-regulated genes [FDR < 0.05, $\log_2$FC < –log2(1.5)], and TENT-5 substrates. *P* value for enrichment of genes encoding secreted proteins in each group versus whole genome is $<2.2 \times 10^{-16}$ (Fisher's exact test). (**E**) Representative fluorescence and DIC microscopy images of distal parts of the adult worms' intestines, which coexpress TENT-5-GFP and ER marker protein CemOrange2–TRAM-1. Both tagged proteins partially colocalize at the ER (Pearson correlation coefficient: 0.4 ± 0.11 SD, 18 worms). Analysis was restricted to cells expressing CemOrange2–TRAM-1, and the strongest unspecific signal from the green channel (examples are marked with asterisks) was excluded from the analysis. White dashed squares in the merge picture indicate the region magnified on the right. Scale bars, 20 μm (all pictures) and 2 μm (magnification). (**F**) TENT5C is present in the ER fraction. Subcellular fractionation of proteins isolated from wild-type, *tent-5::gfp::3xflag*, and *vgln-1::mKate2:::3xmyc* strains followed by Western blot. Anti-tubulin (TBA-1) and anti–red fluorescent protein (VGLN-1–mKate2) antibodies and Ponceau S staining were used as fractionation controls. IN, input; fractions: C, cytoplasmic; M, membrane; N, nuclear. Asterisks indicate nonspecific bands.

The role of TENT-5 in innate immunity is evolutionarily conserved

Because TENT-5 is orthologous to mammalian TENT5 ncPAPs, we hypothesized that mammalian TENT5 proteins might also have a role in innate immunity. Murine macrophages, key innate immune effector cells, express two of the four *Tent5* genes, *Tent5a* and *Tent5c*. We found that both genes were induced in bone marrow–derived macrophages (BMDMs) upon stimulation with lipopolysaccharide (LPS), which is a potent activator of inflammatory and immune response in M1 macrophages (Fig. 7A) (*68*). To test whether TENT5A

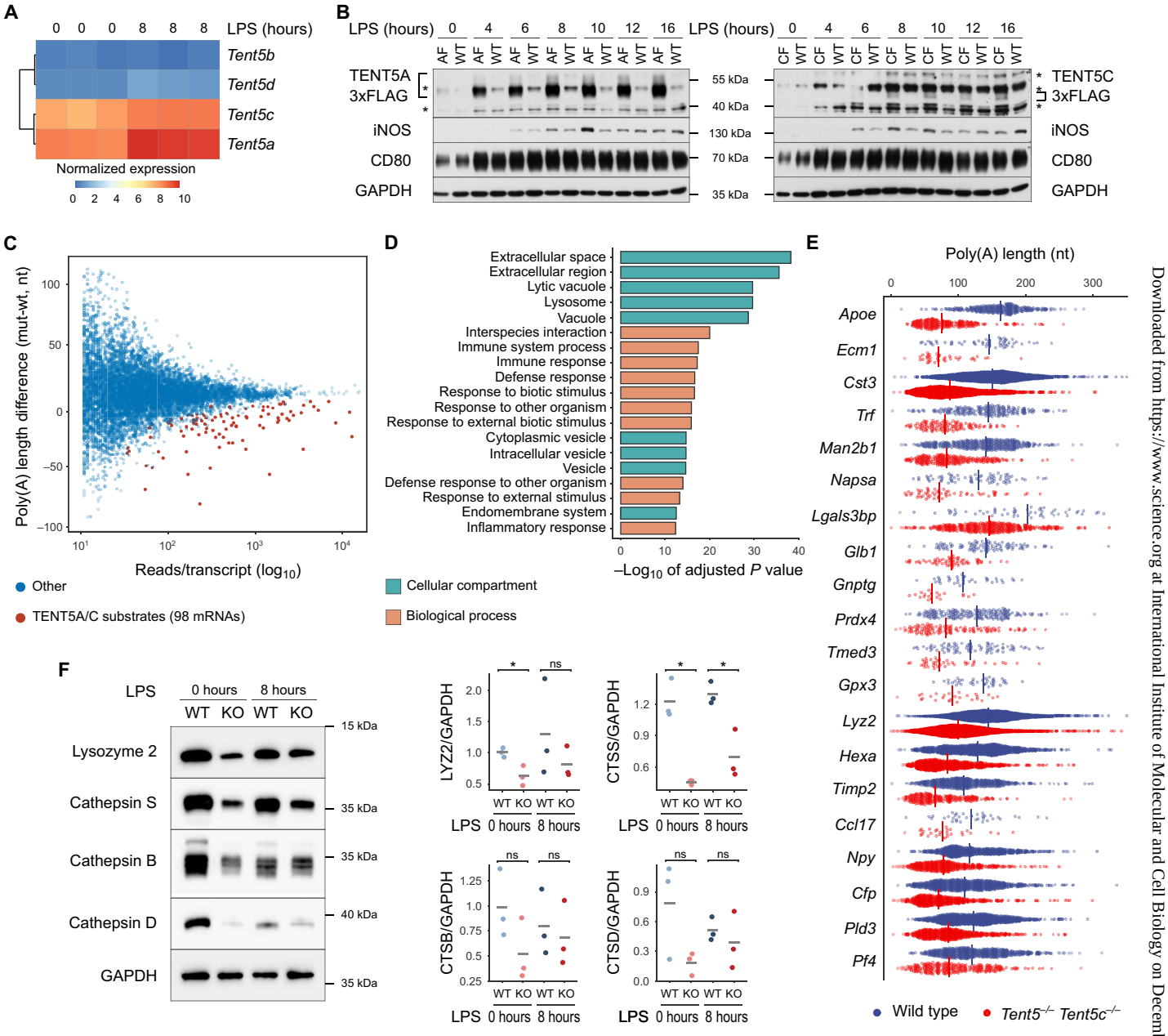


Fig. 7. The role of TENT-5 in innate immunity is evolutionarily conserved. (**A**) Heatmap showing expression levels of *Tent5* genes in BMDMs isolated from the wild-type mice. Stimulation of BMDMs with LPS led to increased expression levels of *Tent5a/c*. (**B**) TENT5A/C were elevated in macrophages stimulated with LPS as illustrated by anti-FLAG immunoblots on extracts prepared from wild-type (WT), TENT5A-3xFLAG (AF), and TENT5C-3xFLAG (CF) BMDMs. iNOS and CD80 levels serve as a control of BMDM activation. The lines and asterisks indicate specific and unspecific bands, respectively. (**C**) Plot showing expression levels and difference in the median poly(A) tail length of mRNAs that were isolated from *Tent5a*$^{-/-}$ *Tent5c*$^{-/-}$ and wild-type BMDMs. TENT5A/C-regulated transcripts ($n = 98$) are marked in red (FDR < 0.05). (**D**) Top 20 (ordered by adjusted *P* value) GO terms of TENT5A/C substrate mRNAs (data S6). (**E**) Bee swarm plots of poly(A) tail length distributions for the top 20 TENT5A/C molecular substrates ordered by the median poly(A) tail length fold change (data S6). (**F**) Western blots showing that the lack of TENT5A/C leads to the lower abundance of indicated peptides both in unstimulated (LPS; 0 hours) and stimulated (LPS; 8 hours) BMDMs. Glyceraldehyde phosphate dehydrogenase (GAPDH) serves as a loading control. Shown is the representative result of three biological replicates (fig. S7K). Graphs on the right show values of the band intensity ratio between tested protein and GAPDH. Points indicate values of each of three biological replicates, and lines denote the average value for biological replicas. *$P \leq 0.05$ (two-tailed *t* test).

and TENT5C proteins also were activated during infection, we examined their levels in the time course of LPS treatment in BMDM isolated from TENT5A-3xFLAG (*26*) and TENT5C-3xFLAG knockin mouse lines (*21*). The effectiveness of BMDM activation was examined by analysis of M1 macrophage polarization markers inducible nitric oxide synthase (iNOS) and CD80 (Fig. 7B) (*69*). Similarly, TENT5A-3xFLAG and TENT5C-3xFLAG levels were elevated in response to LPS stimulation (Fig. 7B), consistent with the notion that the role of TENT5 ncPAPs in innate immunity could be evolutionarily conserved.

To determine the physiological relevance of TENT5A and TENT5C in macrophages, we aimed to analyze the phenotypes of mice devoid of *Tent5a* and *Tent5c* genes. *Tent5c* KO mice do not exhibit any gross phenotypes (*21*, *23*, *70*), whereas *Tent5a* KO are smaller than wild type and display skeletal abnormalities (*26*). The double KO of *Tent5a* and *Tent5c* leads to preweaning lethality that suggests potential redundancy between two ncPAPs. To overcome the lethality of the double *Tent5a Tent5c* KO, we used the $Tent5a^{Flox/Flox}$ $Tent5c^{-/-}$ mouse line that was obtained by crossing $Tent5a^{Flox/Flox}$ conditional KO (cKO) with the previously described $Tent5c^{-/-}$ KO line (*21*). We have isolated BMDMs from $Tent5a^{Flox/Flox}$ $Tent5c^{-/-}$ and wild-type mice and delivered Cre recombinase to macrophages by lentiviral transduction that allowed us to obtain vital $Tent5a^{-/-}$ $Tent5c^{-/-}$ cells. Next, we treated BMDMs with LPS for 8 hours, a time point when expression of TENT5A and TENT5C is highly elevated (Fig. 7, A and B), and collected cells for subsequent analyses. To identify potential TENT5A and TENT5C substrates in activated macrophages, we performed transcriptome-wide poly(A) tail profiling by DRS. We obtained approximately 8 million transcriptome-wide reads from two $Tent5a^{-/-}$ $Tent5c^{-/-}$ and two wild-type samples (data S6). Again, the global distribution of the median poly(A) lengths was consistent with previous reports (*52*). We observed no differences in the global polyadenylation status between $Tent5a^{-/-}$ $Tent5c^{-/-}$ and wild-type samples (fig. S7A), as well as changes in the median poly(A) tail length of mRNAs encoding ribosomal proteins (fig. S7B). We found that in BMDMs isolated from $Tent5a^{-/-}$ $Tent5c^{-/-}$ BMDMs, 98 mRNAs had median poly(A) tail lengths shorter (FDR < 0.05) in comparison to the samples obtained from the wild type (Fig. 7C and data S6). Many of these mRNAs were also identified as molecular substrates for TENT5A or TENT5C in osteoblasts (*26*) and multiple myeloma cell lines (fig. S7C) (*21*), indicating that they represent direct substrates for polyadenylation by TENT5A/C. GO analysis of TENT5A/C substrate mRNAs revealed their strong association with the extracellular space, defense and innate immune response, and lytic vacuole (Fig. 7D and data S6), recapitulating observations obtained with *tent-5* mutant worms. Notably, only 14 mRNAs had median poly(A) tail length at least 5 nt longer in $Tent5a^{-/-}$ $Tent5c^{-/-}$ cells compared to wild type (data S6). Some of these mRNAs also show an association with immune response and extracellular space (data S6) that, similar to our observations in worms, may indicate compensatory effects in cells with defective immune responses. We did not attempt to establish the relationship between the poly(A) tail length and the expression levels of the respective mRNAs, as only one gene (*Lgals3bp*) was differentially expressed between $Tent5a^{-/-}$ $Tent5c^{-/-}$ and wild-type samples (data S6).

In line with our observations in worms, a comprehensive analysis of TENT5A/C-regulated transcripts revealed that they were shorter (fig. S7, G and H), tended to have shorter 3′UTRs (fig. S7I), and tended to have moderately shorter 5′UTRs (fig. S7J) compared to other transcripts detected in macrophages by DRS. Analysis of TENT5A/C substrates identified no sequence motif enrichment across their 3′UTRs. Fifty-nine of 98 (~60%) TENT5A/C substrate mRNAs encode extracellular space proteins (data S6), suggesting that TENT5A/C preferentially polyadenylates mRNAs that encode secreted proteins, many of which (32 of 59) have a role in the immune response. Among mRNAs with shorter median poly(A) tail length in $Tent5a^{-/-}$ $Tent5c^{-/-}$ macrophages (Fig. 7E and data S6) were mRNAs that encode lysozyme (*Lyz2*) (fig. S7D) and lysosomal proteases cathepsins (*Ctsd* and *Ctsb*) (fig. S7, E and F). These proteins transit the ER during their maturation and are homologous to TENT-5 substrates in worms, and their role in the mammalian immune response is well documented (*71*–*77*). Therefore, we examined whether the lack of TENT5A and TENT5C affects the protein levels of lysozyme and cathepsins. We have observed a definite but modest drop in LYZ2 and cathepsins (Fig. 7F and fig. S7K) levels in $Tent5a^{-/-}$ $Tent5c^{-/-}$ BMDMs, confirming the direct effect of cytoplasmic polyadenylation by TENT5A/C on innate immunity-related protein expression. Our observation that in BMDM devoid of *Tent5a* and *Tent5c*, the shortening of the poly(A) tails of a specific group of mRNAs is not accompanied by significant down-regulation of these mRNAs' expression but affects levels of respective proteins allows speculating that at least in activated BMDM, cytoplasmic polyadenylation may have a more prominent influence on translation efficiency than on mRNA stability. On the basis of our results, we concluded that the enzymatic activity of TENT5A/C plays a role in the physiology of murine macrophages and, similarly to its ortholog in worms, may facilitate the innate immune response.

DISCUSSION

In its natural habitat, *C. elegans* encounters numerous and diverse pathogens. Worms' fitness and rapid adaptation to an ever-changing microbial environment require dynamic and highly efficient modulation of the immune response. It is widely appreciated that in animals, innate immunity is orchestrated by a plethora of transcriptional and posttranscriptional mechanisms (*7*, *78*–*80*). The posttranscriptional aspect of innate immune regulation in worms is a fast-growing field of research (*58*, *81*–*84*). *C. elegans* can sense bacterial noncoding RNAs to induce avoidance of a pathogen (*85*). A few reports have also implicated microRNAs in host defense (*86*–*88*), whereas viral RNA uridylation, by one member of the TENT family, CDE-1, has been shown to play a role in the antiviral response (*89*). Cytoplasmic polyadenylation is a powerful posttranscriptional mechanism that shapes the transcriptome and consequently also the proteome, through the regulation of mRNA stability and translation efficiency. This work shows that cytoplasmic polyadenylation by ncPAP TENT-5 positively regulates innate immunity in *C. elegans*.

TENT-5 is a cytoplasmic protein that is expressed through the whole life cycle of the worm in multiple cells of the body, including the intestine (Fig. 1). In worms, the immune response and digestion are connected. Many enzymes responsible for the macromolecular degradation of food participate in the degradation of pathogen-derived macromolecules (*61*). Upon infection with bacterial pathogens that infect through the gut, intestinal cells secrete a large amount of enzymes and antibacterial proteins. Intestinal cells must have an enormous capacity for protein synthesis and secretion. Our results show that TENT-5 preferentially polyadenylates and stabilizes mRNAs that encode short secreted proteins with a role in digestion and immunity (Figs. 5 and 6). The length of the poly(A) tail is critical for mRNA stability, and the posttranscriptional lengthening of mRNAs' poly(A) tails in the cytoplasm may thus extend their half-life. Such a mechanism, extending mRNA longevity and promoting translation, would be not only energy effective but also extremely fast, a valuable feature when a rapid reaction to changing environmental conditions is needed, for example, during immune or stress responses.

The relevance of TENT-5 in innate immunity is also supported by the reduced survival of *tent-5*–deficient worms upon infection with a range of bacterial pathogens (Fig. 4). Mutant worms display

a moderately reduced life span also when grown on *E. coli*. Furthermore, this survival defect is associated with the increased bacterial load in the intestine of the *tent-5*–deficient worms. This observation may suggest that because of the decreased basal expression of genes encoding digestive enzymes and cytoprotective proteins, mutant worms have limited ability to deal even with relatively nonpathogenic food, exhibiting reduced life span. In such an interpretation, TENT-5 may not be solely dedicated to innate immunity but rather influence stability and potentially translation efficiency of many mRNAs encoding secreted proteins. The consequences of *tent-5* deficiency are, however, most evident under physiological conditions when secretion plays a life-saving role. The fact that TENT-5 is up-regulated upon infection may be explained by specific pathogen–induced transcriptional regulation or by a mechanism that senses the demand for efficient protein secretion. Further research will be needed to determine what other secretion-dependent processes are regulated by TENT-5.

Direct sequencing of RNA samples prepared from *tent-5* mutants and wild-type worms allowed us to uncover substrates of TENT-5 enzymatic activity (Fig. 5). Our analysis of poly(A) tails lengths in L4 *C. elegans* is consistent with previous results (*52*, *53*). Those studies identified a negative correlation of poly(A) tail length with mRNA expression. Given the counterintuitive nature of these observations, we sought to take advantage of the high quality of our DRS data and analyzed the relationship between these features. In accordance with (*52*, *53*), our data showed a similar phenomenon (fig. S5). However, TENT-5 substrate mRNAs showed significantly decreased expression levels in *tent-5(tm3504)* mutants, suggesting that shortening of the poly(A) tail from its wild-type length lowers mRNAs expression (Fig. 5). Thus, our results show that for TENT-5 substrates, an increase in poly(A) tail length sustains mRNA steady-state levels.

Our data indicate that TENT-5 affects mRNAs that encode short secreted proteins. Among the prominent TENT-5 targets are mRNAs of the *nspc* family of genes for which mRNA polyadenylation and expression are strongly decreased in the *tent-5* mutant (Fig. 5). In our previous studies, we also found that TENT5 family members regulate the expression of genes encoding secreted proteins. In B cells, TENT5C polyadenylates mRNAs that encode immunoglobulins (*23*), and in osteoblasts, the main substrates of TENT5A are mRNAs encoding collagens (*26*). In line with that, we show that mRNAs that encode secreted proteins constitute a large fraction of TENT5A and TENT5C direct substrates in macrophages (Fig. 7). In agreement with the nature of TENT5 substrates, these poly(A) polymerases are associated with the ER, and this is probably the main determinant of their substrate specificity. It was recently shown that TENT5C is actively recruited to ER in human cells through interaction with fibronectin type III domain–containing proteins FND3CA and FNDC3B (*64*, *65*). Disturbance of TENT5C function leads to ER shrinking, destabilization of ER-translated mRNAs, and defects in ER-mediated protein folding and secretion (*64*). It is probable that TENT5C's binding to the ER, mediated by FNDC3A and FNDC3B, may be enough to target its substrates. Whether the worm's TENT-5 target selection corresponds to an analogous mechanism requires further investigation, because among proteins with fibronectin type III domains, there are no obvious FNDC3 orthologs encoded in the genome. Controlling the secretory capacity of the ER is key for resistance to infection (*90*, *91*). Cytoplasmic polyadenylation by TENT5 proteins of mRNAs encoding proteins that transit the ER provides a new layer to the regulation of secretion essential for the proper immune response and other physiological processes such as bone formation in vertebrates.

Last, we demonstrate that the role of TENT-5 in innate immunity is evolutionarily conserved. In murine macrophages, TENT-5 orthologs, TENT5A and TENT5C, polyadenylate mRNAs that encode lysozyme and cathepsin proteases (Fig. 7), increasing their protein level. Lysozyme is one of the most abundant antimicrobial proteins secreted by macrophages, and deletion of *Lyz2* increases susceptibility to infection with *P. aeruginosa*, *Micrococcus luteus*, and *Klebsiella pneumoniae* due to impaired clearance of pathogen and can lead to higher host mortality (*71*–*74*). Cathepsins not only play a notable role in lysosomal protein breakdown but also regulate the immune response (*92*). BMDMs isolated from *Ctsd*-deficient mice display enhanced susceptibility to *Listeria monocytogenes* infection and increased intraphagosomal viability of bacteria (*75*). Moreover, cathepsin D protein levels were up-regulated after infection of murine macrophages with *Bacillus subtilis*, *P. aeruginosa*, *L. monocytogenes*, *S. aureus*, and *E. coli*, and again, its deficiency led to an increase in the amounts of each of these bacteria inside populations of macrophages (*76*). Cathepsin B was also required for optimal posttranslational processing of tumor necrosis factor–α (TNF-α) in response to LPS, and BMDMs from *Ctsb*-deficient mice secrete significantly less TNF-α than wild-type macrophages (*77*).

Our work demonstrates the conserved role of mRNA polyadenylation and TENT5 family ncPAPs in the regulation of innate immunity in animals. Taking into account that, in worms, TENT-5 is expressed in multiple tissues, we expect that its functions go beyond protection against pathogens and may be generally important for physiological processes that involve protein secretion. In mammals, TENT5A to TENT5D are expressed in different tissues and developmental stages, opening the possibility for them to have broad biological significance and functional interactions too.

MATERIALS AND METHODS

Bacterial strains and culture

E. coli HB101 was cultured at 37°C in LB medium supplemented with streptomycin (final concentration of 0.1 mg/ml). *P. aeruginosa* PAO1 and *S. marcescens* Db10 were grown at 37°C in LB medium without antibiotics (*48*, *93*). *P. luminescens* Hb was grown at 30°C in LB medium without antibiotics (*48*). *S. aureus* NCTC 8325 (Argenta Ltd.) was cultured at 37°C on tryptic soy agar (TSA) or tryptic soy broth (TSB) (both from BD Biosciences) supplemented with nalidixic acid (Nal) (final concentration of 10 μg/ml). *E. coli* DH5α and MH1 strains were cultured at 37°C in LB supplemented with appropriate antibiotics. All antibiotics were purchased from Sigma-Aldrich.

C. elegans strains and maintenance

C. elegans was maintained on nematode growth medium (NGM) plates seeded with *E. coli* HB101 at 20°C unless otherwise specified. Strains obtained from the National Bioresource Project of Japan (NBRP) and strains generated by the CRISPR-Cas9 were outcrossed two to nine times to the wild-type strain. The following *C. elegans* strains were used: N2 Bristol (wild type) and VK2664 (*vkEx2664 [nhx-2p::CemOrange2::tram-1, myo-2p::gfp]*) strains were obtained from the Caenorhabditis Genetics Center; *tent-5(tm3504) I* was obtained from NBRP; ADZ20 (*tent-5(rtt5) I*), ADZ21 (*tent-5(rtt6[tent-5::gfp::3xflag]) I*), and ADZ24 (*vgln-1(rtt9[vgln-1::mKate2:::3xmyc] II*) were generated in this study; ADZ87 (*tent-5(rtt6[tent-5::gfp::3xflag])*

I, vkEx2664 [nhx-2p::CemOrange2::tram-1, myo-2p::gfp]) strain was obtained by crossing ADZ21 with VK2664.

C. elegans transgenic strain generation

Transgenic worm strains were generated using standard microinjection protocols. Plasmids that were used for microinjections were purified with PureLink mini-prep kit (Thermo Fisher Scientific, K210002). Injections were conducted using the Axio Observer D1 inverted microscope (Zeiss) equipped with a Femto Jet 4i microinjection system (Eppendorf). For each transformation, at least two independent transgenic strains were obtained. All oligonucleotides and DNA constructs used for transgenic strain generation are listed in tables S2 and S3, respectively.

The KO strain ADZ20 *tent-5(rtt5) I* that harbors a 2909–base pair deletion, which spans from the start to stop codon of *tent-5* isoform d, was generated using an adapted version of the CRISPR-Cas9 protocol (*94*). Mutation in the *dpy-10* gene was used as a CRISPR co-conversion marker. Single-stranded DNA (ssDNA) encoding tracrRNA (trans-activating CRISPR RNA) and CRISPR RNA (crRNA) with 20 N of single-guide RNA (sgRNA) sequences was used for the preparation of templates for in vitro sgRNA synthesis. Briefly, 5 μl of 100 μM VL311 tracrRNA oligo was annealed with 5 μl of 100 μM sgRNA oligo (VL312, VL313, and VL315) in the 50-μl mix containing deoxynucleotide triphosphates, Phusion buffer, and Phusion Hot Start II Polymerase and incubated at 98°C for 3 min; 98°C for 10 s, 65°C for 20 s, and 72°C for 5 s for 10 cycles; and 72°C for 5 min. The reaction was purified with AMPure XP magnetic beads [1:1.4 (v/v) mix to beads; Beckman Coulter, A63882]. In vitro transcription was assembled by mixing 400 ng of DNA template in 35 μl of RNase-free water, 5 μl of ribonucleoside triphosphates (NTPs) mix (20 mM each), 10× transcription buffer [200 mM tris-HCl (pH 7.9), 30 mM $MgCl_2$, 50 mM dithiothreitol (DTT), 50 mM NaCl, and 10 mM spermidine], 1.25 μl of RiboLock RNase Inhibitor (40 U/μl) (Thermo Fisher Scientific, EO0384), and 4 μl of T7 polymerase (homemade). Eight reactions were set up for the single sgRNA transcription. Following incubation at 37°C for 3 hours, each sample was treated with 0.5 μl of TURBO deoxyribonuclease (DNase) (2 U/μl; Thermo Fisher Scientific, AM2239) at 37°C for 30 min. RNA was purified from the pooled reactions with the phenol/chloroform extraction and purified further through electrophoresis in 6% urea polyacrylamide gel electrophoresis (PAGE). Animals were injected with the following mix: 15.5 μM Cas9 (Cas9::NLS_{SV40}::His_6 protein; homemade), 5.9 μM sgRNA–dpy-10, 11 μM sgRNA–tent-5-1, 11 μM sgRNA–tent-5-2, 0.44 μM single-stranded oligodeoxynucleotide (ssODN)–dpy-10, 0.88 μM ssODN–tent-5, 150 mM KCl, and 20 mM Hepes (pH 8.0). Worms showing dumpy phenotype in F_1 progeny were screened for *tent-5* deletion using PCR, and, later, the deletion was confirmed by Sanger sequencing. Mutant worms were backcrossed two times to wild-type worms to cross out the *dpy-10* mutation and CRISPR off-targets.

The knockin strain ADZ21 *tent-5(rtt6[tent-5::gfp::3xflag]) I* was generated by CRISPR-Cas9 according to (*95*). The gRNA sequence (5′-TGCCACCAGATGCAGCTACA-3′) was cloned into pDD162 to generate pDD162_*sgRNA425*. Homology arm regions were amplified by PCR using genomic DNA (gDNA) as a template and were inserted into pDD282, resulting in the pDD282-*tent-5::gfp::3xflag* construct. N2 worms were injected with the following mix: pDD282-*tent-5::gfp::3xflag* (10 ng/μl), pDD162_*sgRNA425* (50 ng/μl), *myo-2p::mCherry* pharyngeal coinjection marker pCFJ90 (2.5 ng/μl), and *myo-3p::mCherry* body wall muscle coinjection marker pCFJ104 (5 ng/μl). The selection of positive knockin candidates was performed as described (*95*). Animals with successful GFP-tag insertion were backcrossed three times with the wild-type strain to get rid of CRISPR off-target effects. The knockin strain ADZ24 *vgln-1(rtt9[vgln-1::mKate2:::3xmyc] II* was generated by CRISPR-Cas9 using constructs encoding gRNA, pDD162_*sgRNA456* (gRNA sequence: 5′-CGTTCTTTACCAACGACGAG-3′), and homology arm regions, pDD287-*vgln-1::mKate2::3xmyc*.

Plasmid construction

General cloning techniques were conducted according to the well-established protocols (*96*) or manuals provided by the manufacturers of kits. All plasmids were generated using either classical restriction enzyme digestion and ligation or sequence- and ligation-independent cloning (SLIC) (*97*, *98*) and validated by digestion with restriction enzymes and sequencing. All oligonucleotides and DNA constructs are listed in tables S2 and S3. To generate pDD162-*sgRNA425* (pVL060), two PCRs were performed using pDD162 as a template, with primers VL342 and VL344 and with VL343 and VL345. Fragments were gel-purified and used in a 1:1 molar ratio as templates for PCR with primers VL386 and VL387. The product was purified from a gel and used for the SLIC with pDD162 that has been digested with Nde I and Sph I. pDD162-*sgRNA456* (pVL069) was cloned in a similar way using primers VL365 and VL364 instead of VL344 and VL345. pDD282-*tent-5::gfp::3xflag* (pVL062) was prepared as follows: Arms homologous to *tent-5* were amplified with VL347 and VL348 and with VL349 and VL350 on gDNA isolated from N2. pDD282 was digested with Avr II and Spe I, and the reaction was purified using a Clean-Up kit (A&A Biotechnology). For pDD287-*vgln-1::mKate2::3xmyc* construct (pVL070), homology arms were amplified from gDNA with primers VL366 and VL367 as well as VL368 and VL369. The pDD287 vector was digested with Avr II and Ngo MIV. In both cases, 200 ng of vector and 50 ng of each homology arm products were used for SLIC. The pClneo-NHA (N-terminal lambda symbolN boxB-binding domain and an HA-tag) constructs for tethering assays pCI-*NHA-tent-5*^{WT} and pCI-*NHA-tent-5*^{MUT} were cloned as follows: PCR products were generated with primers NHATENT-5_fw and NHATENT-5_rev on plasmids carrying *tent-5*^{WT} and *tent-5*$^{D151A, D153A}$ (isoform a) genes and subsequently cloned into Sal I and Not I sites of the pClneo-NHA.

Mice

All mice lines were generated by CRISPR-Cas9 in the Genome Engineering Unit (https://crisprmice.eu/) using methods described in (*21*, *23*, *26*). Briefly, a cKO *Tent5a*$^{Flox/Flox}$ (B6;CBA-*Tent5a*$^{Flox/Flox}$/Tar) mouse line was created by insertion of *LoxP* sites in introns flanking exon 2, which contains triplets encoding the catalytic center of the protein (D144N and D146N). Cas9-generated double-strand breaks in gDNA were targeted using two chimeric sgRNA. Bam HI restriction sites were inserted next to *LoxP* sites to facilitate genotyping. Donor mice were handled and injected as described before (*26*). The CRISPR cocktail consisted of mRNA Cas9 (25 ng/μl), sgRNAs (15 ng/μl), and ssDNA repair template (6 ng/μl). Correct integration of *LoxP* sites was confirmed by Sanger sequencing and followed by routine mice genotyping. Sequences of the sgRNAs, ssDNA donor, and primers used for sequencing and genotyping of *Tent5a*$^{Flox/Flox}$ mice can be found in table S2. Double *Tent5a*$^{Flox/Flox}$ *Tent5c*$^{-/-}$ mouse line was obtained by crossing *Tent5a*$^{Flox/Flox}$ cKO with the previously described *Tent5c*$^{-/-}$ (B6;CBA-*Tent5c*^{em1}/Tar) KO line (*21*). *Tent5c*-3xFLAG (B6;CBA-*Tent5c*$^{3xFLAG/3xFLAG}$/Tar) knockin mouse line

was generated as described in (*21*), with the exception that 3xFLAG was added instead of 1xFLAG. *Tent5a*-3xFLAG (B6;CBA-*Tent5a*$^{3xFLAG/3xFLAG}$) knockin line were described previously (*26*). Mice were bred in the animal house of Faculty of Biology, University of Warsaw and maintained under conventional conditions (*21*, *26*). All animal experiments were approved by the First Local Ethical Committee in Warsaw affiliated to the University of Warsaw, Faculty of Biology (approval numbers: WAW/176/2016 and WAW/772/2018) and were performed according to Polish Law (act number 266/15.01.2015) and in agreement with the corresponding European Union directive.

Primary BMDM cell culture

The primary BMDM cell cultures were established from the bone marrow monocytes isolated from *Tent5a*$^{Flox/Flox}$ *Tent5c*$^{-/-}$, *Tent5a*-3xFLAG, *Tent5c*-3xFLAG, and wild-type mice. Mice were euthanized by cervical dislocation at ages 13 to 22 weeks. Femurs and tibias were isolated, the ends of bones were cut, and the bone marrow was flushed with medium using a 25-gauge needle. Bone marrow cells were plated in Iscove's modified Dulbecco's medium (Thermo Fisher Scientific, 21980065) supplemented with 10% fetal bovine serum (FBS) (Gibco), penicillin (100 U/ml)/streptomycin (0.1 mg/ml) solution (Sigma-Aldrich), and macrophage colony-stimulating factor (10 ng/ml; PeproTech, 315-02) and cultured at 37°C in 5% CO_2 as described previously (*99*). For conditional gene targeting, BMDMs derived from *Tent5a*$^{Flox/Flox}$ *Tent5c*$^{-/-}$ and wild-type mice were transduced on the 8th day after isolation with 1 ml of concentrated lentivirus solution per 1 million cells. The medium was changed 16 hours after transduction. The lentivirus production was performed as described previously (*21*). Lentiviral packaging (pMD2.G) and envelope (psPAX2) plasmids were provided by D. Trono (École Polytechnique Fédérale de Lausanne, Switzerland). The pCAG-Cre-IRES2-GFP plasmid was a gift from J. Jaworski (International Institute of Molecular and Cell Biology, Warsaw, Poland). The floxed locus was genotyped 3 days after transduction using gDNA isolated from 0.5 million cells with a Genomic Mini kit (A&A Biotechnology). Sequences of primers used for genotyping are listed in table S2. For BMDM stimulation with LPS, on the 14th day after isolation, cells were treated with LPS (100 ng/ml; Santa Cruz Biotechnology, sc3535) for 4 to 16 hours depending on the experiment.

Phylogenetic analysis

Sequences of TENT proteins used for the phylogenetic analysis have been obtained from the WormBase WS272 (*C. elegans*) and UniProt (other organisms), and their IDs are listed in table S4. Sequences were aligned using the PROMALS3D server (*100*). Input sequence alignment for phylogenetic analysis was performed with MUSCLE (*101*). The phylogenetic tree was built using the neighbor-joining method and visualized with iTOL v5 (*102*).

Worms' brood size, body parameters, and locomotion analyses

For brood size analysis, four individual L4 larvae per replicate were placed onto single 35-mm NGM plates seeded with *E. coli* HB101 and were allowed to lay eggs at 20°C. Worms were transferred to fresh plates every 12 hours until they no longer produced embryos. Eggs were counted after the adult was moved. For each strain, 10 worms have been analyzed in two independent trials (two-tailed unpaired *t* test). For worm's body and movement analysis, eight age-synchronized young adult worms per strain were placed onto 35-mm NGM plates seeded with *E. coli*, and the worm movement was recorded for 2 min using the WormLab system (MBF Bioscience). The frame rate, exposure time, and gain were set to 7.5 frames/s, 0.0031 s, and 1, respectively. The worms' body length and width, track length, center point speed, and the overall track pattern of individual worms were analyzed using the WormLab software (MBF Bioscience). For each strain, 80 worms have been analyzed, and data were compared using the two-tailed unpaired Student's *t* test with Welch's correction and presented as mean values ± SD. A $P < 0.05$ was considered significantly different from control: ns, not significant; $^{*}P < 0.05$, $^{**}P < 0.001$, and $^{***}P < 0.0001$.

Microscopic analysis

Worms were immobilized with tetramisole, placed on slides coated with 2% agarose, and immediately imaged. The confocal microscopy for Fig. 1 and fig. S1 was performed using an FV1000 system with a 60×/1.2 water immersion lens (Olympus). Images were processed using Fiji/ImageJ software (version 2.0.0-rc69/1.52p) (*103*). For co-localization analysis presented in Fig. 6, worms were imaged using Zeiss LSM800 confocal microscope with 40×/1.2 water immersion apochromatic objective. Z-stack images were processed using Imaris 8.3 software. Median filter with a 3 × 3 × 1 kernel was applied to remove noise. Analysis was restricted to cells expressing CemOrange2–TRAM-1 ER protein. Gating with polygon was used to exclude the strongest unspecific signal from the green channel (Fig. 6, marked with asterisks). Pearson correlation coefficient was calculated with the ImarisColoc module based on three-dimensional data obtained from 18 worms.

Tethering assay

Tethering assays were performed as previously described (*21*, *104*). 293T cells (American Type Culture Collection, CRL-3216) were cultured in Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 10% FBS (Gibco) and penicillin (100 U/ml)/streptomycin (0.1 mg/ml) (Sigma-Aldrich) at 37°C in 5% CO_2. Cells were seeded into six-well plates and allowed to grow until about 70 to 80% confluence. Next, cells were cotransfected with 0.1 μg of pRL-5Box plasmid carrying an *RL* (*104*) and 2 μg of plasmid encoding tethered wild-type or catalytically inactive NHA–TENT-5 using 5 μl of Lipofectamine 2000 and Opti-MEM medium (Thermo Fisher Scientific, 31985047) according to the manufacturer's instructions. Cells were collected 24 hours after transfection for RNA (Northern blot and DRS) and protein level analyses.

C. elegans cultures for RNA analysis

Animal populations were synchronized by bleaching of the gravid adults and starvation of L1 larvae for 16 hours. Synchronized worms were grown on NGM plates seeded with *E. coli* HB101 at indicated temperatures until they reached the L4 stage. Worms were washed three times with 50 mM NaCl (900*g* for 2 min at room temperature) and resuspended in 1 ml of TRI Reagent (Sigma-Aldrich, T9424). Samples were incubated for 5 min at room temperature and stored at −80°C. For RNA analysis after *C. elegans* infection by *S. aureus*, synchronized worms were grown on NGM plates seeded with *E. coli* HB101 at 25°C until they reached the L4 stage. The infection plates were prepared as described (*37*). Briefly, TSA plates with Nal (10 μg/ml) were prepared 1 week before the experiment and stored at 4°C in the dark. *S. aureus* was grown in TSB + Nal overnight at 37°C. Five hundred microliters of the overnight *S. aureus* culture was uniformly spread onto the entire surface of 100-mm TSA + Nal plates and incubated at 37°C for 6 hours. L4 worms were washed three times

with sterile 50 mM NaCl and seeded onto infection TSA plates that were previously warmed to room temperature. After 8 hours of infection at 25°C, animals were washed off the plates and resuspended in TRI Reagent as described above.

RNA extraction

Total RNA was isolated with TRI Reagent according to the manufacturer's instructions (Sigma-Aldrich, T9424). To ensure the highest purity of the RNA samples isolated from worms, the subsequent phenol/chloroform extraction has been performed according to standard protocols. Before RT-qPCR, RNA-seq and DRS library preparation, RNase H treatment, and PAT total RNA was treated with TURBO DNase (Thermo Fisher Scientific). RNA was then purified by phenol/chloroform extraction and ethanol precipitation.

RNA-seq and data analysis

C. elegans cultures and RNA extraction are described above. Three independent replicate sample sets were prepared for each strain [wild type and *tent-5(tm3504)*] and condition (worms that were grown on *E. coli* HB101 or infected by *S. aureus* for 8 hours). Two micrograms (worms) or 1 μg (wild-type BMDMs stimulated with LPS) of DNase-treated total RNA was used for the library preparation. Ribosomal RNA (rRNA) was removed using a Ribo-Zero rRNA removal kit (human/mouse/rat; Epicentre, RZG1224). Sequencing libraries were prepared using a KAPA stranded RNA-seq library preparation kit (KAPA Biosystems, KR0934), and their quality was assessed with an Agilent 2100 Bioanalyzer (Agilent Technologies Inc.). The libraries were sequenced in the 75-nt single-end (*C. elegans* samples) or 75-nt pair-end (BMDMs) mode on the NextSeq500 Illumina platform. RNA-seq reads were adapter-clipped and quality-filtered with cutadapt (version 1.18) to remove adapters, low-quality fragments (minimum quality score was set to 20), and too short sequences (threshold set to 30 nt) (*105*). Quality-filtered reads were mapped to the respective reference genomes of *C. elegans* (WBCel235; ENSEMBL, release 94) or mouse (GRCm38; ENSEMBL, release 94) using the STAR (Spliced Transcripts Alignment to a Reference) aligner (version 2.6.1b or version 2.7.6a for worm and mouse, respectively) (*106*). Read counts were assigned to genes using featureCounts from the Subread package (version 1.6.3) with options -Q 10 -p -B -C -s 2 -g gene_id -t exon and respective annotation files for *C. elegans* (WBCel235; ENSEMBL, release 94) or mouse (Gencode vM25) (*107*). Multimappers and reads overlapping multiple features were not counted. Differential expression analysis was performed with DESeq2 (version 1.22) Bioconductor package (*108*) with default settings. For *C. elegans*, most of the analyses were performed for the genes, which expression was down-regulated at least 1.5-fold [$\log_2$ fold change < $-\log_2(1.5)$, FDR < 0.05] in mutant worms. Venn diagrams were drawn with VENNY (www.stefanjol.nl/venny). Gene sets were submitted for GO enrichment analysis to the WormBase Enrichment Suite (WS278) (*109*) and WormCat tool (*110*).

DRS and data analysis

For tethering assay, technical replicate sample sets were prepared from 293T cells transfected with the wild-type or catalytically inactive NHA–TENT-5. For *C. elegans* DRS, two independent replicate sample sets were prepared for *tent-5(tm3504)* and wild-type worms [the same input RNA samples as for the RNA-seq experiment (replicates 1 and 2) were used]. For mice samples, BMDMs from *Tent5a*$^{Flox/Flox}$ *Tent5c*$^{-/-}$ and wild-type mice were isolated, cultured, and transduced as described above. On the 14th day after isolation, cells were stimulated with LPS (100 ng/ml) for 8 hours. Two replicate sample sets were prepared from BMDMs for DRS analysis. Total RNA from 293T, *C. elegans*, and BMDM samples was isolated with TRI Reagent according to the manufacturer's instructions (Sigma-Aldrich, T9424). The cap-enriched mRNA was prepared from 100 μg of total RNA with GST-eIF4E^{K119A} protein (homemade) and glutathione sepharose 4B (GE Healthcare, 17-0756-01) as described previously (*23*). Nanopore direct RNA libraries were prepared with a DRS Kit [Oxford Nanopore Technologies (ONT), SQK-RNA002] from 3 μg (worms and 293T) and 3.5 μg (BMDMs) of cap-enriched mRNA mixed with 150 ng of *Saccharomyces cerevisiae* oligothymidilate [oligo(dT)]–enriched mRNA to optimize sequencing efficiency. Sequencing was performed with a MinION device (ONT; Flow cell type FLO-MIN106, RevD). Raw reads were basecalled with the standalone version of Guppy 4.0.11 (ONT). Sequencing reads were mapped to Gencode v36 supplemented with sequences of reporter transcripts (293T), WBCel235 (worms), or Gencode vM26 (BMDMs) reference transcriptomes using MiniMap 2.17 (*111*) with options -k 14 -ax map-ont –secondary = no and processed with samtools 1.9 to filter out supplementary alignments and reads mapping to reverse strand (samtools view -b -F 2320). The poly(A) tail lengths were estimated with Nanopolish (version 0.13.2) polya function (*112*). In subsequent analyses, only length estimates with quality control tag reported by Nanopolish as PASS were considered. Statistical analysis was performed using functions provided in the NanoTail R package (*23*). Poly(A) length distributions in analyzed conditions were compared using the Wilcoxon test, filtering out transcripts that had a low number of supporting reads under each condition (<10). Collected *P* values were adjusted for multiple comparisons using the Benjamini-Hochberg method. Transcripts were considered as having a significant change in poly(A) tail length if the adjusted *P* value was <0.05. Transcripts were considered as TENT-5 or TENT5A/C substrates if, in addition to being significantly changed, their median poly(A) tail length was at least 5 nt shorter in the mutant worms or double *Tent5a*$^{Flox/Flox}$ *Tent5c*$^{-/-}$ mutant compared to wild-type worms or mouse BMDMs, respectively. For differential expression estimates, reads were mapped to *C. elegans* (WBCel235; ENSEMBL, release 94) or mouse (GRCm38; ENSEMBL, release 94) reference genomes using MiniMap 2.17 (*111*), with options -ax splice –secondary = no -uf. Read counts were assigned to genes using featureCounts from the Subread package (version 2.0.1) with options -L –fracOverlap 0.5 –fracOverlapFeature 0.2 -s 1 and respective annotation files for *C. elegans* (WBCel235; ENSEMBL, release 94) or mouse (Gencode vM25) (*107*). Multimappers and reads overlapping multiple features were not counted. Differential expression analysis was performed with DESeq2 (version 1.28) Bioconductor package (*108*) with default settings. Gene sets were submitted for GO enrichment analysis to the WormBase Enrichment Suite (WS278) (*109*) or g:Profiler (*113*).

Transcript and UTR length analysis

Data regarding coordinates of the 5′ and 3′UTRs in the WBCel235 genome and percent guanine-cytosine content for each gene were downloaded from ParaSite BioMart (WS276) (https://parasite.wormbase.org/biomart/martview/). Data regarding transcript and coding sequences lengths were obtained from BioMart using biomaRt R package (biomart = "ensembl" and dataset = "celegans_gene_ensembl"). For each gene, only the longest possible 5′ and 3′UTRs or coding sequence was considered for the analysis. The

lengths of the 5′ and 3′UTRs were calculated on the basis of obtained coordinates. Statistics were calculated using the Wilcoxon test.

Motif enrichment analysis

Data regarding coordinates of the 3′UTRs in the WBCel235 genome were downloaded from ParaSite BioMart (WS276) (https://parasite.wormbase.org/biomart/martview/). Only the longest possible 3′UTR sequence for each gene was considered for the analysis. Coordinates of 3′UTRs of TENT-5 substrates (DRS) and genes that expression levels were down-regulated at least 1.5-fold in *tent-5(tm3505)* mutant worms compared to wild type (RNA-seq), as well as coordinates of 3′UTRs of all remaining genes identified by DRS or RNA-seq (background), were saved as bed file and were used for respective FASTA sequences collection using the bedtools getfasta tool (version 2.29.2) (*114*) and WBCel235 (ENSEMBL, release 94) genome sequence. FASTA sequences of 3′UTRs of TENT5A/C substrates and all Gencode-annotated transcripts in mm10 genome (background) were obtained with bedtools getfasta tool (version 2.29.2) (*114*), using bed files with 3′UTR coordinates downloaded from the UCSC Browser Table tool (Gencode vM23 track and known_gene table) and GRCm38 genome sequence. Sequence motifs were searched using the DREME tool (version 5.3.0) (*59*) with options -rna -norc -k 8 -l with the respective background (described above) specified.

Reverse transcription quantitative polymerase chain reaction

One microgram of the DNase-treated total RNA was reverse-transcribed with 1 μl of oligo$(dT)_{25}$ and random primers mix (50 mM and 50 ng/μl, respectively) using the SuperScript III Reverse Transcriptase (Thermo Fisher Scientific, 18080085). cDNA samples were diluted 10× and used for RT-qPCR analysis using the LightCycler 480 SYBR Green I Master Mix (Roche, 04887352001) and 0.25 μM primers on the LightCycler 480 Instrument (Roche). Primers were designed with Primer-BLAST (National Center for Biotechnology Information) to be exon-junction spanning where possible and tested for amplification efficiencies with a series of template dilutions. Each experimental replicate was measured in technical triplicate. Expression levels for each sample were normalized to *act-1*. Gene expression changes were calculated using the $2^{-\Delta\Delta C(t)}$ method. Unpaired two-sample two-sided *t* test using ΔC_t values were performed for most comparisons except for induction of gene expression during infection, where one-sided *t* tests were performed. A $P < 0.05$ was considered statistically significant: ns, not significant; $^*P \leq 0.05$, $^{**}P \leq 0.01$, and $^{***}P \leq 0.001$. Primer sequences are listed in table S2.

RNase H treatment

Twenty micrograms of DNase-treated total RNA was mixed with 2 μl of 50 mM oligo$(dT)_{25}$, 1 μl 10× hybridization buffer [25 mM tris-HCl (pH 7.5), 1 mM EDTA, and 50 mM NaCl], and water in 10 μl. RNA was denatured at 70°C for 10 min and slowly cooled down to 42°C. Next, 10 μl of prewarmed to 37°C 2× reaction buffer [40 mM tris-HCl (pH 7.5), 20 mM $MgCl_2$, 200 mM KCl, 2 mM DTT, and 10% sucrose] and 1 μl of RNase H (2 U; Thermo Fisher Scientific, EN0201) were added, and the reactions were carried out at 37°C for 1 hour. RNA was recovered with phenol/chloroform extraction, precipitated with 96% ethanol, and analyzed using Northern blots or PAT.

Poly(A) tail analysis

One microgram of DNA-free total RNA was ligated with 125 pmol of RA3_15N 3′ adaptor at 18°C for 16 hours in 20-μl mixtures containing 1× T4 RNA ligase buffer, 10% PEG 8000 (polyethylene glycol, molecular weight 8000), 50 U of RiboLock RNase inhibitor, and 300 U of T4 RNA ligase 2 truncated KQ [New England Biolabs (NEB), M0373L]. The samples were purified with AMPure XP magnetic beads [1:0.75 (v/v) RNA-to-beads ratio] to discard nonligated adapters and short RNA fragments. RNA was eluted in 15 μl of RNase-free water and reverse-transcribed using 200 U of the SuperScript III Reverse Transcriptase (Thermo Fisher Scientific) and 100 pmol of the RPI PCR Index Primer (TruSeq Illumina) according to the manufacturer's protocol. The cDNAs were purified using AMPure XP beads [1:1 (v/v) ratio], eluted with 20 μl of RNase-free water, and used for the nested PCR. Briefly, 1 μl of the cDNA was used for PCR-1 (Phusion Hot Start II; 25 cycles), with a gene-specific forward primer and a universal reverse primer RPuni. The PCR-1 samples were diluted 100× and used as a template for PCR-2, with a second forward gene-specific primer and RPuni. The PCR-1 and PCR-2 amplicons were analyzed in the 2% agarose gels in 1× TBE buffer (90 mM tris-borate and 2 mM EDTA). All primers used for PAT are described in table S2.

Northern blotting

High-molecular weight RNA samples were separated on 1.2% agarose gels containing 1.7% formaldehyde in 1× NBC buffer (50 mM boric acid, 1 mM sodium acetate, and 5 mM NaOH) and transferred to Hybond-N+ membranes (GE Healthcare) by overnight capillary elution using 8× SSC buffer (1.2 M NaCl and 120 mM sodium citrate). Low–molecular weight RNA samples were separated on 4 to 6% acrylamide gels containing 7 M urea in 0.5× TBE buffer (45 mM tris-borate and 1 mM EDTA) and electrotransferred to membranes in 0.5× TBE buffer at 300 to 350 mA at 4°C for 3 hours. RNA was immobilized on membranes by 254-nm UV light using a CL-1000 cross-linker (UVP) with the auto cross-link function (120 mJ/cm^2). Next, membranes were stained with 0.03% methylene blue in 0.3 M NaAc (pH 5.3), and staining was digitized. Random primed *RL* probes were PCR-amplified with primers RL_Fw and RL_Rev using pRL-5Box plasmid as a template and radioactively labeled with 20 μCi of [α-^{32}P]-dATP and the DECAprime II DNA Labeling Kit (Thermo Fisher Scientific, AM1456). Membranes were prehybridized in the PerfectHyb Plus Hybridization Buffer (Sigma-Aldrich) at 65°C for 30 min and incubated with probes in PerfectHyb buffer at 65°C overnight with rotation. Membranes were washed three times in prewarmed 0.5× SSC with 0.1% SDS at 65°C for 20 min and then exposed overnight to PhosphorImager screens (Fujifilm). The screens were scanned with a Typhoon FLA 7000 scanner (GE Healthcare) and analyzed with Multi Gauge software version 2.0 (Fujifilm).

Subcellular protein fractionation

Mixed-stage worm populations were grown on NGM plates seeded with *E. coli* HB101 at 20°C. Two independent replicate sample sets were prepared for each strain. Subcellular protein fractionation was performed using the Subcellular Protein Fractionation Kit (Thermo Fisher Scientific, 78840). Briefly, all buffers were supplemented with protease inhibitors (Invitrogen), and the entire procedure was performed at 4°C. For each strain, worms were washed from three 100-mm plates and washed three times with 50 mM NaCl (900*g* for 2 min at room temperature). Worms were resuspended in ice-cold 1.5 ml of cytoplasmic extraction buffer and lysed using Omni tissue homogenizer for 1 min following 10 min of incubation with

gentle mixing. Lysates were centrifuged at 500*g*, and the supernatant was collected as a cytoplasmic fraction, while the pellet was resuspended in membrane extraction buffer. After 10 min of incubation, the sample was centrifuged at 3000*g*, and the supernatant was collected as a membrane fraction. Last, the pellet was resuspended in nuclear extraction buffer, incubated for 30 min, and centrifuged at 5000*g*. The supernatant was collected as a nuclear fraction. The concentration of protein was measured by Bradford assay, and samples for SDS-PAGE and Western blot were prepared using the same amount of protein from each fraction. Control total protein input samples were prepared by boiling worms in 1× SDS sample buffer.

Total worm extract preparation for Superdex 200 chromatography

Protein extracts were prepared from mixed-stage worm populations grown at 20°C. Animals were washed three times with 50 mM NaCl and then briefly with lysis buffer [50 mM Hepes (pH 7.4), 100 mM NaCl, 3 mM $MgCl_2$, 0.5 mM DTT, 0.05% NP-40 substitute, and 10% (v/v) glycerol]. The supernatant was discarded, and the pellet was resuspended in 5 V of lysis buffer supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF), 1× chymostatin, and 1× protease inhibitors and drop-frozen in liquid nitrogen. Lysis was performed by grinding frozen worms in liquid nitrogen. The extract was allowed to melt on ice and supplemented with 1 mM PMSF, 1× chymostatin, and 1×protease inhibitors. The lysate was cleared by two centrifugation steps at 20,000*g* at 4°C for 20 min. Five hundred microliters of the lysate (7 μg/μl) was subjected to the Superdex 200 chromatography column equilibrated with lysis buffer. Two hundred microliters of each collected fraction was precipitated using methanol/chloroform; pellets were resuspended in 20 μl of 1× SDS sample buffer, boiled, and used for the SDS-PAGE and Western blot.

Western blotting

BMDMs from *Tent5a*$^{Flox/Flox}$ *Tent5c*$^{-/-}$, *Tent5a*-3xFLAG, *Tent5c*-3xFLAG, and wild-type mice were isolated, cultured, and transduced as described above. On the 13th day after isolation, cells were counted and seeded on six-well plates, with 0.5 million cells per well. The next day, cells were stimulated with LPS (100 ng/ml) for 8 hours (*Tent5a*$^{Flox/Flox}$ *Tent5c*$^{-/-}$ and wild type) or for 0 to 16 hours of time points (*Tent5a*-3xFLAG, *Tent5c*-3xFLAG, and wild type). BMDMs were scratched and pelleted by centrifugation for 3 min at 350*g*. Cells were lysed with 0.1% NP-40 in phosphate-buffered saline (PBS) supplemented with protease inhibitors and viscolase (final concentration of 0.1 U/ml; A&A Biotechnology). The samples were incubated at 37°C for 30 min with shaking before 3× SDS sample buffer [187.5 mM tris-HCl (pH 6.8), 6% SDS, 150 mM DTT, 0.02% bromophenol blue, 30% glycerol, and 3% 2-mercaptoethanol] was added, and the lysates were boiled for 10 min. Lysates from the 293T cells following tethering assay were prepared using the same protocol. Protein samples from a mixed population of worms were prepared by boiling ~100 worms in 3× SDS sample buffer for 5 min. Boiled protein mixtures from cells or worms were cleared by centrifugation for 5 min at maximum speed at room temperature and resolved on 10 to 12% SDS-PAGE gels. Proteins were transferred to Protran nitrocellulose membranes (GE Healthcare) by wet transfer at 300 mA at 4°C for 2 hours in 1× transfer buffer [25 mM tris base, 192 mM glycine, and 20% methanol (v/v)]. After transfer, the membranes were stained with 0.3% (w/v) Ponceau S in 3% (v/v) acetic acid, and the staining was digitized. Next, membranes were soaked in 5% (w/v) nonfat milk in 1× TBS-T (20 mM tris base, 150 mM NaCl, and 0.01% Tween 20) for 1 hour with gentle agitation at room temperature, followed by the overnight incubation at 4°C with specific primary antibodies diluted 1:3000 (anti-*RL* antibody, clone 5B11.2; Millipore, MAB4400), 1:3000 (GFP B-2; Santa Cruz Biotechnology, sc-9996), 1:1000 (FLAG; Proteintech, 20543-1-AP), 1:10,000 RFP (red fluorescent protein); Erdogan, AB233), 1:5000 [glyceraldehyde phosphate dehydrogenase (GAPDH); Proteintech, 10494-1-AP], 1:10,000 (α-tubulin, clone DM1A; Millipore, MABT205), 1:1000 (iNOS; Cell Signaling Technology, 13120), 1:1000 (CD80; Cell Signaling Technology, 54521), 1:30,000 (lysozyme; Abcam, ab108508), 1:30,000 (cathepsin S; Invitrogen, MA5-29695), 1:10,000 (cathepsin B; Abcam, ab214428), and 1:20,000 (cathepsin D; Abcam, ab75852). Membranes were washed three times for 10 min each in 1× TBS-T and then incubated for 2 hours with gentle agitation at room temperature with horseradish peroxidase–conjugated secondary anti-mouse (Millipore, 401215) or anti-rabbit (Millipore, 401393) antibodies diluted 1:5000. Following three washes in 1× TBS-T, blots were incubated with the Clarity Western ECL Substrate (Bio-Rad) for 1 to 3 min, and signals were detected either through exposure to a CL-Exposure film (Thermo Fisher Scientific) and developed in an AGFA Curix CP-1000 device or visualized using the ChemiDoc Imaging System (Bio-Rad). Protein bands from Western blots were quantified with ImageJ as described in www.yorku.ca/yisheng/Internal/Protocols/ImageJ.pdf. Final relative quantification values represent the ratio of net band intensity from the protein of interest to net GAPDH (loading control).

Mass spectrometry

Protein extracts were prepared in eight replicate sample sets. The *tent-5(tm3504)* mutant and wild-type worms were grown at 20°C on *E. coli* HB101 until they reached the L4 stage. Worms were washed three times with 50 mM NaCl and once in 1× lysis buffer [50 mM Hepes (pH 7.4), 100 mM NaCl, 3 mM $MgCl_2$, 0.5 mM DTT, 0.05% NP-40 substitute, and 10% (v/v) glycerol]. Pellets were resuspended in 900 μl of 1× lysis buffer supplemented with 1 mM PMSF, 1× chymostatin, and 1× protease inhibitors; transferred to 2-ml tubes containing 200 μl of Zirconia beads (BioSpec Products); and drop-frozen in liquid nitrogen. Next, tubes were inserted into the Fast Prep-24 machine (MP Biomedicals), and worms were crushed for 1 min at maximum speed. Following centrifugation at 14,000*g* for 10 min at 4°C, lysates were transferred to the new tubes and subjected to sonication at high amplitude for 20 min (30-s on/30-s off cycle) (Diagenode Bioruptor XL) and then cleared by centrifugation at 14,000*g* for 30 min at 4°C. A Millipore Direct Detect infrared spectrometer was used to determine the total protein concentration of the lysate. Sample preparation was done on the basis of modified FASP (Filter-aided sample preparation) protocol (*115*). Briefly, a supernatant was placed at Vivacon 30-kDa filter (Sartorius), centrifuged, and washed three times with 200 μl of 8 M urea in 100 mM NH_4HCO_3. Next, samples were reduced (DTT at room temperature for 30 min) and alkylated (indole-3-acetic acid at room temperature for 15 min) following overnight digestion with trypsin (Promega) and acidified with trifluoroacetic acid to a final concertation of 0.1%. Mass spectrometry (MS) analysis was performed by liquid chromatography–MS in the Laboratory of Mass Spectrometry (Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw) using a nanoACQUITY UPLC system (Waters, 176016000) coupled to an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher

Scientific). Peptides were separated by a 180-min linear gradient of 95% solution A (0.1% formic acid in water) to 35% solution B (acetonitrile and 0.1% formic acid). The measurement of each sample was preceded by three washing runs to avoid cross-contamination; the final MS washing run was searched for the presence of cross-contamination between samples. If the protein of interest was identified in the washing run and the next measured sample at the same or smaller intensity, then the sample was regarded as contaminated and excluded from the final graphs. The mass spectrometer was operated in the data-dependent MS-MS2 mode, and data were acquired in the mass/charge ratio range of 300 to 2000. MS raw data files were used to calculate protein abundance in the samples using the MaxQuant (version 1.6.3.4) platform (*116*). The reference proteome of *C. elegans* database from UniProt (27,805 protein entries) and common contaminates list included in MaxQuant were used, and analysis was performed with the following settings: match between runs, variable modification: oxidation (M), and 4.5 parts per million of error tolerance. Label-free quantification (LFQ) intensity values were calculated using the MaxLFQ algorithm to estimate quantities of identified proteins. Protein abundance was defined as the LFQ value calculated by MaxQuant software for a protein (sum of intensities of identified peptides of a given protein) divided by its molecular weight. The Scafold4 Q + S platform was used for statistical analysis. Protein abundance in analyzed samples was compared using the Mann-Whitney test with Benjamini-Hochberg correction.

Life-span and killing assays

The 35-mm plates with TSA + Nal (final concentration of 10 μg/ml) for killing assays on *S. aureus* NCTC8325 were prepared 1 week before the experiments and stored at 4°C protected from light (*37*). Survival analysis on *S. marcescens* Db10, *P. luminescens*, and respective control *E. coli* OP50 were performed on the 35-mm plates with NGM without the addition of drugs or antibiotics. Life-span assays on *E. coli* HB101, *P. aeruginosa* PAO1, and some assays on *E. coli* OP50 were performed with the addition of FUdR (5-Fluoro-2′-deoxyuridine) (final concentration of 0.1 mg/ml; Sigma-Aldrich, F0503) to prevent progeny production. All bacteria strains were freshly seeded from the −80°C stock 3 days before an experiment to the appropriate solid medium. *E. coli* OP50 was cultured overnight in LB, and 50 μl of overnight culture was spread onto the center of 35-mm NGM plates and incubated at 37°C for 16 hours. *E. coli* HB101 was cultured overnight in LB + streptomycin (0.1 mg/ml). Fifty microliters of overnight culture was spread onto the center of 35-mm NGM or NGM + FUdR plates (depending on the experiment setup) and incubated at 37°C for 16 hours. For survival analysis on UV-killed *E. coli* HB101, NGM + FUdR plates seeded with bacteria were exposed to UV light in a UV Stratalinker 2400 for 30 min at maximum. For survival analysis on heat-killed bacteria, *E. coli* HB101 and OP50 were cultured overnight in 50 ml of LB and centrifuged at 7000*g* for 10 min at room temperature, and pellets were resuspended in 5 ml of fresh LB. Next, to kill bacteria, mixtures were incubated in the water bath at 65°C for 30 min. Fifty microliters of culture containing dead bacteria was spread onto the center of 35-mm plates and incubated at 25°C for 24 hours. Bacterial killing was evaluated by inoculating LB medium with UV- or heat-treated bacteria, and lack of growth at 37°C confirmed effective killing. *S. aureus* was grown overnight in TSB + Nal (10 μg/ml). Ten microliters of overnight culture was spread onto the center of 35-mm TSA + Nal plates and incubated at 37°C for 6 hours and then cooled down to 25°C and used for the killing assays. *P. aeruginosa* and *S. marcescens* were grown overnight at 37°C in LB. Ten microliters of overnight cultures was spread onto the center of 35-mm NGM or NGM + FUdR plates (for *S. marcescens* and *P. aeruginosa*, respectively) and incubated at 37°C for 24 hours and then at 25°C for another 24 hours. *P. luminescens* was cultured overnight at 30°C in LB, and 10 μl of the overnight culture was spread onto the center of 35-mm NGM plates and incubated at 30°C for 24 hours and after that at 25°C for 24 hours. All worms were grown on NGM + *E. coli* HB101 at 20° or 25°C (depending on the temperature in which the assays were performed) for three to four generations before the experiments. Animal populations were synchronized by bleaching of the gravid adults and the starvation of L1 larvae for 16 hours at appropriate temperatures. Forty to 60 L4-staged worms were transferred to each of the three replicate assay plates per strain. Beginning on the next day, the number of dead and live worms on each plate was recorded daily (*S. marcescens*, *P. aeruginosa*, *P. luminescens*, and *E. coli*) or twice a day (*S. aureus*). Live worms were transferred daily to new plates to avoid contamination with the progeny (*S. marcescens*, *P. luminescens*, and related *E. coli* OP50 control). For *P. aeruginosa* tests, worms were not transferred to the new plates. For *E. coli* HB101 life spans performed on the NGM + FUdR plates, worms were transferred to the new plates every 3 days until day 12 and then left on the same plates. Worms that left the plates in the first several days of the assay were removed from the counts of subsequent days. Animals were considered dead if they failed to respond to a gentle touch. For each survival experiment, at least two biological replicates were carried out. Kaplan-Meier survival analyses were performed using GraphPad Prism 7 software. Life-span survival data were compared using the log-rank significance test and presented as median survival. A $P < 0.05$ was considered significantly different from control: ns, not significant; $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$.

CFU assays

Exposure of wild-type and mutant worms to *E. coli* HB101 and *P. aeruginosa* PAO1 was carried out exactly as for life-span and killing assays described above. CFU assays were performed essentially as described in (*49*, *50*) with minor modifications. Briefly, on the 5th (HB101) and 4th (PAO1) day of adulthood, 30 worms of each genotype from each of the three technical replicate plates were collected into 50 μl of 1× M9 supplemented with 25 mM levamisole (Sigma-Aldrich, L9756) to inhibit pharyngeal pumping and expulsion. Worms were washed in 1× M9 + 25 mM levamisole three times and then surface-sterilized in 1× M9 + 25 mM levamisole + kanamycin (100 μg/ml) for 45 min at room temperature. Following three washes with 1× M9 + 25 mM levamisole, worms were resuspended in 150 μl of PBS containing 0.1% Triton X-100. A 100-μl aliquot of the supernatant was removed from each replicate to test for external bacterial contamination. Animals were homogenized in the remaining 50 μl of PBS + 0.1% Triton X-100 with pellet pestle (Bel-Art, BAF199230001) and motor for 30 s, and then 450 μl of PBS was added to each sample. Dilution series of homogenates was spread to the LB plates without antibiotics and grown at 37°C for 24 hours. CFU value per worm was counted as follows: number of CFU/worm = (number of colonies × dilution factor)/number of worms in lysate − external CFU. For each CFU experiment, three biological replicates were carried out, each comprising at least two technical replicates. A $P < 0.05$ was considered significantly different from control: ns, not significant; $^{*}P < 0.05$ and $^{**}P < 0.01$.

Statistical analysis

Statistical analysis was performed in Microsoft Excel (RT-qPCR), GraphPad Prism 7 (life span assays), or with R 4.0 (*117*). Details of the particular statistical analyses, significance, number of replicates and sample sizes, and the features of all plots are described in the figure legends. Data plotted as box plots have the following features: whiskers (25th and 75th percentiles), minima and maxima (5th and 95th percentiles), and thick lines (median). Data presented as heatmaps were normalized to a sequencing depth using DESeq2 and transformed with regularized log transformation for visualization purposes.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.add9468

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. J. A. Hoffmann, F. C. Kafatos, C. A. Janeway, R. A. B. Ezekowitz, Phylogenetic perspectives in innate immunity. *Science* **284**, 1313–1318 (1999).
2. A. Iwasaki, R. Medzhitov, Control of adaptive immunity by the innate immune system. *Nat. Immunol.* **16**, 343–353 (2015).
3. K. Buchmann, Evolution of innate immunity: Clues from invertebrates via fish to mammals. *Front. Immunol.* **5**, 1–8 (2014).
4. D. H. Kim, J. J. Ewbank, Signaling in the innate immune response. *WormBook* **2018**, 1–35 (2018).
5. J. E. Irazoqui, J. M. Urbach, F. M. Ausubel, Evolution of host innate defence: Insights from *Caenorhabditis elegans* and primitive invertebrates. *Nat. Rev. Immunol.* **10**, 47–58 (2010).
6. G. Gasteiger, A. D'osualdo, D. A. Schubert, A. Weber, E. M. Bruscia, D. Hartl, Cellular innate immunity: An old game with new players. *J. Innate Immun.* **9**, 111–125 (2017).
7. S. Carpenter, E. P. Ricci, B. C. Mercier, M. J. Moore, K. A. Fitzgerald, Post-transcriptional regulation of gene expression in innate immunity. *Nat. Rev. Immunol.* **14**, 361–376 (2014).
8. Y. Shi, J. L. Manley, The end of the message: Multiple protein–RNA interactions define the mRNA polyadenylation site. *Genes Dev.* **29**, 889–897 (2015).
9. J. Neve, R. Patel, Z. Wang, A. Louey, A. M. Furger, Cleavage and polyadenylation: Ending the message expands gene regulation. *RNA Biol.* **14**, 865–890 (2017).
10. A. L. Jalkanen, S. J. Coleman, J. Wilusz, Determinants and implications of mRNA poly(A) tail size—Does this protein make my tail look big? *Semin. Cell Dev. Biol.* **34**, 24–32 (2014).
11. Y. Bin Yan, Deadenylation: Enzymes, regulation, and functional implications. *Wiley Interdiscip. Rev. RNA* **5**, 421–443 (2014).
12. G. Martin, W. Keller, RNA-specific ribonucleotidyl transferases. *RNA* **13**, 1834–1849 (2007).
13. S. Yu, V. N. Kim, A tale of non-canonical tails: gene regulation by post-transcriptional RNA tailing. *Nat. Rev. Mol. Cell Biol.* **21**, 542–556 (2020).
14. V. Liudkovska, A. Dziembowski, Functions and mechanisms of RNA tailing by metazoan terminal nucleotidyltransferases. *Wiley Interdiscip. Rev. RNA* **12**, e1622 (2021).
15. J. H. Kim, J. D. Richter, Opposing polymerase-deadenylase activities regulate cytoplasmic polyadenylation. *Mol. Cell* **24**, 173–183 (2006).
16. K. W. Kim, T. L. Wilson, J. Kimble, GLD-2/RNP-8 cytoplasmic poly(A) polymerase is a broad-spectrum regulator of the oogenesis program. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 17445–17450 (2010).
17. M. Nousch, A. Yeroslaviz, C. R. Eckmann, Stage-specific combinations of opposing poly(A) modifying enzymes guide gene expression during early oogenesis. *Nucleic Acids Res. Res.* **47**, 10881–10893 (2019).
18. E. K. Jae, E. Drier, S. A. Barbee, M. Ramaswami, J. C. P. Yin, M. Wickens, GLD2 poly(A) polymerase is required for long-term memory. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 14644–14649 (2008).
19. T. Udagawa, S. A. Swanger, K. Takeuchi, J. H. Kim, V. Nalavadi, J. Shin, L. J. Lorenz, R. S. Zukin, G. J. Bassell, J. D. Richter, Bidirectional control of mRNA translation and synaptic plasticity by the cytoplasmic polyadenylation complex. *Mol. Cell* **47**, 253–266 (2012).
20. K. Kuchta, A. Muszewska, L. Knizewski, K. Steczkiewicz, L. S. Wyrwicz, K. Pawlowski, L. Rychlewski, K. Ginalski, FAM46 proteins are novel eukaryotic non-canonical poly(A) polymerases. *Nucleic Acids Res.* **44**, 3534–3548 (2016).
21. S. Mroczek, J. Chlebowska, T. M. Kuliński, O. Gewartowska, J. Gruchota, D. Cysewski, V. Liudkovska, E. Borsuk, D. Nowis, A. Dziembowski, The non-canonical poly(A) polymerase *FAM46C* acts as an onco-suppressor in multiple myeloma. *Nat. Commun.* **8**, 619 (2017).
22. Y. X. Zhu, C. X. Shi, L. A. Bruins, P. Jedlowski, X. Wang, K. M. Kortüm, M. Luo, J. M. Ahmann, E. Braggio, A. K. Stewart, Loss of *FAM46C* promotes cell survival in myeloma. *Cancer Res.* **77**, 4317–4327 (2017).
23. A. Bilska, M. Kusio-kobia, P. S. Krawczyk, O. Gewartowska, B. Tarkowski, K. Koby, D. Nowis, J. Golab, J. Gruchota, E. Borsuk, A. Dziembowski, S. Mroczek, Immunoglobulin expression and the humoral immune response is regulated by the non-canonical poly(A) polymerase TENT5C. *Nat. Commun.* **11**, 1–17 (2020).
24. A. B. Herrero, D. Quwaider, L. A. Corchete, M. V. Mateos, R. García-Sanz, N. C. Gutiérrez, FAM46C controls antibody production by the polyadenylation of immunoglobulin mRNAs and inhibits cell migration in multiple myeloma. *J. Cell. Mol. Med.* **00**, 1–12 (2020).
25. J. L. Hu, H. Liang, H. Zhang, M. Z. Yang, W. Sun, P. Zhang, L. Luo, J. X. Feng, H. Bai, F. Liu, T. Zhang, J. Y. Yang, Q. Gao, Y. Long, X. Y. Ma, Y. Chen, Q. Zhong, B. Yu, S. Liao, Y. Wang, Y. Zhao, M. S. Zeng, N. Cao, J. Wang, W. Chen, H. T. Yang, S. Gao, FAM46B is a prokaryotic-like cytoplasmic poly(A) polymerase essential in human embryonic stem cells. *Nucleic Acids Res.* **48**, 2733–2748 (2020).
26. O. Gewartowska, G. Aranaz-Novaliches, P. S. Krawczyk, S. Mroczek, M. Kusio-Kobiałka, B. Tarkowski, F. Spoutil, O. Benada, O. Kofroňová, P. Szwedziak, D. Cysewski, J. Gruchota, M. Szpila, A. Chlebowski, R. Sedlacek, J. Prochazka, A. Dziembowski, Cytoplasmic polyadenylation by TENT5A is required for proper bone formation. *Cell Rep.* **35**, 109015 (2021).
27. J. Coller, M. Wickens, Tethered function assays: An adaptable approach to study RNA regulatory proteins. *Methods Enzymol.* **429**, 299–321 (2007).
28. E. R. Troemel, S. W. Chu, V. Reinke, S. S. Lee, F. M. Ausubel, D. H. Kim, p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans*. *PLOS Genet.* **2**, 1725–1739 (2006).
29. T. Roeder, M. Stanisak, C. Gelhaus, I. Bruchhaus, J. Grötzinger, M. Leippe, Caenopores are antimicrobial peptides in the nematode *Caenorhabditis elegans* instrumental in nutrition and immunity. *Dev. Comp. Immunol.* **34**, 203–209 (2010).
30. H. Schulenburg, M. P. Hoeppner, J. Weiner, E. Bornberg-Bauer, Specificity of the innate immune system and diversity of C-type lectin domain (CTLD) proteins in the nematode *Caenorhabditis elegans*. *Immunobiology* **213**, 237–250 (2008).
31. I. Engelmann, A. Griffon, L. Tichit, F. Montañana-Sanchis, G. Wang, V. Reinke, R. H. Waterston, L. D. W. Hillier, J. J. Ewbank, A comprehensive analysis of gene expression changes provoked by bacterial and fungal infection in *C. elegans*. *PLOS ONE* **6**, e19055 (2011).
32. C. Taffoni, N. Pujol, Mechanisms of innate immunity in *C. elegans* epidermis. *Tissue Barriers* **3**, 1–8 (2015).
33. J. H. Thomas, Concerted evolution of two novel protein families in *Caenorhabditis species*. *Genetics* **172**, 2269–2281 (2006).
34. D. Wong, D. Bazopoulou, N. Pujol, N. Tavernarakis, J. J. Ewbank, Genome-wide investigation reveals pathogen-specific and shared signatures in the response of *Caenorhabditis elegans* to infection. *Genome Biol.* **8**, R194 (2007).
35. G. V. Mallo, C. L. Kurz, C. Couillault, N. Pujol, S. Granjeaud, Y. Kohara, J. J. Ewbank, Inducible antibacterial defense system in *C. elegans*. *Curr. Biol.* **12**, 1209–1214 (2002).
36. C. T. Murphy, S. A. McCarroll, C. I. Bargmann, A. Fraser, R. S. Kamath, J. Ahringer, H. Li, C. Kenyon, Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. *Nature* **424**, 277–283 (2003).
37. O. Visvikis, N. Ihuegbu, S. A. Labed, L. G. Luhachack, A. M. F. Alves, A. C. Wollenberg, L. M. Stuart, G. D. Stormo, J. E. Irazoqui, Innate host defense requires TFEB-mediated transcription of cytoprotective and antimicrobial genes. *Immunity* **40**, 896–909 (2014).
38. A. Sinha, R. Rae, I. Iatsenko, R. J. Sommer, System wide analysis of the evolution of innate immunity in the nematode model species *Caenorhabditis elegans* and *Pristionchus pacificus*. *PLOS ONE* **7**, e44255 (2012).
39. S. Alper, S. J. McBride, B. Lackford, J. H. Freedman, D. A. Schwartz, Specificity and complexity of the *Caenorhabditis elegans* innate immune response. *Mol. Cell. Biol.* **27**, 5544–5553 (2007).
40. W. Yang, K. Dierking, P. C. Rosenstiel, H. Schulenburg, GATA transcription factor as a likely key regulator of the *Caenorhabditis elegans* innate immune response against gut pathogens. *Fortschr. Zool.* **119**, 244–253 (2016).
41. E. A. Evans, T. Kawli, M. W. Tan, Pseudomonas aeruginosa suppresses host immunity by activating the DAF-2 insulin-like signaling pathway in *Caenorhabditis elegans*. *PLOS Pathog.* **4**, e1000175 (2008).
42. R. P. Shivers, M. J. Youngman, D. H. Kim, Transcriptional responses to pathogens in *Caenorhabditis elegans*. *Curr. Opin. Microbiol.* **11**, 251–256 (2008).
43. X. X. Lin, I. Sen, G. E. Janssens, X. Zhou, B. R. Fonslow, D. Edgar, N. Stroustrup, P. Swoboda, J. R. Yates, G. Ruvkun, C. G. Riedel, DAF-16/FOXO and HLH-30/TFEB function as combinatorial transcription factors to promote stress resistance and longevity. *Nat. Commun.* **9**, 4400 (2018).
44. D. Gems, D. L. Riddle, Genetic, behavioral and environmental determinants of male longevity in *Caenorhabditis elegans*. *Genetics* **154**, 1597–1610 (2000).

45. D. Garigan, A. L. Hsu, A. G. Fraser, R. S. Kamath, J. Abringet, C. Kenyon, Genetic analysis of tissue aging in *Caenorhabditis elegans*: A role for heat-shock factor and bacterial proliferation. *Genetics* **161**, 1101–1112 (2002).
46. J. E. Irazoqui, E. R. Troemel, R. L. Feinbaum, L. G. Luhachack, B. O. Cezairliyan, F. M. Ausubel, Distinct pathogenesis and host responses during infection of *C. elegans* by *P. aeruginosa* and *S. aureus*. *PLOS Pathog.* **6**, 1–24 (2010).
47. C. L. Kurz, S. Chauvet, E. Andrès, M. Aurouze, I. Vallet, G. P. F. Michel, M. Uh, J. Celli, A. Filloux, S. De Bentzmann, I. Steinmetz, J. A. Hoffmann, B. B. Finlay, J.-P. Gorvel, D. Ferrandon, J. J. Ewbank, Virulence factors of the human opportunistic pathogen *Serratia marcescens* identified by in vivo screening. *EMBO J.* **22**, 1451–1460 (2003).
48. C. Couillault, J. J. Ewbank, Diverse bacteria are pathogens of *Caenorhabditis elegans*. *Infect. Immun.* **70**, 4705–4707 (2002).
49. M. F. Palominos, A. Calixto, Quantification of bacteria residing in *Caenorhabditis elegans* intestine. *Bio Protoc.* **10**, 1–12 (2020).
50. F. Rodriguez Ayala, S. Cogliati, C. Bauman, C. Leñini, M. Bartolini, J. Villalba, F. Argañaraz, R. Grau, Culturing bacteria from *Caenorhabditis elegans* gut to assess colonization proficiency. *Bio Protoc.* **7**, e2345 (2017).
51. M. E. Hoinville, A. C. Wollenberg, Changes in *Caenorhabditis elegans* gene expression following exposure to *Photorhabdus luminescens* strain TT01. *Dev. Comp. Immunol.* **82**, 165–176 (2018).
52. I. Legnini, J. Alles, N. Karaiskos, S. Ayoub, N. Rajewsky, FLAM-seq: Full-length mRNA sequencing reveals principles of poly(A) tail length control. *Nat. Methods* **16**, 879–886 (2019).
53. S. A. Lima, L. B. Chipman, A. L. Nicholson, Y. H. Chen, B. A. Yee, G. W. Yeo, J. Coller, A. E. Pasquinelli, Short poly(A) tails are a conserved feature of highly expressed genes. *Nat. Struct. Mol. Biol.* **24**, 1057–1063 (2017).
54. J. Tao, Y. Hao, X. Li, H. Yin, X. Nie, J. Zhang, B. Xu, Q. Chen, B. Li, Systematic identification of housekeeping genes possibly used as references in *Caenorhabditis elegans* by large-scale data integration. *Cells* **9**, 786 (2020).
55. C. Y. Ewald, J. N. Landis, J. P. Abate, C. T. Murphy, T. K. Blackwell, Dauer-independent insulin/IGF-1-signalling implicates collagen remodelling in longevity. *Nature* **519**, 97–101 (2015).
56. D. Sellegounder, Y. Liu, P. Wibisono, C. H. Chen, D. Leap, J. Sun, Neuronal GPCR NPR-8 regulates *C. elegans* defense against pathogen infection. *Sci. Adv.* **5**, eaaw4717 (2019).
57. Y. Liu, D. Martinez-Martinez, C. L. Essmann, M. R. Cruz, F. Cabreiro, D. A. Garsin, Transcriptome analysis of *Caenorhabditis elegans* lacking heme peroxidase SKPO-1 reveals an altered response to *Enterococcus faecalis*. *G3 (Bethesda)* **11**, jkaa055 (2021).
58. I. Gallotta, A. Sandhu, M. Peters, M. Haslbeck, R. Jung, S. Agilkaya, J. L. Blersch, C. Rödelsperger, W. Röseler, C. Huang, R. J. Sommer, D. C. David, Extracellular proteostasis prevents aggregation during pathogenic attack. *Nature* **584**, 410–414 (2020).
59. T. L. Bailey, DREME: Motif discovery in transcription factor ChIP-seq data. *Bioinformatics* **27**, 1653–1659 (2011).
60. J. Suh, H. Hutter, A survey of putative secreted and transmembrane proteins encoded in the *C. elegans* genome. *BMC Genomics* **13**, 333 (2012).
61. J. D. McGhee, The Caenorhabditis elegans intestine. *Wiley Interdiscip. Rev. Dev. Biol.* **2**, 347–367 (2013).
62. C. Coburn, D. Gems, The mysterious case of the *C. elegans* gut granule: Death anthranilic acid and the kynurenine pathway. *Front. Genet.* **4**, 151 (2013).
63. B. J. Thomas, I. E. Wight, W. Y. Y. Chou, M. Moreno, Z. Dawson, A. Homayouni, H. Huang, H. Kim, H. Jia, J. R. Buland, J. A. Wambach, F. S. Cole, S. C. Pak, G. A. Silverman, C. J. Luke, CemOrange2 fusions facilitate multifluorophore subcellular imaging in *C. elegans*. *PLOS ONE* **14**, 1–25 (2019).
64. C. Fucci, M. Resnati, E. Riva, T. Perini, E. Ruggieri, U. Orfanelli, F. Paradiso, F. Cremasco, A. Raimondi, E. Pasqualetto, M. Nuvolone, L. Rampoldi, S. Cenci, E. Milan, The interaction of the tumor suppressor FAM46C with p62 and FNDC3 proteins Integrates protein and secretory homeostasis. *Cell Rep.* **32**, 108162 (2020).
65. N. Manfrini, M. Mancino, A. Miluzio, S. Oliveto, M. Balestra, P. Calamita, R. Alfieri, R. L. Rossi, M. Sassoe-Pognetto, C. Salio, A. Cuomo, T. Bonaldi, M. Manfredi, E. Marengo, E. Ranzato, S. Martinotti, D. Cittaro, G. Tonon, S. Biffo, FAM46C and FNDC3A are multiple myeloma tumor suppressors that act in concert to impair clearing of protein aggregates and autophagy. *Cancer Res.* **80**, 4693–4706 (2020).
66. R. A. Zabinsky, B. M. Weum, M. Cui, M. Han, RNA binding protein vigilin collaborates with miRNAs to regulate gene expression for *Caenorhabditis elegans* larval development. *G3 (Bethesda)* **7**, 2511–2518 (2017).
67. Y. S. Ooi, K. Majzoub, R. A. Flynn, M. A. Mata, J. Diep, J. K. Li, N. van Buuren, N. Rumachik, A. G. Johnson, A. S. Puschnik, C. D. Marceau, L. Mlera, J. M. Grabowski, K. Kirkegaard, M. E. Bloom, P. Sarnow, C. R. Bertozzi, J. E. Carette, An RNA-centric dissection of host complexes controlling flavivirus infection. *Nat. Microbiol.* **4**, 2369–2382 (2019).
68. P. J. Murray, Macrophage polarization. *Annu. Rev. Physiol.* **79**, 541–566 (2017).
69. M. Benoit, B. Desnues, J.-L. Mege, Macrophage polarization in bacterial infections. *J. Immunol.* **181**, 3733–3739 (2008).
70. C. Zheng, Y. C. Ouyang, B. Jiang, X. Lin, J. Chen, M. Z. Dong, X. Zhuang, S. Yuan, Q. Y. Sun, C. Han, Non-canonical RNA polyadenylation polymerase FAM46C is essential for fastening sperm head and flagellum in mice. *Biol. Reprod.* **100**, 1673–1685 (2019).
71. T. Ganz, V. Gabayan, H.-I. Liao, L. Liu, A. Oren, T. Graf, A. M. Cole, Increased inflammation in lysozyme M-deficient mice in response to *Micrococcus luteus* and its peptidoglycan. *Blood* **101**, 2388–2392 (2003).
72. P. Markart, T. R. Korfhagen, T. E. Weaver, H. T. Akinbi, Mouse lysozyme M is important in pulmonary host defense against *Klebsiella pneumoniae* infection. *Am. J. Respir. Crit. Care Med.* **169**, 454–458 (2004).
73. H. T. Akinbi, R. Epaud, H. Bhatt, T. E. Weaver, Bacterial killing is enhanced by expression of lysozyme in the lungs of transgenic mice. *J. Immunol.* **165**, 5760–5766 (2000).
74. A. M. Cole, D. R. Thapa, V. Gabayan, H.-I. Liao, L. Liu, T. Ganz, Decreased clearance of *Pseudomonas aeruginosa* from airways of mice deficient in lysozyme M. *J. Leukoc. Biol.* **78**, 1081–1085 (2005).
75. E. del Cerro-Vadillo, F. Madrazo-Toca, E. Carrasco-Marín, L. Fernandez-Prieto, C. Beck, F. Leyva-Cobián, P. Saftig, C. Alvarez-Dominguez, Cutting edge: A novel nonoxidative phagosomal mechanism exerted by cathepsin-D controls *Listeria monocytogene*s intracellular growth. *J. Immunol.* **176**, 1321–1325 (2006).
76. Q. Fu, J. Yuan, L. Wang, H. Ran, F. Li, F. Liu, J. Zhang, W. Liu, W. Huang, Y. Huang, X. Xia, Proteomic analysis of murine macrophages mitochondria and lysosomes reveal Cathepsin D as a potential broad-spectrum antimicrobial protein. *J. Proteomics* **223**, 103821 (2020).
77. S.-D. Ha, A. Martins, K. Khazaie, J. Han, B. M. C. Chan, S. O. Kim, Cathepsin B is involved in the trafficking of TNF-α-containing vesicles to the plasma membrane in macrophages. *J. Immunol.* **181**, 690–697 (2008).
78. R. Elling, J. Chan, K. A. Fitzgerald, Emerging role of long noncoding RNAs as regulators of innate immune cell development and inflammatory gene expression. *Eur. J. Immunol.* **46**, 504–512 (2016).
79. M. K. Atianand, D. R. Caffrey, K. A. Fitzgerald, Immunobiology of long noncoding RNAs. *Annu. Rev. Immunol.* **35**, 177–198 (2017).
80. J. Blin, K. A. Fitzgerald, Perspective: The RNA exosome, cytokine gene regulation and links to autoimmunity. *Cytokine* **74**, 175–180 (2015).
81. C. Kew, W. Huang, J. Fischer, R. Ganesan, N. Robinson, A. Antebi, Evolutionarily conserved regulation of immunity by the splicing factor RNP-6/PUF60. *eLife* **9**, 1–23 (2020).
82. A. O. Olaitan, A. Aballay, Non-proteolytic activity of 19S proteasome subunit RPT-6 regulates GATA transcription during response to infection. *PLOS Genet.* **14**, e1007693 (2018).
83. V. Tiku, C. Kew, P. Mehrotra, R. Ganesan, N. Robinson, A. Antebi, Nucleolar fibrillarin is an evolutionarily conserved regulator of bacterial pathogen resistance. *Nat. Commun.* **9**, 1–10 (2018).
84. C.-W. Wu, K. Wimberly, A. Pietras, W. Dodd, M. B. Atlas, K. P. Choe, RNA processing errors triggered by cadmium and integrator complex disruption are signals for environmental stress. *BMC Biol.* **17**, 56 (2019).
85. R. Kaletsky, R. S. Moore, G. D. Vrla, L. R. Parsons, Z. Gitai, C. T. Murphy, *C. elegans* interprets bacterial non-coding RNAs to learn pathogenic avoidance. *Nature* **586**, 445–451 (2020).
86. B. A. Kudlow, L. Zhang, M. Han, Systematic analysis of tissue-restricted miRISCs reveals a broad role for microRNAs in suppressing basal activity of the *C. elegans* pathogen response. *Mol. Cell* **46**, 530–541 (2012).
87. Z. Ren, V. R. Ambros, *Caenorhabditis elegans* microRNAs of the let-7 family act in innate immune response circuits and confer robust developmental timing against pathogen stress. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E2366–E2375 (2015).
88. L. Sun, L. Zhi, S. Shakoor, K. Liao, D. Wang, microRNAs involved in the control of innate immunity in Candida infected *Caenorhabditis elegans*. *Sci. Rep.* **6**, 36036 (2016).
89. J. Le Pen, H. Jiang, T. Di Domenico, E. Kneuss, J. Kosałka, C. Leung, M. Morgan, C. Much, K. L. M. Rudolph, A. J. Enright, D. O'Carroll, D. Wang, E. A. Miska, Terminal uridylyltransferases target RNA viruses as part of the innate immune system. *Nat. Struct. Mol. Biol.* **25**, 778–786 (2018).
90. C. E. Richardson, T. Kooistra, D. H. Kim, An essential role for XBP-1 in host protection against immune activation in *C. elegans*. *Nature* **463**, 1092–1095 (2010).
91. E. J. Tillman, C. E. Richardson, D. J. Cattie, K. C. Reddy, N. J. Lehrbach, R. Droste, G. Ruvkun, D. H. Kim, Endoplasmic reticulum homeostasis is modulated by the forkhead transcription factor FKH-9 during infection of *Caenorhabditis elegans*. *Genetics* **210**, 1329–1337 (2018).
92. E. Vidak, U. Javoršek, M. Vizovišek, B. Turk, Cysteine cathepsins and their extracellular roles: Shaping the microenvironment. *Cells* **8**, 264 (2019).
93. K. Lasocki, A. A. Bartosik, J. Mierzejewska, C. M. Thomas, G. Jagura-Burdzy, Deletion of the parA (soj) homologue in *Pseudomonas aeruginosa* causes ParB instability and affects growth rate, chromosome segregation, and motility. *J. Bacteriol.* **189**, 5762–5772 (2007).
94. A. Paix, A. Folkmann, D. Rasoloson, G. Seydoux, High efficiency, homology-directed genome editing in *Caenorhabditis elegans* using CRISPR-Cas9 ribonucleoprotein complexes. *Genetics* **201**, 47–54 (2015).

95. D. J. Dickinson, A. M. Pani, J. K. Heppert, C. D. Higgins, B. Goldstein, Streamlined genome engineering with a self-excising drug selection cassette. *Genetics* **200**, 1035–1049 (2015).
96. M. R. Green, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (CSH Press, ed. 4, 2012).
97. J. Y. Jeong, H. S. Yim, J. Y. Ryu, H. S. Lee, J. H. Lee, D. S. Seen, S. G. Kang, One-step sequence-and ligation-independent cloning as a rapid and versatile cloning method for functional genomics Studies. *Appl. Environ. Microbiol.* **78**, 5440–5443 (2012).
98. M. Z. Li, S. J. Elledge, SLIC: A method for sequence- and ligation-independent cloning. *Methods Mol. Biol.* **852**, 51–59 (2012).
99. D. Graczyk, R. J. White, K. M. Ryan, Involvement of RNA polymerase III in immune responses. *Mol. Cell. Biol.* **35**, 1848–1859 (2015).
100. J. Pei, B. H. Kim, N. V. Grishin, PROMALS3D: A tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.* **36**, 2295–2300 (2008).
101. R. C. Edgar, MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004).
102. I. Letunic, P. Bork, Interactive tree of life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **47**, W256–W259 (2019).
103. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
104. M. Chekulaeva, H. Mathys, J. T. Zipprich, J. Attig, M. Colic, R. Parker, W. Filipowicz, miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. *Nat. Struct. Mol. Biol.* **18**, 1218–1226 (2011).
105. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10 (2011).
106. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
107. Y. Liao, G. K. Smyth, W. Shi, FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
108. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
109. D. Angeles-Albores, R. Y. Raymond, J. Chan, P. W. Sternberg, Tissue enrichment analysis for *C. elegans* genomics. *BMC Bioinformatics* **17**, 366 (2016).
110. A. D. Holdorf, D. P. Higgins, A. C. Hart, P. R. Boag, G. J. Pazour, A. J. M. Walhout, A. K. Walker, WormCat: An online tool for annotation and visualization of *Caenorhabditis elegans* genome-scale data. *Genetics* **214**, 279–294 (2020).
111. H. Li, Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
112. R. E. Workman, A. D. Tang, P. S. Tang, M. Jain, J. R. Tyson, R. Razaghi, P. C. Zuzarte, T. Gilpatrick, A. Payne, J. Quick, N. Sadowski, N. Holmes, J. G. de Jesus, K. L. Jones, C. M. Soulette, T. P. Snutch, N. Loman, B. Paten, M. Loose, J. T. Simpson, H. E. Olsen, A. N. Brooks, M. Akeson, W. Timp, Nanopore native RNA sequencing of a human poly(A) transcriptome. *Nat. Methods* **16**, 1–37 (2019).
113. U. Raudvere, L. Kolberg, I. Kuzmin, T. Arak, P. Adler, H. Peterson, J. Vilo, g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).
114. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
115. J. R. Wiśniewski, P. Ostasiewicz, M. Mann, High recovery FASP applied to the proteomic analysis of microdissected formalin fixed paraffin embedded cancer tissues retrieves known colon cancer markers. *J. Proteome Res.* **10**, 3040–3049 (2011).
116. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
117. R Core Team, R: A language and environment for statistical computing (2020).
118. Y. Perez-Riverol, A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D. J. Kundu, A. Inuganti, J. Griss, G. Mayer, M. Eisenacher, E. Pérez, J. Uszkoreit, J. Pfeuffer, T. Sachsenberg, Ş. Yilmaz, S. Tiwary, J. Cox, E. Audain, M. Walzer, A. F. Jarnuczak, T. Ternent, A. Brazma, J. A. Vizcaíno, The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).
119. C. Frøkjaer-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).
120. D. J. Dickinson, J. D. Ward, D. J. Reiner, B. Goldstein, Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. *Nat. Methods* **10**, 1028–1034 (2013).
121. G. J. Woodhead, C. A. Mutch, E. C. Olson, A. Chenn, Cell-autonomous β-catenin signaling regulates cortical precursor proliferation. *J. Neurosci.* **26**, 12620–12630 (2006).

Acknowledgments: We thank M. Krzyszton for helpful comments and the Dziembowski laboratory members for discussions. We are grateful to B. Goldstein, W. Filipowicz, D. Trono, and J. Jaworski for sharing plasmids; G. Jagura-Burdzy for sharing *P. aeruginosa* PAO1; D. Adamska for assistance with RNA-seq; O. Gewartowska, M. Szpila, E. Borsuk, and J. Gruchota for mice lines generation; and M. Hyjek-Składanowska for help with S200 chromatography. Some *C. elegans* strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). **Funding:** This work was supported by National Science Center (OPUS 17 UMO-2019/33/B/NZ2/01773 to A.D., OPUS 14 UMO-2017/27/B/NZ2/01234 to S.M., and PRELUDIUM 19 UMO-2020/37/N/NZ2/02893 to A.B.). This research was supported by the funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 810425. A.B. was also supported by the Foundation for Polish Science (FNP). **Author contributions:** A.D. acquired funding and directed the studies. A.D., V.L., S.M., and A.B. designed the experiments. V.L. performed all *C. elegans* experiments and analyzed initial RNA-seq data. A.B. established primary BMDM cell cultures and performed all subsequent BMDM experiments. P.S.K. analyzed RNA-seq and *C. elegans* DRS data. S.M. performed tethering assays and preliminary BMDM experiments. N.G. analyzed BMDM DRS data. T.W. performed statistical colocalization analysis. D.C. analyzed MS data. Z.M. participated in *C. elegans* life span, CFU analysis, and transgenic strains generation. K.D. and J.J.E. contributed resources and helped to design experiments. V.L. and A.D. wrote the manuscript with input from P.S.K., A.B., and J.J.E. **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. RNA-seq data have been deposited to GEO database under the accession number GSE163549. Nanopore DRS data have been deposited to ENA with the following accession numbers: *C. elegans* and BMDM DRS, PRJEB40892; tethering experiment DRS, ERS12230818, ERS12230819, ERS12230820, and ERS6477295. Proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (*118*) partner repository with identifier PXD023238. All plasmids and *C. elegans* strains generated in this study are available upon request from A.D. The mouse lines can be provided by A.D.'s pending scientific review and a completed material transfer agreement. Requests for the mice lines should be submitted to A.D.

Submitted 14 July 2022
Accepted 29 September 2022
Published 16 November 2022
10.1126/sciadv.add9468

Science Advances
AAAS

Supplementary Materials for

TENT5 cytoplasmic noncanonical poly(A) polymerases regulate the innate immune response in animals

Vladyslava Liudkovska *et al.*

Corresponding author: Andrzej Dziembowski, adziembowski@iimcb.gov.pl

Sci. Adv. **8**, eadd9468 (2022)
DOI: 10.1126/sciadv.add9468

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

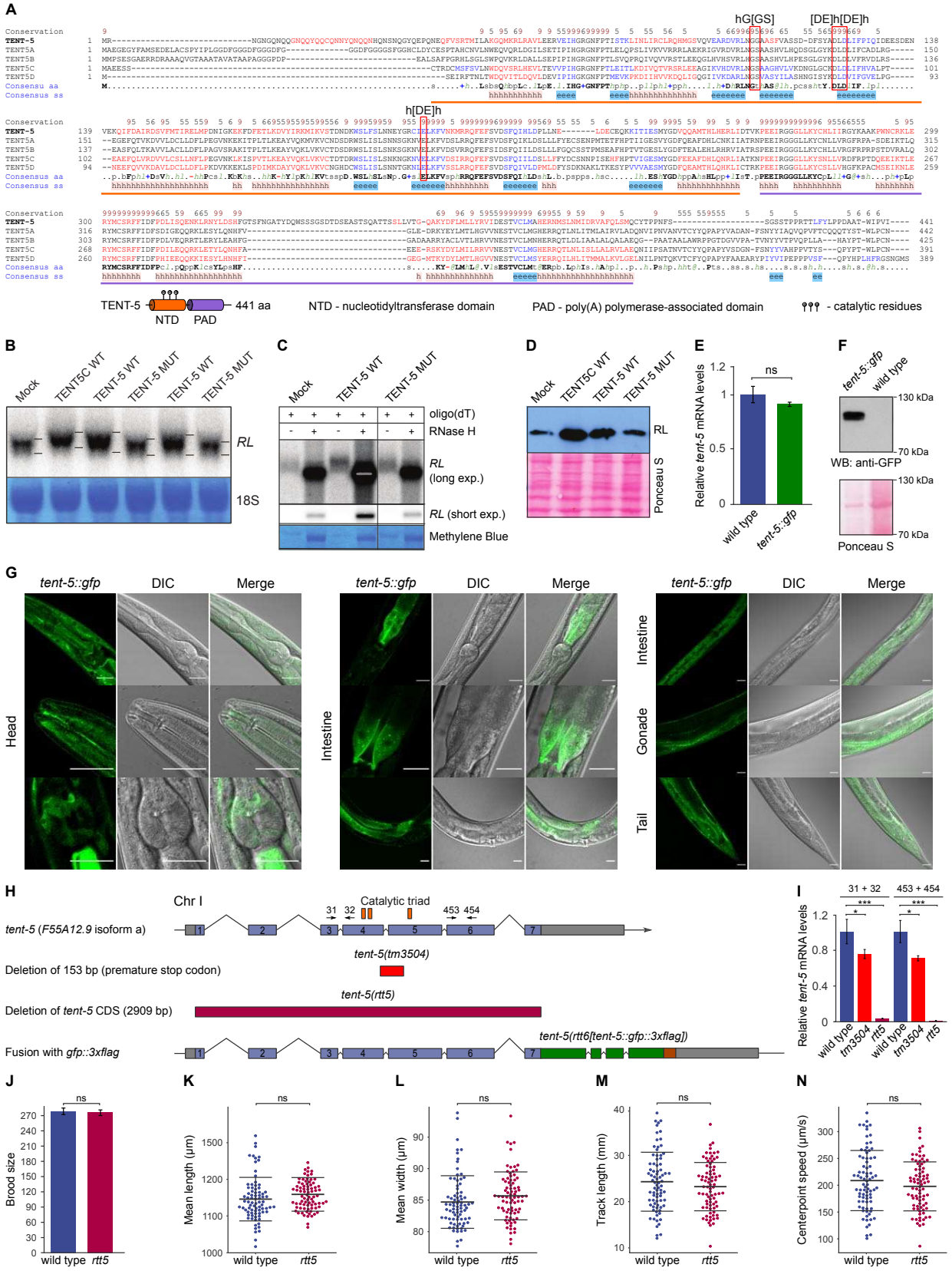

A
hG[GS]
[DE]h[DE]h
h[DE]h
TENT-5
441 aa
NTD
PAD
NTD - nucleotidyltransferase domain
PAD - poly(A) polymerase-associated domain
- catalytic residues
B
Mock
TENT5C WT
TENT-5 WT
TENT-5 MUT
RL
18S
C
oligo(dT)
RNase H
RL (long exp.)
RL (short exp.)
Methylene Blue
D
RL
Ponceau S
E
Relative tent-5 mRNA levels
ns
wild type
tent-5::gfp
F
tent-5::gfp
wild type
130 kDa
70 kDa
WB: anti-GFP
Ponceau S
G
tent-5::gfp
DIC
Merge
Head
Intestine
Gonade
Tail
H
Chr I
Catalytic triad
tent-5 (F55A12.9 isoform a)
Deletion of 153 bp (premature stop codon)
tent-5(tm3504)
Deletion of tent-5 CDS (2909 bp)
tent-5(rtt5)
Fusion with gfp::3xflag
tent-5(rtt6[tent-5::gfp::3xflag])
I
31 + 32
453 + 454
Relative tent-5 mRNA levels
wild type
tm3504
rtt5
J
Brood size
K
Mean length (µm)
L
Mean width (µm)
M
Track length (mm)
N
Centerpoint speed (µm/s)

Fig. S1. TENT-5 is a cytoplasmic ncPAP in worms.

(**A**) upper panel: Multiple sequence alignment of TENT-5 and human TENT5A-D proteins. Residues critical for the TENT5 proteins catalytic activity are marked with red boxes. TENT5 ncPAPs contain NTase domain with catalytic triad ([DE]h[DE]h, h[DE]h) and PAP-associated domain, underlined in orange and violet, respectively, and shown schematically in the lower panel (based on (*25*)). (**B**) Northern blot detection of *RL* reporter mRNA using total RNA from mock-transfected 293T cells or cells transfected with constructs that express NHA-hsTENT5C^{WT} (TENT5C WT, positive control), NHA-TENT-5^{WT} (TENT-5 WT), or NHA-TENT-5^{D151A,D153A} (TENT-5 MUT). Samples from two independent transfections are shown for TENT-5. In the cells transfected with NHA-hsTENT5C^{WT} or NHA-TENT-5^{WT} the *RL* mRNA migrates slowly, indicating its bigger length. Methylene blue staining of 18S rRNA serves as a loading control. (**C**) Northern blot detection of *RL* reporter mRNA using poly(A)+ RNA isolated from 293T cells transfected as in B. Observed differences in *RL* mRNA gel-migration upon tethering of NHA-TENT-5^{WT} resulted from mRNA polyadenylation, as revealed by the fact that *RL* mRNA from mock and NHA-TENT-5^{WT}-transfected cells migrated equivalently following RNA treatment with RNase H in the presence of oligo(dT)$_{25}$. Methylene blue staining serves as a loading control. (**D**) NHA-TENT-5^{WT} tethering increases RL reporter protein levels. Western blot detection of RL from mock-transfected 293 cells or from cells transfected with constructs that express NHA-hsTENT5C^{WT} (TENT5C WT, positive control), NHA-TENT-5^{WT} (TENT-5 WT), or NHA-TENT-5^{D151A,D153A} (TENT-5 MUT). Ponceau S staining serves as a loading control. (**E**) RT-qPCR quantification of *tent-5* mRNA levels in wild-type and *tent-5(rtt6[tent-5::gfp::3xflag] I*) knock-in strain. Data represent mean ± standard deviation (SD) of three biological replicates. ns – not significant (two-tailed t-test). (**F**) Anti-GFP immunoblot blot analysis of total protein lysate prepared from the *tent-5::gfp* worms. Total protein lysate from wild type was used as a negative control. Ponceau S staining of membrane serves as a loading control. (**G**) Representative fluorescence and DIC images of TENT-5-GFP expression. Scale bar is 20 µm. (**H**) Schematic of *tent-5* gene structure and mutant alleles. Exons are depicted as blue (*tent-5*), green (*gfp*), and brown (3x*flag*) boxes, UTRs as grey boxes, introns as lines. The approximate positions of triplets encoding catalytic residues are marked with orange boxes. Arrows correspond to the position of primers used for RT-qPCR. (**I**) RT-qPCR quantification of *tent-5* mRNA levels in L4 wild-type and mutant worms. Data represent mean ± SD of three biological replicates; $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$ (two-tailed t-test). (**J**) Total brood size of wild-type and *tent-5*-deficient worms (n=10, average of two independent trials). Data represent mean ± SD; ns – not significant (two-tailed unpaired t-test). Analysis of the worms' body length (**K**) and width (**L**), track length (**M**), and centerpoint speed (**N**) with the WormLab software. 80 young adult worms per strain have been analyzed. Data represent mean ± SD; ns – not significant (two-tailed unpaired t-test with Welch's correction).

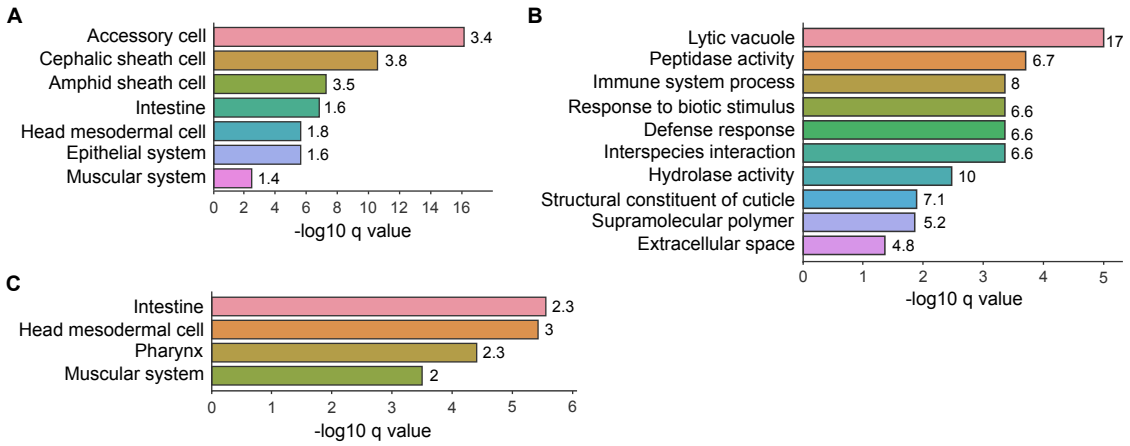


Fig. S2. Loss of *tent-5* leads to the downregulation of genes that encode innate immune effectors. (**A** and **C**) Tissue enrichment analysis of downregulated genes (log2FC < -log2(1.5), RNA-seq) (**A**) and proteins (p value < 0.05, Mass Spectrometry) (**C**) which levels were decreased in *tent-5(tm3504)* mutant compared to the wild-type worms. (**B**) Overrepresented functional GO terms of proteins with decreased abundance in *tent-5*-deficient worms in comparison to the wild type (Mass Spectrometry). Numbers on the right indicate the enrichment fold change for each term (observed/expected genes).

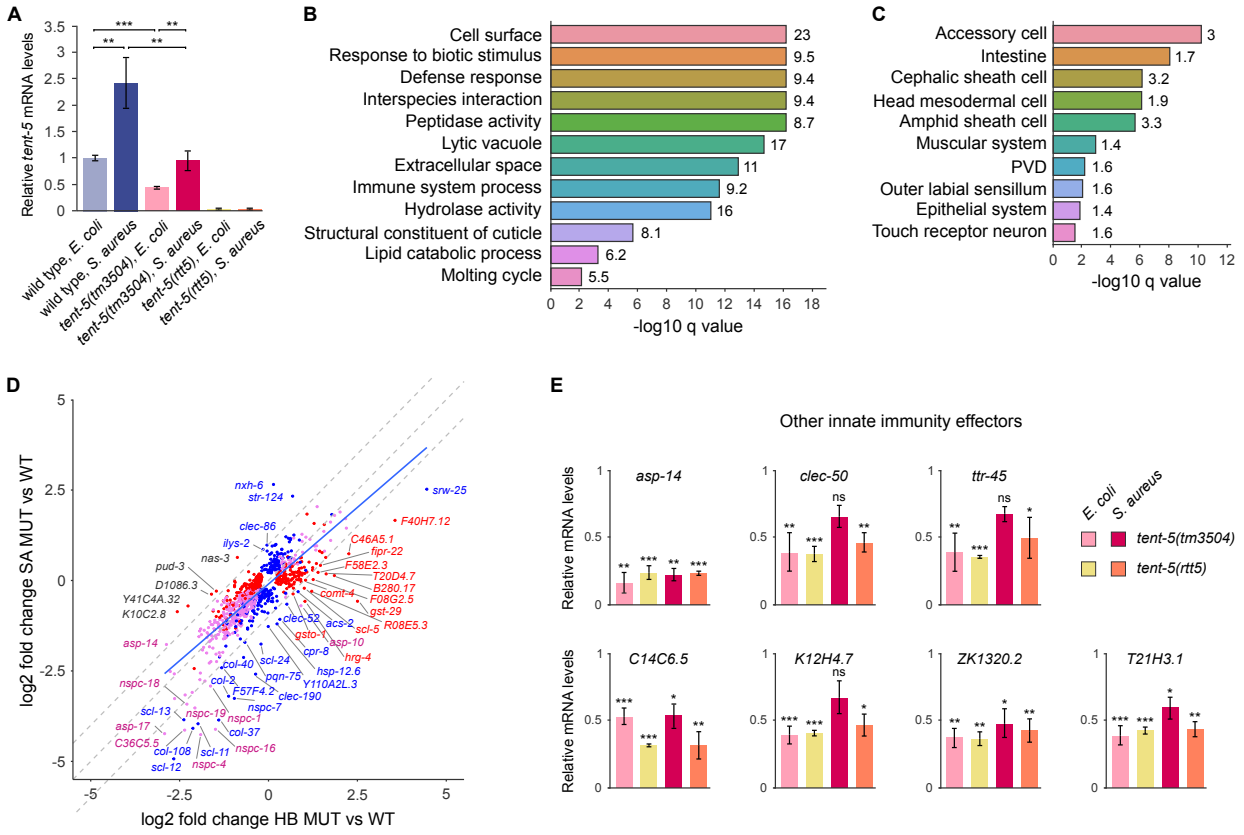


Fig. S3. TENT-5 deficiency leads to the downregulation of genes that encode innate immune effectors. (**A**) RT-qPCR analysis of *tent-5* expression in wild type, *tent-5(tm3504)*, and *tent-5(rtt5)* worms that were exposed to *S. aureus* or control *E. coli* for 8 hours. Relative *tent-5* mRNA level was normalized to *act-1* and wild type. Data represent mean ± SD of three biological replicates; $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$ (two-tailed t-test). (**B** and **C**) Overrepresented functional GO and TEA terms of genes downregulated (log2FC < -log2(1.5)) in *tent-5(tm3504)* mutant compared to wild type upon infection of L4 worms with *S. aureus* for 8 hours. Numbers on the right indicate the enrichment fold change for each term. (**D**) Scatter plot illustrates a comparison of the differential expression of genes (|log2FC| > 0, FDR < 0.05) between mutant (MUT) and wild type (WT) in infected (SA) and non-infected (HB) animals. Red and blue dots indicate genes differentially expressed in wild-type and mutant worms, respectively. Dots in magenta indicate the common part. The middle diagonal dashed line shows the equal response of both genotypes, and two additional dashed lines show the borders of the area where the response is +/- 1 log2FC. The solid blue line shows a linear trend created using the linear regression model for all data points. (**E**) RT-qPCR analysis of relative levels of mRNA expression in two *tent-5*-deficient worm strains grown on *E. coli* or exposed to *S. aureus* for 8 hours. Relative abundance of mRNAs was normalized to *act-1* and wild type. Data represent mean ± SD of three biological replicates; ns – not significant; $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$ (two-tailed t-test).

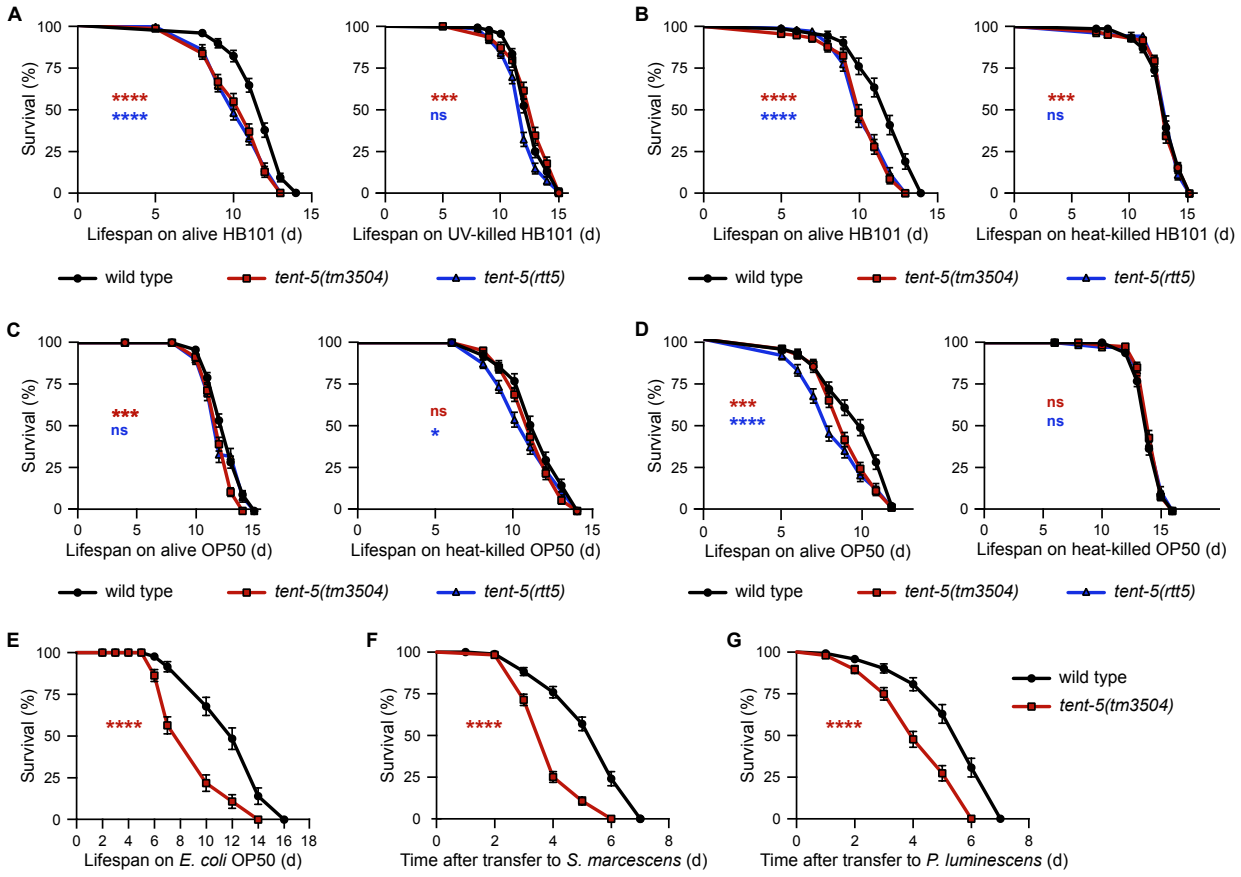


Fig. S4. TENT-5 is required for the host defense against various bacterial strains. (**A-D**) The lifespan of wild type, *tent-5(tm3504),* and *tent-5(rtt5)* worms grown at 25°C on plates seeded with alive or UV-killed *E. coli* HB101 (**A**), alive or heat-killed *E. coli* HB101 (**B**), and alive or heat-killed *E. coli* OP50 (**C** and **D**). For these experiments, NGM plates contained FUdR. **P* < 0.05; ***P* < 0.01; ****P* < 0.001; *****P* < 0.0001; ns – not significant (Log-Rank test). (**E-G**) Survival of wild type and *tent-5(tm3504)* worms grown on *E. coli* OP50 (**E**) or infected with *S. marcescens* (**F**) and *P. luminescens* (**G**). Experiments presented in **E-G** are representative of at least two independent trials; **P* < 0.05; ***P* < 0.01; ****P* < 0.001; *****P* < 0.0001; ns – not significant (Log-Rank test). Survival statistics related to lifespan and survival assays can be found in table S1.

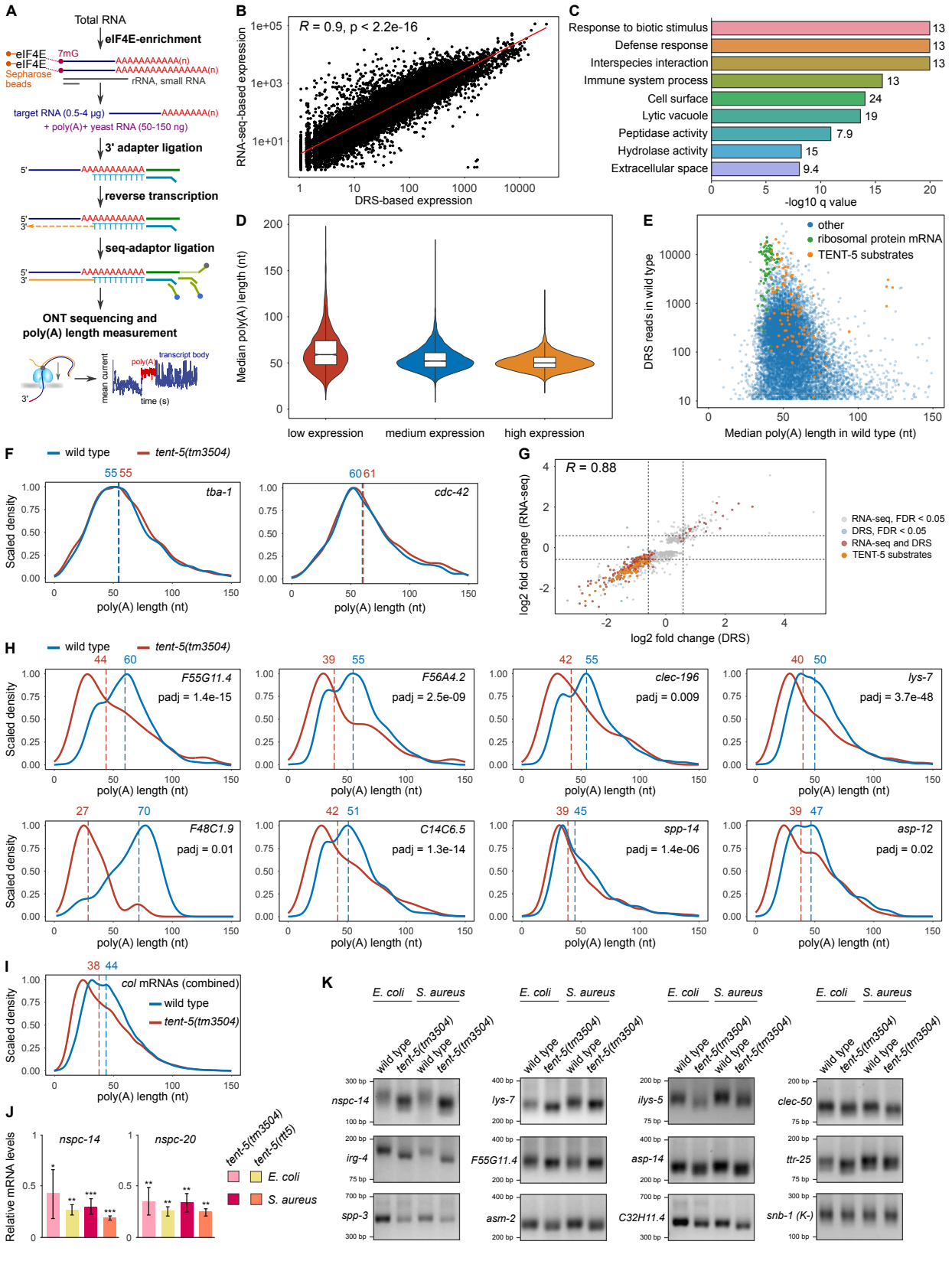
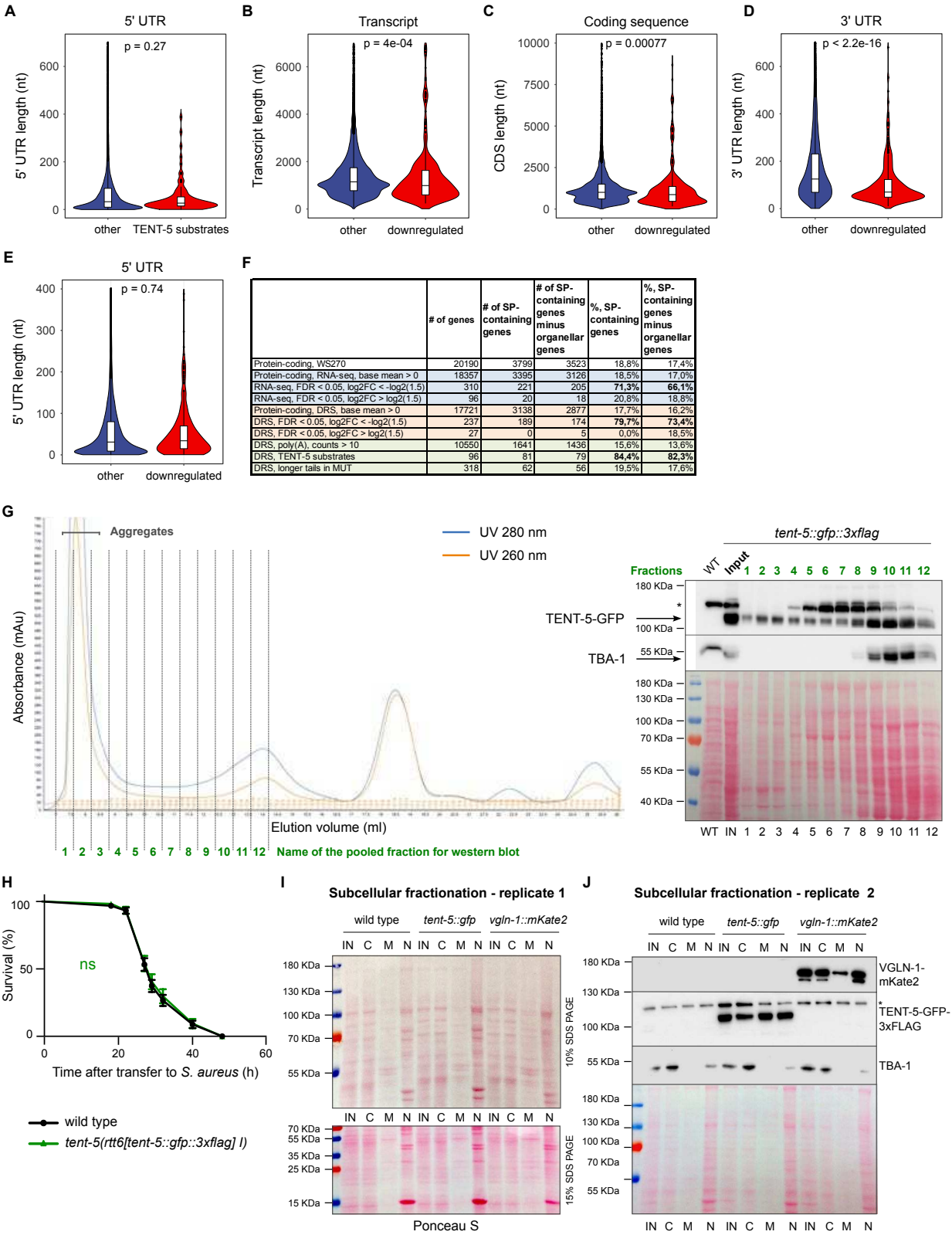

A
Total RNA
eIF4E-enrichment
eIF4E
eIF4E
7mG
Sepharose beads
AAAAAAAAAA(n)
AAAAAAAAAAAAAAAA(n)
rRNA, small RNA
target RNA (0.5-4 µg)
AAAAAAAA(n)
+ poly(A)+ yeast RNA (50-150 ng)
3' adapter ligation
reverse transcription
seq-adaptor ligation
ONT sequencing and
poly(A) length measurement
transcript body
poly(A)
mean current
time (s)
B
R = 0.9, p < 2.2e-16
RNA-seq-based expression
DRS-based expression
C
Response to biotic stimulus
Defense response
Interspecies interaction
Immune system process
Cell surface
Lytic vacuole
Peptidase activity
Hydrolase activity
Extracellular space
-log10 q value
D
Median poly(A) length (nt)
low expression
medium expression
high expression
E
other
ribosomal protein mRNA
TENT-5 substrates
DRS reads in wild type
Median poly(A) length in wild type (nt)
F
wild type
tent-5(tm3504)
tba-1
cdc-42
Scaled density
poly(A) length (nt)
G
R = 0.88
log2 fold change (RNA-seq)
log2 fold change (DRS)
RNA-seq, FDR < 0.05
DRS, FDR < 0.05
RNA-seq and DRS
TENT-5 substrates
H
F55G11.4
padj = 1.4e-15
F56A4.2
padj = 2.5e-09
clec-196
padj = 0.009
lys-7
padj = 3.7e-48
F48C1.9
padj = 0.01
C14C6.5
padj = 1.3e-14
spp-14
padj = 1.4e-06
asp-12
padj = 0.02
I
col mRNAs (combined)
J
nspc-14
nspc-20
Relative mRNA levels
tent-5(tm3504)
tent-5(rt5)
E. coli
S. aureus
K
E. coli
S. aureus
wild type
tent-5(tm3504)
nspc-14
irg-4
spp-3
lys-7
F55G11.4
asm-2
ilys-5
asp-14
C32H11.4
clec-50
ttr-25
snb-1 (K-)

Fig. S5. Transcripts downregulated in the *tent-5*-deficient mutants are the direct targets for polyadenylation by TENT-5. (**A**) Schematic of library preparation for Direct RNA sequencing. (**B**) Correlation between gene expression levels in the wild-type worms measured by DRS and RNA-seq. The red line represents a linear regression line. Genes with normalized expression of minimum 1 were included in the analysis (n=13,632). Values of Pearson's correlation coefficient and statistical test are indicated on the plot. (**C**) Overrepresented functional GO terms of transcripts significantly downregulated (log2FC < -log2(1.5)) in *tent-5(tm3504)* mutant in comparison to the wild-type worms. Numbers on the right indicate the enrichment fold change for each term. (**D**) Violin plots showing the median poly(A) tail length distribution for genes with low (first quartile), medium (interquartile range) and highly (third quartile) expressed mRNA in the wild-type worms. (**E**) Plot illustrating the relationship between median poly(A) tail length and mRNA expression in wild-type worms. TENT-5 substrates are indicated in orange, mRNAs that encode ribosomal proteins are marked in green. (**F-I**) DRS-based poly(A) length profiling of mRNAs that encode established reference genes (**F**), indicated innate immune effectors (**H**) and collagens (**I**). Shown are density distribution plots scaled to 1. Vertical dashed lines represent median poly(A) lengths (nt); padj values (Wilcoxon test) are shown on the top right corner below the mRNA ID. (**G**) Scatter plot illustrates a comparison of the differential expression of genes (FDR < 0.05) between *tent-5* mutant and wild-type worms, as identified by RNA-seq and DRS. Red dots indicate genes differentially expressed in wild-type and mutant worms according to both sequencing methods (n=231), orange dots present TENT-5 substrates (n=96). Dashed lines mark the threshold of the log2(1.5) fold change. Pearson's correlation test. (**J**) RT-qPCR analysis of relative levels of *nspc-14* and *nspc-20* mRNA expression in two *tent-5*-deficient worm strains grown on *E. coli* or exposed to *S. aureus* for 8 hours. Relative abundance of mRNAs was normalized to *act-1* and wild type. Data represent mean ± SD of three biological replicates; $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$ (two-tailed t-test). (**K**) Poly(A) tail PCR tests of selected TENT-5-regulated transcripts. Samples were prepared from *tent-5(tm3504)* mutant and wild-type worms infected with *S. aureus* for 8 hours or grown on control *E. coli.* PCR products were analyzed on 2% agarose gels.



	# of genes	# of SP-containing genes	# of SP-containing genes minus organellar genes	%, SP-containing genes	%, SP-containing genes minus organellar genes
Protein-coding, WS270	20190	3799	3523	18,8%	17,4%
Protein-coding, RNA-seq, base mean > 0	18357	3395	3126	18,5%	17,0%
RNA-seq, FDR < 0.05, log2FC < -log2(1.5)	310	221	205	**71,3%**	**66,1%**
RNA-seq, FDR < 0.05, log2FC > log2(1.5)	96	20	18	20,8%	18,8%
Protein-coding, DRS, base mean > 0	17721	3138	2877	17,7%	16,2%
DRS, FDR < 0.05, log2FC < -log2(1.5)	237	189	174	**79,7%**	**73,4%**
DRS, FDR < 0.05, log2FC > log2(1.5)	27	0	5	0,0%	18,5%
DRS, poly(A), counts > 10	10550	1641	1436	15,6%	13,6%
DRS, TENT-5 substrates	96	81	79	**84,4%**	**82,3%**
DRS, longer tails in MUT	318	62	56	19,5%	17,6%

Fig. S6. TENT-5 regulates mRNAs that encode secreted proteins. (**A**) Violin plot showing length distribution of 5′ UTRs of TENT-5 substrates (n=96) compared to other transcripts

identified by DRS (n=16,568); (p value, Wilcoxon test). Violin plots illustrating length distribution of transcripts (**B**), coding sequences (**C**), 3′ UTRs (D), and 5′ UTRs (**E**) of transcripts, whose expression levels were decreased at least 1.5-fold in *tent-5(tm3504)* mutant worms (n=308) compared to other transcripts identified by RNA-seq (n=18,229); (p value, Wilcoxon test). (**F**) Fractions of genes that encode secreted proteins (%) as defined by (*60*) across indicated datasets. For more details also see data S4. (**G**) The chromatography analysis of total protein extracts from the *tent-5(rtt6)* worms on Superdex 200 (on the left) followed by western blot analysis (on the right). Fractions pooled for western blot analysis are indicated with dotted lines and numbered in green. Most of TENT-5-GFP is present in the fraction corresponding to the TENT-5-GFP molecular weight (pooled fractions 9-11) and a small amount of TENT-5 can be found in macromolecular complexes or aggregates (pooled fractions 1-3). Anti-tubulin (TBA-1) antibodies were used as control, and Ponceau S staining illustrates the relative amount of protein in each pooled fraction. (**H**) Survival of wild-type and *tent-5::gfp::3xflag* worms infected with *S. aureus*; ns – not significant (Log-Rank test). (**I-J**) Subcellular fractionation of proteins isolated from wild-type, *tent-5::gfp::3xflag,* and *vgln-1::mKate2:::3xmyc* strains. (**I**) Ponceau S staining of the whole membrane used for the western blot presented in Fig. 6F (10% SDS PAGE) and membrane containing the same extracts resolved on the 15% SDS PAGE indicating successful fractionation of proteins. (**J**) Second biological replicate of subcellular fractionation shown in Fig. 6F. Anti-tubulin (TBA-1) and anti-RFP (VGLN-1-mKate2) antibodies and Ponceau S staining were used as fractionation controls. IN – input sample, C – cytoplasmic fraction, M – membrane fraction, N – nuclear fraction. Asterisks indicate nonspecific bands.

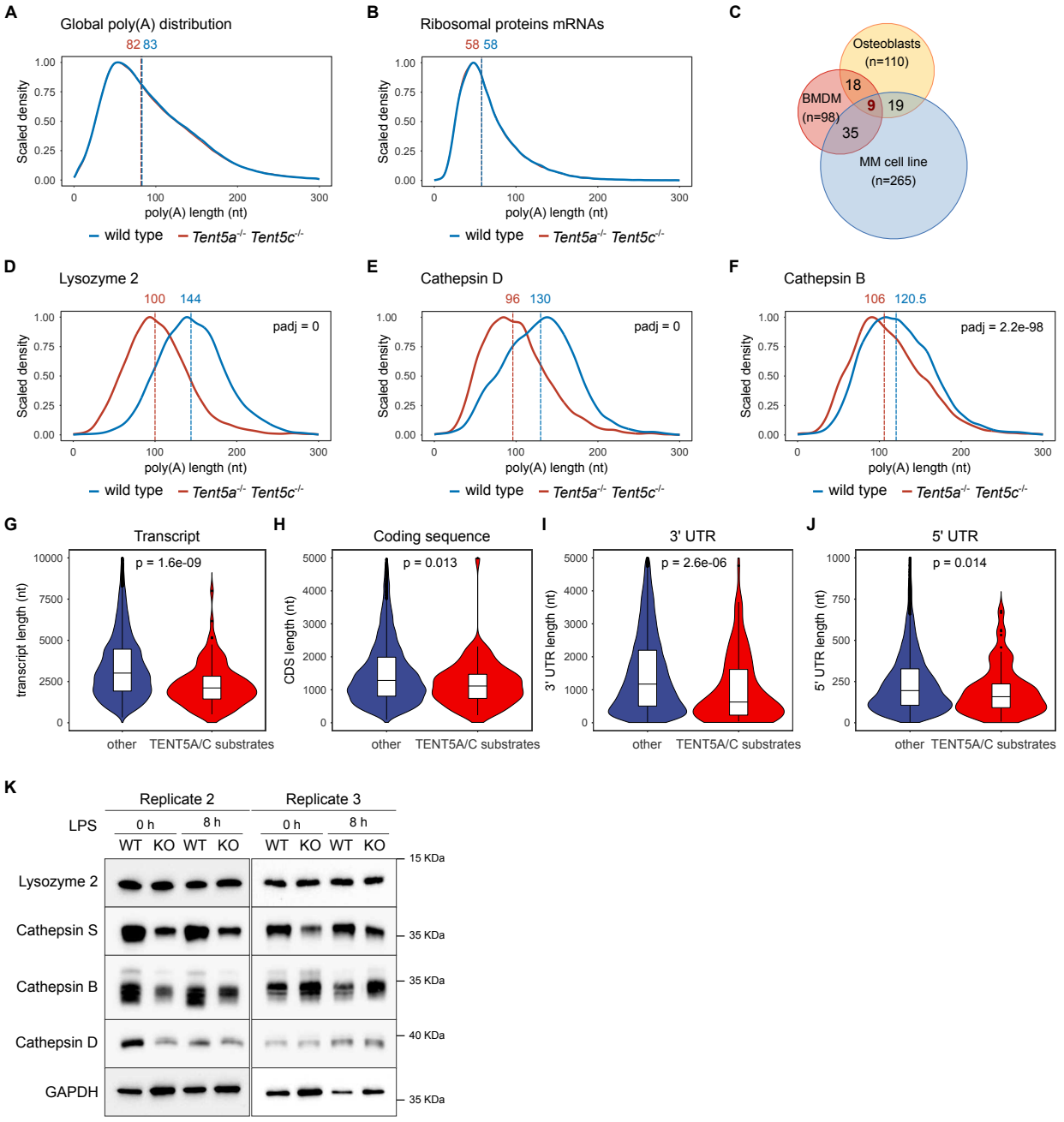


Fig. S7. The role of TENT-5 in innate immunity is evolutionarily conserved. DRS-based poly(A) length profiling of mRNA isolated from *Tent5a*$^{-/-}$ *Tent5c*$^{-/-}$ and wild-type BMDM. Shown are density distribution plots for all transcripts (**A**), mRNAs encoding ribosomal proteins (**B**), *Lyz2* (**D**), *Ctsd* (**E**), and *Ctsb* (**F**) mRNAs scaled to 1. Vertical dashed lines represent median poly(A) lengths (nt); for *Lyz2*, *Ctsd*, and *Ctsb* transcripts padj values (Wilcoxon test) are shown. (**C**) Overlap between TENT5A/C molecular substrates identified in BMDM (red, DRS; this study), osteoblasts (yellow, DRS) (*26*), and multiple myeloma cell lines (blue, RNA-seq after RNA poly(A) length fractionation) (*21*). (**G-I**) Violin plots showing length distribution of transcripts (**G**), coding sequences (**H**), 3′ UTRs (**I**), and 5′ UTRs (**J**) of TENT5A/C substrates (n=98) compared to all other transcripts identified by DRS (n=25,678); (p value, Wilcoxon test).

(**K**) Western blots showing that the lack of TENT5A and TENT5C leads to the lower abundance of LYZ2, CTSS, CTSD, and CTSB both in unstimulated (LPS, 0 hours) and stimulated (LPS, 8 hours) BMDM derived from mutant compared to the wild-type mouse. GAPDH serves as a loading control. Shown are replicates 2 and 3.

Table S1. Survival statistics related to lifespan and pathogenesis assays.

MS: median survival
N: total number of animals that died
P value: compared to the wild type, based on Log-rank (Mantel-Cox) test

	Replicate 1			Replicate 2			Replicate 3		
	MS, days	**N**	**p value**	**MS, days**	**N**	**p value**	**MS, days**	**N**	**p value**
***E. coli HB101*, 20°C, FUdR (Fig. 4A)**									
wild type	26	80		28	91				
tent-5(tm3504)	19	97	p < 0.0001	26	84	0.0091			
tent-5(rtt5)	26	82	ns	28	110	ns			
***E. coli HB101*, 25°C, FUdR (Fig. 4B)**									
wild type	12	151		11	179		14	104	
tent-5(tm3504)	10	129	0.0042	10	176	p < 0.0001	12	113	p < 0.0001
tent-5(rtt5)	12	131	ns	10	184	p < 0.0001	12	138	p < 0.0001
***E. coli HB101* alive, 25°C, FUdR, control to UV-killed HB101 (Fig. S4A)**									
wild type	12	132							
tent-5(tm3504)	11	115	p < 0.0001						
tent-5(rtt5)	10	138	p < 0.0001						
***E. coli HB101* UV-killed, 25°C, FUdR (Fig. S4A)**									
wild type	13	134							
tent-5(tm3504)	13	95	ns						
tent-5(rtt5)	12	130	0,001						
***E. coli HB101* alive, 25°C, FUdR, control to heat-killed HB101 (Fig. S4B)**									
wild type	11	113							
tent-5(tm3504)	11	139	p < 0.0001						
tent-5(rtt5)	10	149	p < 0.0001						
***E. coli* HB101 heat-killed, 25°C, FUdR (Fig. S4B)**									
wild type	13	154							
tent-5(tm3504)	13	144	ns						
tent-5(rtt5)	13	107	ns						
***E. coli OP50* alive, 25°C, FUdR, control to heat-killed HB101 (Fig. S4C-D)**									
wild type	13	170		10	120				
tent-5(tm3504)	12	144	p < 0.0001	9	143	0,0004			
tent-5(rtt5)	12	104	ns	8	126	p < 0.0001			
***E. coli OP50* heat-killed, 25°C, FUdR (Fig. S4C-D)**									
wild type	12	100		14	177				
tent-5(tm3504)	11	130	ns	14	146	ns			
tent-5(rtt5)	11	138	0,0414	14	161	ns			
***S. aureus* (Fig. 4D)**									
wild type	40	75		40	86		38	103	
tent-5(tm3504)	36	111	p < 0.0001	40	107	p < 0.0001	31	144	0.0002
tent-5(rtt5)	36	104	p < 0.0001	40	109	p < 0.0001	31	127	0.0010

***P. aeruginosa PAO1* (Fig. 4E)**									
wild type	7	102		6	132		6	95	
tent-5(tm3504)	6	110	$p < 0.0001$	5	142	$p < 0.0001$	5	90	$p < 0.0001$
tent-5(rtt5)	5	94	$p < 0.0001$	5	161	$p < 0.0001$	5	97	$p < 0.0001$
***E. coli OP50*, 25°C (Fig. S4E)**									
wild type	13	92		12	101				
tent-5(tm3504)	8	127	$p < 0.0001$	10	116	$p < 0.0001$			
***S. marcescens Db10*, 25°C (Fig. S4F)**									
wild type	6	115		4	84				
tent-5(tm3504)	4	173	$p < 0.0001$	3	121	$p < 0.0001$			
***P. luminescens*, 25°C (Fig. S4G)**									
wild type	6	90		4	96				
tent-5(tm3504)	4	103	$p < 0.0001$	3	112	$p < 0.0001$			

Table S2. List of oligonucleotides and sgRNAs used in this study.

ID	Name	Sequence 5′ to 3′	Description
Primers used for PCR genotyping and sequencing of *C. elegans* strains (this study)			
VL001	tent-5-wt-Rev	tcaggtttccactgacaatg	Used for: *tent-5(tm3504), tent-5(rtt5), tent-5(rtt6),* ADZ47
VL002	tent-5-ko-Rev	tgatctcgacctgatattcc	
VL003	tent-5-Fw	gttcacactcgtccaactc	
VL318	tent-5-crispr-Fw	gatcgtgtcgcattccagg	
VL319	tent-5-crispr-Rev	gtgccgacaaattgatggtg	
VL340	tent-5-crispr-Fw2	acgtaagatctccgatccactaattcaaaatttac	
VL355	tent-5-crispr-Fw3	ggtcttcatctggatcagatac	Used for *tent-5(rtt6)*, ADZ47
VL356	tent-5-crispr-Fw4	caaggtaagtttaaacagttc	
VL357	tent-5-crispr-Rev2	ctttaaaatttcagattttgatgta	
Primers used for cloning (this study)			
VL342	sgRNA-universal-Fw	aagcatgcaatttttgagaaactc	Cloning of pDD162-*sgRNA425* (pVL060)
VL343	sgRNA-universal-Rev	gagtgcaccatatgcggtgtgaa	
VL344	sgRNA425-Rev1	tgtagctgcatctggtggcacaagacatctcgcaatagg	
VL345	sgRNA425-Fw1	tgccaccagatgcagctacagttttagagctagaaatagcaagt	
VL346	sgRNA-universal-seq	cgttttacaacgtcgtgactggg	
VL386	sgRNA-universal-Fw2	ctataatttgtcaacctttttcaaaaaaagcatgcaatttttgagaaactc	
VL387	sgRNA-universal-Rev2	cggcatcagagcagattgtactgagagtgcaccatatgcggtgtgaa	
VL347	tent-5-gfp-5′arm-Rev (silent mutations)	catcgatgctcctgaggctcccgatgctccgatgactggaatccaagttgcagcgtccggcggcaagtaga	Cloning of pDD282-*tent-5::gfp::3xflag* (pVL062)
VL348	tent-5-gfp-5′arm-Fw	acgttgtaaaacgacggccagtcgccggcatatgtcgttgaacatgattgatcgtgtcgc	
VL349	tent-5-gfp-3′arm-Fw	cgtgattacaaggatgacgatgacaagagatgatccaattgaacctcttccccctcttac	
VL350	tent-5-gfp-3′arm-Rev	ggaaacagctatgaccatgttatcgatttccagattttgatgtataaacttcgtaggaaatatacag	
VL364	sgRNA-456-F	cgttctttaccaacgacgaggttttagagctagaaatagcaagt	Cloning of pDD162-*sgRNA456* (pVL069)
VL365	sgRNA-456-R	ctcgtcgttggtaaagaacgcaagacatctcgcaatagg	
VL366	vgln-1-mKate2-5arm-R	tatgaccatgattacgccaagctattaaaactccgaactatgatgac	Cloning of pDD287-*vgln-1::mKate2::3xmyc* (pVL070)
VL367	vgln-1-mKate2-5arm-F	catcgatgctcctgaggctcccgatgctccccaacgacgagaggatccccaagca	
VL368	vgln-1-mKate2-3arm-F	ctcgagcagaagttgatcagcgaggaagacttgtaaagaacgttcgaatctggatgattca	
VL369	vgln-1-mKate2-3arm-R	ggaaacagctatgaccatgttatcgatttctctcacatgtacatttagcctacatc	
21	NHATENT-5-Fw	ctagagtcgacatgcgtggaagacgttcccg	Cloning of pCI-*NHA-tent-5*
23	NHATENT-5-Rev	gggaagcggccgctcagatgactggaatccatg	
Primers for RT-qPCR analysis (this study)			
VL031	q-tent-5-Fw	ggcaacttcccaacaatctcc	Efficiency: 1.93
VL032	q-tent-5-Rev	gaagaggctacaaatgatgcgg	
VL453	q-tent-5-Fw2	gtgctcccgcttcttcatcg	2.0
VL454	q-tent-5-Rev2	ccattgatcgtaagtggctccg	
VL033	q-act-1-Fw	gctggacgtgatcttactgattacc	1.95

VL034	q-act-1-Rev	gtagcagagcttctccttgatgtc	
VL096	q-ilys-5-Fw	tgcctggaagagatgtgctg	1.96
VL097	q-ilys-5-Rev	tgcactgagacttgtagcgg	
VL101	q-spp-14-Fw	gaccaccttctgccttctgg	2.0
VL102	q-spp-14-Rev	gcattgtcctgttggcttgg	
VL104	q-spp-3-Fw	gtcgacgctgaatgcaagaag	1.93
VL105	q-spp-3-Rev	ttgggtcgagcttggtgtc	
VL108	q-clec-160-Fw	ggcagcaatgcagtatggac	2.0
VL109	q-clec-160-Rev	acgctcgtctagaaatcggac	
VL111	q-clec-50-Fw	accgattccaactacctgcc	1.98
VL112	q-clec-50-Rev	tccatccttggtcgcaagac	
VL116	q-irg-4-Fw	tcggtcgaaacccaattccc	1.92
VL117	q-irg-4-Rev	aaattggatcgacgcagcac	
VL119	q-F55G11.4-Fw	tgttgcatctattgggcagg	1.92
VL120	q-F55G11.4-Rev	ctcctggtgcagcgtattcc	
VL122	q-irg-5-Fw	tcatggacctgattacaccgc	1.94
VL123	q-irg-5-Rev	agcgaatcggatggagtacg	
VL124	q-asp-14-Fw	ttaccttgacaccaccaccg	1.99
VL125	q-asp-14-Rev	aagggagtcaaacacctggc	
VL130	q-T21H3.1-Fw	ttcggcctttgtgctaccag	1.98
VL131	q-T21H3.1-Rev	tggcgtcgttgtatgggatg	
VL132	q-K12H4.7-Fw	acattcggagaggctcaatctg	1.94
VL133	q-K12H4.7-Rev	tccgaactcagtgcaagtctg	
VL206	q-nspc-14-Fw	gccctctgccttgtttctgc	1.94
VL207	q-nspc-14-Rev	aagattcgccaccattgacacc	
VL208	q-nspc-20-Fw	tgtgtgcataactgtgctgctg	1.90
VL209	q-nspc-20-Rev	gcttctttgctgggaatgttgc	
VL224	q-spp-5-Fw	ccgataaggatgccaatgtc	1.98
VL225	q-spp-5-Rev	gggtctaccttgctgttgac	
VL231	q-lys-7-Fw	atccggtcgtgatctgattc	2.0
VL232	q-lys-7-Rev	gattgttgactcatcccttc	
VL236	q-C14C6.5-Fw	tggtgtcagtgactgcccag	2.0
VL237	q-C14C6.5-Rev	agccacagtatccgcaggtc	
VL406	q-F28B4.3-Fw1	agaccatctctggtggaagcg	1.96
VL407	q-F28B4.3-Rev	tcctccattggtgctagcgg	
VL409	qr-ttr-45-Fw1	gcttgtggctgcaatgaggg	1.98
VL410	q-ttr-45-Rev	ccgagtcctcttcccagagc	
VL414	q-ZK1320.2-Fw	aatgccggatggaacctggg	1.96
VL415	q-ZK1320.2-Rev	acatctgcactgttagttgtggg	
Primers used for Poly(A) PCR tests (this study)			
	RA3-15N	/5rApp/ctgacnnnnnnnnnnnnnnnntggaattctcgg gtgccaagg/3ddC/ (5′ pre-adenylated, with 15 nt unique molecular identifier)	3′ adaptor
	RPI	caagcagaagacggcatacgagatnnnnnngtgactgga gttccttggcacccgagaattcca	RNA PCR Rev Primer
VL294	RP-uni-Rev	caagcagaagacggcatacgagat	
VL116	r-irg-4-Fw1	tcggtcgaaacccaattccc	
VL446	r-irg-4-Fw2	gacaaacatcactcgtccatcg	
VL156	r-lys-7-Fw1	actcctgccaacttcaatgact	

VL235	r-lys-7-Fw2	aatgatacggcgaccaccgagatctacacgttcagagttcta cagtccgacgatcatgctggaggatttgtgcaaggaa	
VL212	r-clec-50-Fw1	ggatctcgcgtggatcggta	
VL213	r-clec-50-Fw2	aatgatacggcgaccaccgagatctacacgttcagagttcta cagtccgacgatcctgcaagaagggatctattcatgca	
VL164	r-F55G11.4-Fw1	gtgaacgcgggagcacata	
VL211	r-F55G11.4-Fw2	aatgatacggcgaccaccgagatctacacgttcagagttcta cagtccgacgatcgtggcaaggctgtattcacgttta	
VL217	r-nspc-14-Fw1	gagttctgtgctggtgtcaatg	
VL218	r-nspc-14-Fw2	tgtgtatgttgcttcagtcgt	
VL104	r-spp-3-Fw1	gtcgacgctgaatgcaagaag	
VL221	r-spp-3-Fw2	cgtctgcaccaagcttggaatgtg	
VL165	r-asp-14-Fw1	tcagcacaaacgttatgaatgct	
VL216	r-asp-14-Fw2	aatgatacggcgaccaccgagatctacacgttcagagttcta cagtccgacgatccactccagcgcttctccaagattag	
VL214	r-asm-2-Fw1	ctcacctcaatcgtggtccg	
VL215	r-asm-2-Fw2	aatgatacggcgaccaccgagatctacacgttcagagttcta cagtccgacgatcgatcatttaacaaaatagtttcac	
VL222	r-ttr-25-Fw1	cacagatgcaaccacgtcgg	
VL223	r-ttr-25-Fw2	ggagagggtactgattgcattca	
VL268	r-snb-1-Fw1	cggcgcgataagctgcgtga	
VL269	r-snb-1-Fw2	cagactaacaacatgccttctc	
VL449	r-C32H11.4-Fw1	cggatctgaaactgcgttcg	
VL450	r-C32H11.4-Fw2	cactcacgttctacggcgaa	
VL096	r-ilys-5-Fw1	tgcctggaagagatgtgctg	
VL475	r-ilys-5-Fw2	ggatactatgaagactgcgg	
Primers for RL northern blot probe preparation (this study)			
	RL_Fw	atgattactggtccacaatgg	
	RL_Rev	ggacacgctcaacaaacgat	
Oligonucleotides used to generate *tent-5(rtt5)* strain by CRISPR/Cas9			
VL311	tracrRNA-fragment	AAAAAGCACCGACTCGGTGCCACTTT TTCAAGTTGATAACGGACTAgccttattttaa cttgctatttctagctcta	
VL312	crRNA01-tent-5	gaaattaatacgactcactatagggGTTCAATTGGA TCAGATGACgttttagagctagaaatagcaagttaaaa taaggc	Description: underlined – T7 RNA promoter; blue – 5′N20 guide RNA sequence; red – crRNA scaffold
VL315	crRNA02-tent-5	gaaattaatacgactcactatagggCTCATTCTGTT TGTTCCGTGgttttagagctagaaatagcaagttaaaa taaggc	
VL313	crRNA-dpy-10	gaaattaatacgactcactatagggGCTACCATAGG CACCACGAGgttttagagctagaaatagcaagttaaa ataaggc	
VL314	ssODN-dpy-10	cacttgaacttcaatacggcaagatgagaatgactggaaacc gtaccgcatgcggtgcctatggtagcggagcttcacatggctt cagaccaacagcctat	(*94*)
VL326	ssODN-tent-5	ctggattttaaattattaatttaaaaaaaaattttcagttggctcat catggaacaaacagatccaattgaacctcttccccctcttacat tctgtccacccttcttcgaaaatcaaaccac	upstream of the start, down-stream of the stop codon of *tent-5* (d)

Primers for genotyping of the previously published mouse lines (with reference)			
	Tent5c KO-Fw	aggtcctgactgaggtcgtg	(*21*)
	Tent5c KO-Rev	ttcctcaaaatccccgtaca	
	Tent5a-3xFLAG-Fw	tggtaaacgagagtacggtg	(*26*)
	Tent5a-3xFLAG-Rev	ctggctgctttttaccaaca	
	Tent5c-3xFLAG-Fw	gaccctgaacctcatctctc	(*21*)
	Tent5c-3xFLAG-Rev	tatgggtccttttagggttg	
sgRNAs, ssDNA and primers used for *Tent5a*$^{Flox/Flox}$ cKO mouse line			
	sgRNA_LoxP_left	tatgggcgtcacgatcggggagg	guide RNA sequence
	sgRNA_LoxP_right	actaatgcgcgtgagtggtgtgg	
	ssDNA donor	GCTGGGCATTGGGGTCGGTGAGGGTTGGG GGAGGAGGGCTACTATGGGCGTCACGATC GGATAACTTCGTATAGCATACATTATACG AAGTTATGGatccAGGGGAAGCCCAGGCCC CTcggcgggaccctgtcccctcacccgcacacctctctgccgct ccctaaatccactctttcctcttagGACTGACCAAGGCC AAGTGGCTTTCGGAGGGCACTACATGGCC GAGGGCGAAGGGTACTTTGCCATGGCAG AGGACGAGCTGACCGGCGGCCCTTACATC CCCCTGGGTGGTGACTTCGGCGGCGGCGG CAGCAGCTTCGGCGACCGCTGCTCGGACT ATTGCGAAAGCCCCACGGCGCACTGCAAT GTGCTGAACTGGGAGCAAGTGCAGCGGTT GGACGGCATTCTGAGCGAGACCATCCCGA TCCACGGGCGCGGCAACTTCCCCACGCTC GAGCTGCAGCCAAGCCTGATTGTGAAGGT GGTGAGGAGGCGCCTGGAAGAGAAGGGC ATAGGTGTCCGCGACGTGCGCCTCAACGG CTCAGCCGCCAGTCACGTCCTGCACCAGG ACAGTGGCCTAGGCTACAAGGACCTGGAC CTCATCTTCTGCGCTGACCTGCGTGGGGA AGAGGAGTTTCAGACTGTGAAGGACGTCG TTCTGGATTGCCTGTTAGACTTCTTGCCCG AAGGGGTGAACAAGGAGAAGATCACACC GCTCACTCTCAAGgtaaagccacctgatgggcttgggca gagcaggggtggggagaaagcaagcccagcttgccttgactgccg tcagtatggccatagtgagttcgtttatttatttttttattatttttttttcacg agcgcaCTAATGCGCGTGAGTGGatccATAAC TTCGTATAGCATACATTATACGAAGTTAT TGTGGAGTGTTTTAGTTTCCTGGGAGATC CAGCGCCACAACTCTGTGCTTCCTCCTCTG CGTTTC	Description: *LoxP* sites are underlined, BamHI sites are highlighted in grey
	Tent5a_cKO_sFw	actctcgcttgtgctttcca	sequencing
	Tent5a_cKO_sRv	tcagttctgcttcaggctgc	
	Tent5a_cKO_gFw	caagcctgattgtgaaggtg	Mice genotyping
	Tent5a_cKO_gRv	aaggaagagaaggaaacgca	
	For genotyping of *Tent5a* cKO mice, PCR products generated with Tent5a_cKO_gFw/gRv primers were digested with BamHI: • WT: undigested 489 bp, digested 489 bp • cKO: undigested 527 bp, digested: 117 bp		
	A_FloX_F1	actctcgcttgtgctttcca	BMDM genotyping
	A_FloX_F2	tcagactgtgaaggacgtcg	
	A_FloX_R1	tcagttctgcttcaggctgc	

Table S3. List of DNA constructs used in this study.

pDNA	Source	Reference
pCFJ90-*Pmyo-2::mCherry::unc-54utr*	Dr. J. J. Ewbank	Addgene plasmid # 19327; see ref. (*119*)
pCFJ104-*Pmyo-3::mCherry::unc-54utr*	Dr. J. J. Ewbank	Addgene Plasmid # 19328; see ref. (*119*)
pCFJ151	Dr. J. J. Ewbank	Addgene Plasmid # 19330; see ref. (*119*)
pDD162 (*Peft-3::Cas9* + Empty sgRNA)	Dr. B. Goldstein	Addgene Plasmid # 47549; see ref. (*120*)
pDD282	Dr. B. Goldstein	Addgene Plasmid # 66823; see ref. (*120*)
pDD287	Dr. B. Goldstein	Addgene Plasmid # 70685; see ref. (*120*)
pRH269-*Pmyo-2::GFP::unc-54utr*	Dr. K. Drabikowski	
pCIneo-NHA	Dr. Witold Filipowicz	see ref. (*104*)
pRL-5BoxB	Dr. Witold Filipowicz	see ref. (*104*)
pMD2.G	Dr. Didier Trono	Addgene Plasmid # 12259
psPAX2	Dr. Didier Trono	Addgene Plasmid # 12260
pCAG-Cre-IRES2-GFP	Dr. Jacek Jaworski	Addgene Plasmid # 26646; see ref. (*121*)
pCIneo-*NHA-tent-5*^{WT}	This work	
pCIneo-*NHA-tent-5*^{MUT} (MUT = D151A, D153A in *tent-5* isoform a)	This work	
pDD162-*sgRNA425* (pVL060)	This work	
pDD162-*sgRNA456* (pVL069)	This work	
pDD282-*tent-5::gfp::3xflag* (pVL062)	This work	
pDD287-*vgln-1::mKate2::3xmyc* (pVL070)	This work	

Table S4. Protein sequences used for the phylogenetic analysis.

Organism	Protein	WormBase or UniProt ID
Caenorhabditis elegans		
	TENT-5	F55A12.9.d1
	GLD-2	ZC308.1c.1
	GLD-4	ZK858.1
	USIP-1	ZK863.4
	PUP-1/CID-1	K10D2.3
	PUP-2	K10D2.2
	PUP-3	F59A3.9
	MUT-2	K04F10.6b.1
Homo sapiens		
	TENT5A	Q96IP4
	TENT5B	Q96A09
	TENT5C	Q5VWP2
	TENT5D	Q8NEK8
	TENT2	Q6DFA8
	TENT4A	Q5XG87
	TENT4B	Q8NDF8
	TENT6	Q9NVV4
	TENT1	Q9H6E5
	TUT4	Q5TAX3
	TUT7	Q5VYS8
Mus musculus		
	TENT5A	A0A087WS27
	TENT5B	Q8C152
	TENT5C	Q5SSF7
	TENT5D	B1ATX6
	TENT2	Q91YI6
	TENT4A	Q6PB75
	TENT4B	Q68ED3
	TENT6	Q9D0D3
	TENT1	Q8R3F9
	TUT4	B2RX14
	TUT7	Q5BLK4
Xenopus laevis		
	TENT2-a	Q641A1
	TENT2-b	Q6DFA8
	TENT4B	A0A1L8GKF9
	TUT4	A0A1L8GMG3
	TUT7	A0A1L8HZ12
	TENT5A	Q08AY7
	TENT5B	Q66J95
	TENT5D	A0A1L8F355

Appendix 2

Manuscript 2

Mackiewicz Z., Liudkovska V., Dziembowski A. TENT-5 regulates the expression of male-specific genes in *Caenorhabditis elegans*. *bioRxiv* (2024). DOI: 10.1101/2024.06.18.599341

Warsaw, 7.01.2025

Zuzanna Mackiewicz
Laboratory of RNA Biology
International Institute of Molecular and Cell Biology
4 Ks. Trojdena
02-109 Warsaw

PHD CANDIDATE'S
CONTRIBUTION STATEMENT

Article's title: TENT-5 regulates the expression of male-specific genes in *Caenorhabditis elegans*.

Authors: Zuzanna Mackiewicz, Vladyslava Liudkovska, Andrzej Dziembowski

Journal: bioRxiv

Date of publishing: 18th June 2024

DOI: 10.1101/2024.06.18.599341

I hereby declare that my contribution to the article titled *"TENT-5 regulates the expression of male-specific genes in Caenorhabditis elegans"* was as following:

- designing experiments
- conducting all experiments
- analyzing all Nanopore RNA sequencing data
- writing the initial draft of the manuscript and creating all figures

.....
PhD candidate signature

.....
Corresponding author signature

TENT-5 regulates the expression of male-specific genes in *Caenorhabditis elegans*

Zuzanna Mackiewicz[1], Vladyslava Liudkovska[2], Andrzej Dziembowski[1]

[1]Laboratory of RNA Biology, International Institute of Molecular and Cell Biology, Trojdena 4, 02-109 Warsaw, Poland
[2]Laboratory of Stem Cell RNA Metabolism, IMol Polish Academy of Sciences, Flisa 6, 02-247 Warsaw, Poland

Abstract

Polyadenylation is an important post-transcriptional process that governs mRNA stability and expression. Advancements in direct RNA sequencing in recent years have clarified many aspects of this intricate regulation, revealing the influence of various factors. Here, we used Nanopore Direct RNA Sequencing to investigate the association between genome-wide mRNA poly(A) tail profiles and sexual dimorphism in *Caenorhabditis elegans*. Our results demonstrate sex-dependent differences in both gene expression and poly(A) tail metabolism. Notably, we discovered that cytoplasmic poly(A) polymerase TENT-5 regulates multiple male-specific transcripts, predominantly encoding putative seminal fluid components with predicted extracellular localization. TENT-5 expression in male-specific tissues, such as seminal vesicle and vas deferens, corroborates its functional significance. Intriguingly, despite extensive TENT-5-mediated polyadenylation of male-specific transcripts, males devoid of TENT-5 show no abnormalities in mating behavior, sperm morphology, or fertility. Our findings suggest that TENT-5 plays a role in regulating sex-related processes in males, although the physiological consequences remain to be fully elucidated.

Introduction

Polyadenylation is a crucial post-transcriptional process that modifies RNA species by adding adenosines to their 3′ ends, thereby influencing RNA stability and function within the cell. This modification primarily occurs in the nucleus, catalyzed by canonical poly(A) polymerase, and is essential for mRNA biogenesis and subsequent export to the cytoplasm (1, 2). Upon reaching the cytoplasm, mRNAs associate with poly(A) binding proteins, which are critical for maintaining mRNA stability and enabling efficient protein synthesis (2, 3). During the later lifespan of mRNA, poly(A) tails undergo further modifications. Deadenylating enzymes, such as CCR4-NOT or PAN2/3, gradually shorten poly(A) tails, directing mRNAs to degradation (2, 4, 5). However, poly(A) tails can also be re-adenylated in the cytoplasm by noncanonical poly(A) polymerases (ncPAPs), which are members of the terminal nucleotidyltransferase family of proteins (TENTs) (6, 7). Cytoplasmic polyadenylation enhances stability, promotes the expression of selected transcripts, and prevents their degradation, contributing to the complex dynamics of poly(A) tails and regulation of gene expression (6, 7).

Recent advances in RNA sequencing technologies have enabled more comprehensive studies of poly(A) tail metabolism (8). Genome-wide poly(A) tail profiling has been carried out for various organisms, including yeast (9, 10), nematodes (11–13), mice (10, 14–16), and humans (10, 12, 17). These studies revealed organism-specific characteristics of poly(A) tails and their dependence on multiple factors such as developmental stage, age, tissue localization, and response to environmental cues. While poly(A) tail length is often positively associated with mRNA half-life and translation efficiency, this relationship is not always consistent. During early embryogenesis, longer poly(A) tails enhance mRNA stability and translation (2, 10, 18). However, in somatic cells, poly(A) tail length often shows a poor or even negative correlation

with mRNA expression levels and translational activity (10–12). For instance, mRNAs encoding highly abundant housekeeping machinery, such as ribosomal proteins, typically have short tails. Conversely, longer tails are predominantly found on lowly expressed mRNAs, which often encode regulatory proteins (11, 19). Moreover, poly(A) tail dynamics is affected by additional layers of regulatory control, including poly(A)-modifying enzymes that influence numerous physiological processes. For example, cytoplasmic polyadenylation by GLD2 (Germ Line Development 2)/TENT2 is indispensable for proper germline development and fertility in nematodes (20–22) and flies (23, 24). In mammals, a similar role is observed for the recently discovered TENT5 ncPAPs (15). Additionally, members of the TENT5 family have been implicated in regulating humoral and innate immune responses by stabilizing mRNAs encoding immunoglobulins and secreted defense proteins (13, 16). Despite these findings, the high complexity of cytoplasmic polyadenylation requires further examination to elucidate the interplay between poly(A)-modifying enzymes, the mechanisms driving their substrate specificity, and their effects on various groups of transcripts during multiple physiological processes and conditions (6).

Here, we demonstrate that sexual dimorphism is a critical factor that strongly influences polyadenylation profiles. For our studies, we used *Caenorhabditis elegans*, which features two distinct sexes: self-fertile hermaphrodites (XX) and males (X0), comprising approximately 0,05% of the population. Utilizing Oxford Nanopore Direct RNA Sequencing (DRS), the most reliable methodology for studying poly(A) tails, we uncovered profound differences in gene expression and poly(A) length distributions between hermaphrodites and males. Notably, we showed that TENT-5 ncPAP is responsible for the regulation of male-specific transcripts that mainly encode secreted seminal fluid components. Our work provides new and interesting insights into the polyadenylation process, expanding our knowledge of both TENT-5 functions and *C. elegans* male physiology and genetics.

Methods

C. elegans culture and growth conditions

The following *C. elegans* strains were used: wild type N2 Bristol, BS553 (*fog-2(oz40) V*), and DR94 (*unc-45(m94) III*), obtained from the Caenorhabditis Genetics Center (CGC); *tent-5(tm3504) I*, obtained from the National Bioresource Project of Japan (NBRP); and ADZ21 (*tent-5(rtt6[tent-5::gfp::3xflag]) I*), previously generated in our laboratory (13). All strains were cultured at 20°C on nematode growth medium (NGM) plates seeded with *E. coli* HB101 as a food source. Male enrichment of N2 and *tent-5(tm3504) C. elegans* cultures was achieved through mating, with one hermaphrodite and five males of the same genotype placed on plates with a drop of *E. coli* HB101. Successful mating resulted in up to 50% male enrichment.

Preparation of samples for Direct RNA sequencing (DRS)

Wild-type or *tent-5(tm3504)* young adult males were hand-picked from the male-enriched plates and transferred to an Eppendorf tube containing 10 µl of 50 mM NaCl. The worm pellet was then resuspended in 200 µl of TRI Reagent (Sigma-Aldrich), vortexed for 15 min at room temperature, and stored at -80°C before RNA isolation. This process was repeated three times to obtain sufficient biological material. TRIzol samples from three different pickings were combined just before RNA isolation, resulting in approximately 300 male worms per sample. The wild-type hermaphrodite population was age-synchronized by bleaching gravid adults and seeding isolated embryos on plates. Worms were grown until they reached the young adult stage, then collected and washed three times using 50 mM NaCl. The worm pellet was resuspended in 1 ml of TRI Reagent (Sigma-Aldrich), vortexed for 15 min at room temperature, and stored at -80°C. RNA was isolated from TRIzol according to the manufacturer's

instructions. The samples were additionally purified with KAPA Pure magnetic beads (RNA to beads ratio: 1:3 v/v), and RNA quality was assessed using the Agilent TapeStation system. Samples for wild-type and *tent-5(tm3504)* males and hermaphrodites were prepared in two independent biological replicates.

Direct RNA sequencing and data analysis

Sequencing libraries were prepared using the DRS Kit (Oxford Nanopore Technologies, SQK-RNA002) according to the manufacturer's protocol. In each library, mRNA isolated from male or hermaphrodite worms (0,5 - 1,5 µg depending on the amount of isolated RNA) was mixed with 2 ng of *in vitro* transcribed poly(A) standards. Sequencing was performed with the MinION device, followed by basecalling with Guppy 6.0.0 (ONT), mapping to WBCel235 reference (MiniMap v2.17 (25) with options -k 14 -ax map-ont –secondary = no), and processing with samtools v1.9 (26). Lengths of poly(A) tails were estimated with Nanopolish (v0.13.2) as described before (13, 16). Poly(A) length distributions were compared between conditions for transcripts with a minimum of 5 reads per condition using the Wilcoxon test. Adjusted *p*-values were calculated using the Benjamini-Hochberg method. Differential expression analysis was performed using DESeq2 v1.28 Bioconductor package (27). GO enrichment analysis was performed using WormBase Enrichment Suite (28). The subcellular localization of proteins was assessed with DeepLoc v2.0 (29). DRS data have been deposited to the European Nucleotide Archive (ENA) with the following accession numbers: ERS20270802, ERS20271308, ERS20270803, ERS20271309, ERS20270804, ERS20270805, ERS20227048, and ERS20270807.

The dataset from Ebbing A. *et al.* (2018) for comparisons of sex-enriched genes was reanalyzed from raw *fastq* files deposited by authors in Gene Expression Omnibus (GEO). Reads were mapped to WBCel235 reference (STAR v2.7.10a (30)) and followed by processing with samtools v1.9 (26). Read counts were collected using featureCounts from the Subread package (v2.0.6 (31) with options -Q 10 -p -C -B and WBCel235 annotation). Differential expression analysis was performed using DESeq2 v1.28 Bioconductor package (27).

Spermatids isolation

A day before the spermatids isolation, a few wild-type or *tent-5(tm3504)* L4 stage males were transferred to fresh NGM plates seeded with *E. coli* HB101. Separating males from hermaphrodites allows spermatids to accumulate in the male's gonad before the experimental procedure (32). Males were dissected the next day in a drop of sperm medium (50 mM HEPES, 50 mM NaCl, 25 mM KCl, 5 mM $CaCl_2$, 1 mM $MgSO_4$, and 1 mg/ml BSA) by cutting the posterior end of the male with a needle. The isolated spermatids were visualized immediately (32, 33).

Microscopic observations

Isolated spermatids were imaged using the Nikon Eclipse 80i microscope with a 100×/oil immersion lens. The spermatids' size was assessed using ImageJ2 software (34) and compared between wild-type and *tent-5(tm3504)* males using a two-tailed unpaired Student's *t*-test. To evaluate the localization of TENT-5 in males, *tent-5::gfp* worms were immobilized with 25 µM levamisole on freshly prepared 2% agarose pads and immediately imaged on the Zeiss LSM800 confocal microscope with a 40×/oil immersion lens.

Fertility and mating behavior experiments

To assess male fertility, wild-type or *tent-5(tm3504)* L4 stage males were crossed with L4 stage *fog-2(oz40)* females (which are unable to produce their own sperm (35)). For each cross, one female and four males were placed on an NGM plate with a 5-mm diameter *E. coli* lawn.

Crosses were performed at 20°C and continued throughout the entire female lifespan. The number of progeny was counted and compared between wild-type and *tent-5(tm3504)* males using a two-tailed unpaired Student's *t*-test. For comparison of mating behavior, wild-type or *tent-5(tm3504)* young adult males were crossed with young adult *unc-45(m94)* hermaphrodites (uncoordinated hermaphrodites were used to increase mating efficiency (36)). For each cross, ten hermaphrodites and five males were placed on an NGM plate seeded with *E. coli* HB101 and recorded for 30 minutes. Recorded movies were subsequently analyzed, and mating behavior was described according to the following criteria: 1 – number of different hermaphrodites touched by the male, 2 – number of contacts with any hermaphrodite, 3 – backward locomotion after touching the hermaphrodite with the tail, 4 – turning around the hermaphrodite's body, 5 – successful location of the hermaphrodite's vulva, and 6 – successful spicule insertion (36, 37). The number of each incident was then compared between wild-type and *tent-5(tm3504)* males using a two-tailed unpaired Student's *t*-test (with or without Welch's correction depending on variance distributions).

Results

Male-enriched genes are expressed mainly in sperm, seminal vesicle, and vas deferens

It is widely recognized that gene expression displays significant variability depending on the sex of the organism, contributing to differences in appearance and behavior. Studies of sex-specific gene expression have been conducted in various organisms, including the nematode *C. elegans* (38–42). Here, we used Oxford Nanopore Direct RNA Sequencing (DRS) to expand on previous studies by simultaneously investigating differences in both gene expression and mRNA poly(A) tail length dynamics between young adult *C. elegans* wild-type males and hermaphrodites. Our DRS analysis revealed substantial differences in gene expression between the two sexes, with 1538 genes exhibiting sex-dependent expression patterns (Figure 1A, Supplementary Table S1). Among these, 1203 genes were enriched in males, while 335 were enriched in hermaphrodites. To validate our findings, we compared our results with previously published datasets that utilized different approaches to male collection and sample preparation. For instance, Kim B. et al. (2016) identified 1739 genes enriched in *him-5* mutant males using Illumina sequencing (42), of which 802 genes overlapped with our study (Figure 1B, Supplementary Table S1). Similarly, Ebbing A. et al. (2018) detected 3529 genes upregulated in wild-type males (41), with 1090 genes overlapping with our dataset (Figure 1B, Supplementary Table S1). The substantial overlap between male-enriched genes across different studies underscores the robustness of our approach and the reliability of our data for elucidating sex-specific gene regulation in *C. elegans*.

We then analyzed the gene sets enriched in males and hermaphrodites in more detail. The Gene Ontology (GO) term analysis of male-enriched genes revealed a distinctive association with phosphatase and kinase activities (Figure 1C and D), consistent with previous reports identifying these as characteristic features of sperm-related genes (42–45). Indeed, among male-enriched genes, we detected a large population encoding sperm components: 30 genes from the *msp* family, 17 from the *ssp* family, and 9 from the *nspd* family (Supplementary Table S1). While both *C. elegans* sexes generate sperm, males exhibit a prolonged sperm production period, resulting in higher overall sperm levels (46). This accounts for the observed upregulation of sperm genes in males compared to hermaphrodites (Figure 1C). Intriguingly, apart from the sperm genes, many other male-enriched genes are exclusively expressed in males and encode proteins with unknown functions, presenting an opportunity for further exploration in *C. elegans* male research. Subsequent analyses of genes specific to males allowed us to categorize them into two major groups: seminal vesicle genes and vas deferens genes (Figure

1C). The seminal vesicle and vas deferens are male-specific tissues that are responsible for producing seminal fluid and transporting it into the hermaphrodite uterus during mating (41). Despite the crucial role of these tissues in reproduction, the precise molecular regulatory mechanisms governing their functions are still not fully understood. However, it is known that due to their secretory function, male-specific tissues are rich in endoplasmic reticulum (ER) (47). In line with that, we observed that 82% of male-specific transcripts (defined as those with no expression in hermaphrodites, FDR < 0.05) are predicted to contain a signal peptide-encoding sequence targeting them to the secretory pathway through the ER (Supplementary Figure S1, Supplementary Table S1).

The pool of hermaphrodite-enriched genes and their overlap with datasets from other studies appeared relatively small compared to that of males (Figure 1A and B). However, the genes we identified as enriched closely align with hermaphrodite reproductive functions. As expected, among the genes highly enriched in hermaphrodites (Supplementary Table S1), we found many implicated in hermaphrodite reproduction, germline function, and early embryo development. Notably, all six members of the vitellogenin (*vit*) family exhibited high enrichment in hermaphrodites (Figure 1E). These genes are predominantly expressed in the intestine of young adult and adult hermaphrodites, encoding yolk proteins responsible for transporting essential nutrients to developing embryos within the gonad (48). Similarly, genes from the *egg* family (e.g., *egg-1*, *egg-2*, *egg-5*, and *egg-6*), *perm* family (e.g., *perm-2*, *perm-4*), and proteins involved in RNA regulation (e.g., *gld-1, lin-41, puf-5, puf-7, puf-11, mex-5, mex-6*), which are all known to influence germ cell proliferation, differentiation, and germline function in hermaphrodites (49), showed substantial enrichment compared to males. Interestingly, we also observed significant differences in the expression of genes encoding ribosomal proteins (*rpl* and *rps*) between the sexes (Figure 1E and F). Some of these genes (e.g., *rpl-10*, *rpl-3*, *rpl-7A*, *rpl-9, rpl-34, rps-8, rps-11, rps-12, rps-3*) were proposed to play critical roles in germline and embryonic development (49). More broadly, ribosomal proteins are essential for protein synthesis and cell growth. We suspect that their differential expression between sexes might reflect differences in cellular metabolism and growth between hermaphrodites and males.

In conclusion, our DRS results reveal profound differences in gene expression between *C. elegans* sexes, consistent with previous studies (41, 42). Our approach provides a more robust basis for examining sex-related transcriptomes by analyzing a large population of wild-type males, thus avoiding the potential influence of the *him* (high incidence of males) mutation background on gene expression and poly(A) tail distributions. Notably, we observed the most pronounced sex-dependent expression for genes encoding poorly studied secreted proteins localized in male-specific tissues.

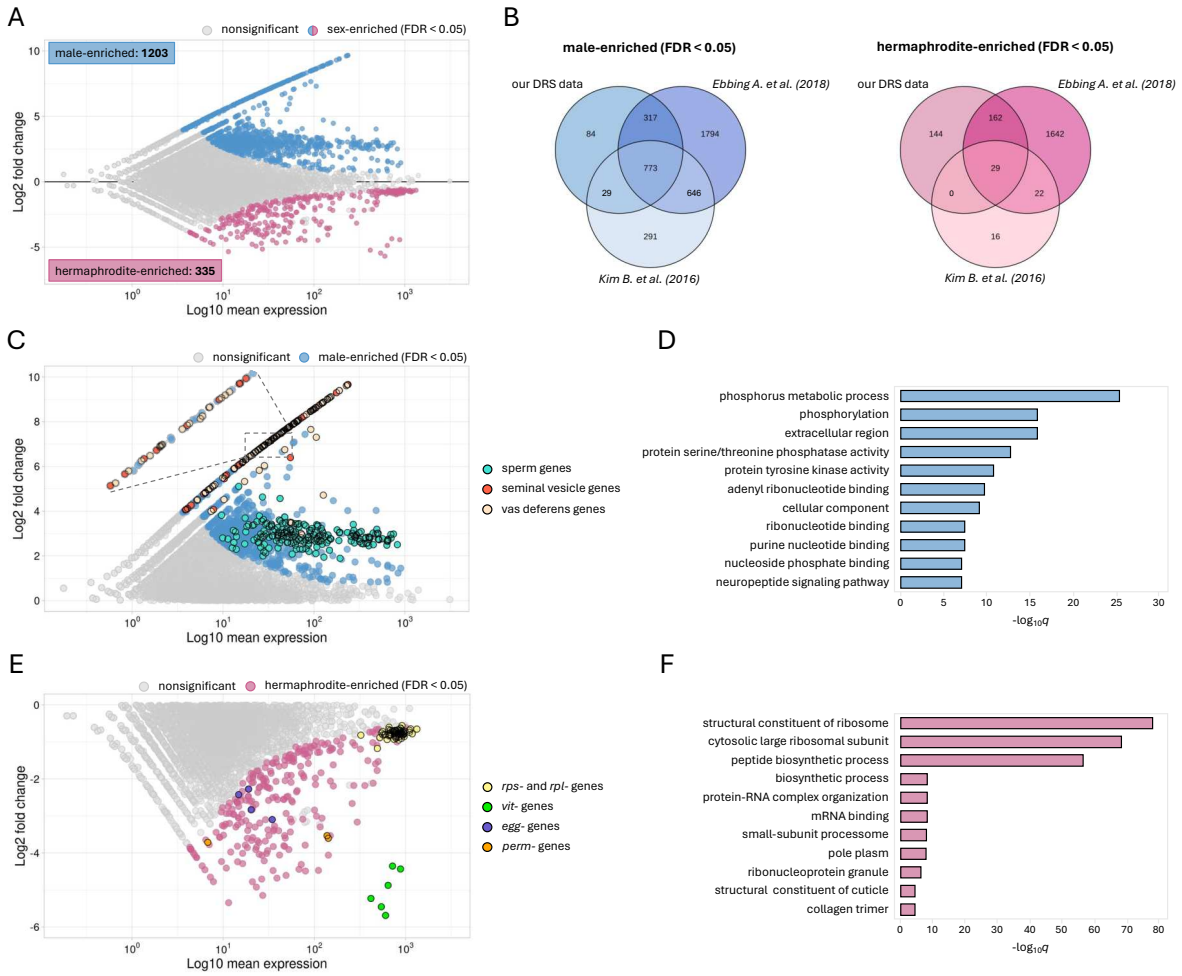


Figure 1. Differences in gene expression patterns between wild-type males and hermaphrodites.
(A) MA plot illustrating differential gene expression between males and hermaphrodites of N2 wild-type worms. Significantly changed genes (FDR < 0.05) are marked with blue and pink dots for male-enriched and hermaphrodite-enriched genes, respectively.
(B) Venn diagrams showing overlaps between our DRS data and two other studies comparing transcriptomes of *C. elegans* males and hermaphrodites (Kim B. *et al.* 2016 and Ebbing A. *et al.* 2018) (41, 42). For each compared group, the significantly enriched pool of genes was included according to the cut-off FDR value set in each study. The overlap for male-enriched genes is shown in blue and for hermaphrodite-enriched genes in pink.
(C) Zoom on the half of the MA plot showing genes upregulated in males. Significantly changed genes (FDR < 0.05) are marked with blue dots. Among these, three of the most interesting groups of genes were selected: sperm genes (cyan dots), seminal vesicle genes (red dots), and vas deferens genes (beige dots). Groups of genes were defined based on Ebbing A. *et al.* 2018 (41). A zoomed-in fragment of the plot is shown for clearer visualization of overlapping dots.
(D) Top GO terms for the male-enriched genes ordered by adjusted *p*-value (WormBase Enrichment Suite).
(E) Zoom on the half of the MA plot showing genes upregulated in hermaphrodites. Significantly changed genes (FDR < 0.05) are marked with pink dots. Four groups of genes are marked: ribosomal genes (*rpl* and *rps*, yellow dots), vitellogenin genes (*vit*, green dots), *egg* genes (purple dots), and *perm* genes (orange dots).
(F) Top GO terms for the hermaphrodite-enriched genes ordered by adjusted *p*-value (WormBase Enrichment Suite).

Poly(A) tail distributions differ for males and hermaphrodites

Given that the poly(A) tail length often correlates with a transcript's stability and expression level (8), we hypothesized that significant differences would be observed between the two sexes of *C. elegans* not only in gene expression but also in mRNA poly(A) tail length distribution. Utilizing the same DRS-based dataset of mRNAs isolated from hermaphrodites and males as described above, we found global mRNA poly(A) tail length profiles consistent with previous

studies (Supplementary Table S2) (11–13, 50). Additionally, we identified a slight difference between the sexes, with global median poly(A) tail lengths of 51 nt for hermaphrodites and 55 nt for males (Figure 2A). Further analysis of poly(A) tail length distributions for differentially polyadenylated mRNAs (at least 5 nt difference between sexes, FDR < 0.05) revealed 208 transcripts with significantly longer poly(A) tails in hermaphrodites and 155 transcripts with longer poly(A) tails in males (Figure 2B, Supplementary Figure S2A and B, Supplementary Table S2). Interestingly, in males, a significant portion of transcripts with longer tails are lowly expressed (Figure 2C). This group of mRNAs primarily contributes to the longer median poly(A) tail length observed in males (Figure 2D). In contrast, many transcripts with longer tails in hermaphrodites encode highly abundant proteins, such as ribosomal proteins, collagens, or actin (Figure 2C and E, Supplementary Figure S2C, Supplementary Table S2). These genes were also upregulated in hermaphrodites compared to males based on our differential expression analysis (Figure 1E and F, Supplementary Table S1). We speculate that the elongation of tails for these transcripts may be attributed to the increased demand for their respective proteins to support embryo development in the hermaphrodite's gonad.

We observed that in males, poly(A) tail elongation is directed towards transcripts functionally associated with the endoplasmic reticulum (Figure 2G), suggesting the importance of efficient secretory machinery for proper male physiology. Since most male-specific genes encode proteins secreted by the ER (Supplementary Figure S1A, Supplementary Table S1), we analyzed their poly(A) tails. We found that globally, male-specific mRNAs exhibit significantly longer poly(A) tails compared to other transcripts expressed in males (Figure 2F, left panel). This observation can be explained by the much higher proportion of lowly expressed genes in this group, which generally have longer poly(A) tails than highly expressed ones (Figure 2D). Further comparisons of poly(A) tails within groups of genes with similar expression levels revealed no significant difference in median tail length between male-specific and other male-expressed genes (Figure 2F, middle and right panels). To investigate further, we also explored the relationship between poly(A) tail lengths and mRNA expression levels. For both hermaphrodites and males, we observed a negative correlation between mRNA tail length and expression level (Figure 2D and F), which is consistent with previous reports (11). Moreover, we found that transcripts significantly upregulated in hermaphrodites also carry globally longer poly(A) tails in hermaphrodites than in males (Figure 2H). A similar trend was observed for male-enriched transcripts (Figure 2H), which proves the importance of poly(A) tail in regulating gene expression levels.

Finally, due to the distinct chromosome compositions of males and hermaphrodites, we examined the global poly(A) length distributions based on gene location within each chromosome. Surprisingly, we identified differences in poly(A) profiles for transcripts from chromosomes IV and X (Figure 2I, Supplementary Figure S2D). Specifically, mRNAs from chromosome IV exhibited significantly longer poly(A) tails in males compared to hermaphrodites, whereas transcripts from chromosome X showed the opposite effect. Although we compared multiple features of genes encoded on each chromosome (expression levels, presence of sequence encoding signal peptides, GC content, transcript length, or 5′ UTR and 3′ UTR length) (Supplementary Figure S3), we were unable to explain how observed sex-dependent polyadenylation is influenced by gene chromosome location.

In addition to the significant differences in gene expression between males and hermaphrodites, our DRS results unveil distinct patterns of poly(A) tail length distribution between the two sexes. These findings underscore the important role of poly(A) metabolism in sex-dependent physiological processes.

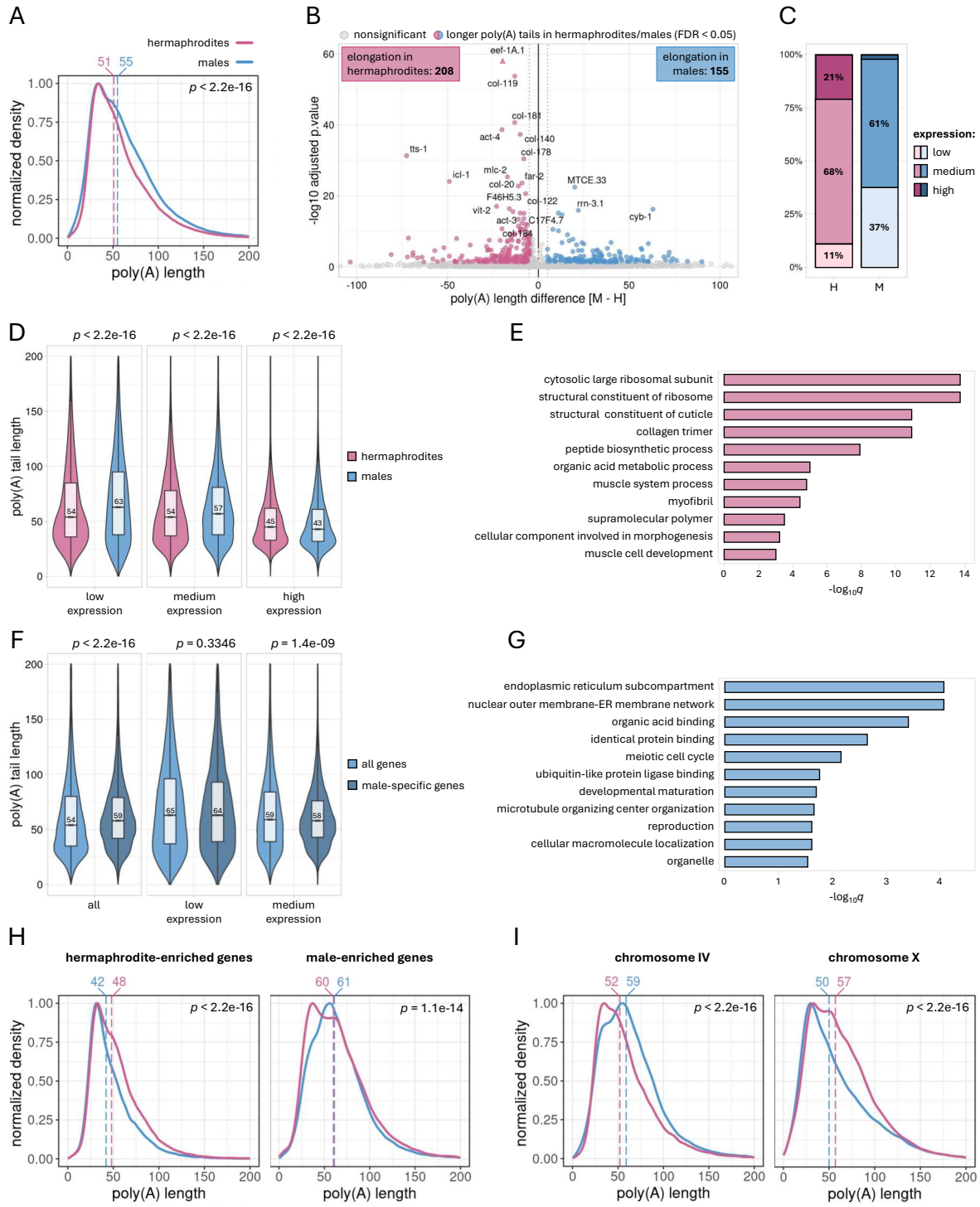


Figure 2. Differences in poly(A) distributions between wild-type males and hermaphrodites.
(A) Density plot showing global differences in the poly(A) tail distribution between males (blue line) and hermaphrodites (pink line). Vertical dashed lines represent the median poly(A) tail length for each condition (in nucleotides). The plot was generated for all identified transcripts and normalized to 1.
(B) Volcano plot displaying differential polyadenylation between males and hermaphrodites for N2 wild-type worms. Genes with significantly changed median poly(A) tail length (FDR < 0.05) by a minimum of 5 nucleotides (dotted line) are marked with blue dots (elongated in males) and pink dots (elongated in hermaphrodites). Triangles represent data points outside the *y*-axis limit. "M" in the *x*-axis title represents males and "H" represents hermaphrodites.
(C) Fractions of genes with elongated poly(A) tails in hermaphrodites (pink bars) or males (blue bars), characterized by low, medium, or high expression levels. Expression levels are defined by baseMean values presented in Supplementary Table S1. Genes with baseMean < 20 are defined as lowly expressed, with 20 <

baseMean < 500 as medium expressed, and with baseMean > 500 as highly expressed. "M" in the *x*-axis title represents males and "H" represents hermaphrodites.
(D) Violin plots showing distributions of poly(A) tail length for genes grouped by their expression levels as defined in panel **C**. Numbers inside the boxes indicate the median poly(A) tail length for each condition.
(E) Top GO terms for genes with poly(A) tails elongated in hermaphrodites ordered by adjusted *p*-value (WormBase Enrichment Suite).
(F) Violin plots illustrating the distribution of poly(A) tail lengths for male-specific genes compared to other male-expressed genes. Male-specific genes are defined as those with no expression in hermaphrodites based on our DRS data (Supplementary Table S1). The left plot represents the distribution for all male-specific genes contrasted with all others, while the middle and right plots represent similar comparisons within groups of genes with similar expression levels as defined in panel **C**. No male-specific genes met our criteria for highly expressed genes. Numbers inside the boxes indicate the median poly(A) tail length for each condition.
(G) Top GO terms for genes with poly(A) tails elongated in males ordered by adjusted *p*-value (WormBase Enrichment Suite).
(H) Density plots showing differences in the poly(A) tail distribution for sex-enriched genes between males (blue line) and hermaphrodites (pink line). Vertical dashed lines represent the median poly(A) tail length for each condition (in nucleotides). Plots were generated for all transcripts significantly upregulated in either hermaphrodites or males based on Supplementary Table S1 and normalized to 1.
(I) Density plots showing differences in the poly(A) tail distribution for all genes localized on chromosome IV or X between males (blue line) and hermaphrodites (pink line). Vertical dashed lines represent the median poly(A) tail length for each condition (in nucleotides). Plots were generated for all transcripts assigned to a particular chromosome and normalized to 1.

TENT-5 regulates the expression of male-specific genes

Within the cell, poly(A) tails can undergo modification by various proteins that either lengthen or shorten them (4–6). We hypothesized that the observed differences in poly(A) distributions between *C. elegans* males and hermaphrodites might result from the sex-specific activities of these proteins. To identify a potential candidate protein, we analyzed our DRS data to determine if the expression of known poly(A) polymerases and deadenylases varied between the sexes. We observed no significant differences in gene expression levels for any of these poly(A)-modifying enzymes (Figure 3A, Supplementary Table S1).

Our previous work in hermaphrodite *C. elegans* demonstrated that the polyadenylation of mRNAs encoding ER-targeted secreted proteins is mediated by the noncanonical poly(A) polymerase TENT-5 (13). Depletion of TENT-5 in hermaphrodites resulted in poly(A) tail shortening and downregulation of mRNAs encoding secreted innate immunity effector proteins, leading to increased susceptibility of *tent-5* mutants to pathogen infection (13). As mentioned above, our functional enrichment analysis showed that transcripts with longer poly(A) tails in males compared to hermaphrodites are associated with ER functionality (Figure 2F). Therefore, we assumed that TENT-5 might be responsible for their polyadenylation and thus contribute to the differential polyadenylation observed between the sexes, particularly evident for lowly expressed genes. Additionally, we hypothesized that since most male-specific transcripts are predicted to encode secreted proteins, TENT-5 might also regulate their expression, which could have been overlooked in our previous studies on hermaphrodites.

Using the *tent-5::gfp* knock-in strain, we examined TENT-5 localization in males. Confocal microscopy revealed a strong expression of TENT-5-GFP in the male intestine and head neurons (Figure 3B), consistent with previous observations in hermaphrodites (13). Notably, we detected a pronounced localization of TENT-5 in spermatids, seminal vesicles, and vas deferens (Figure 3B, Supplementary Figure S4A), where most of the male-enriched genes are expressed (Figure 1C). This observation suggested that, indeed, TENT-5 may regulate male-specific transcripts through their polyadenylation. Of note, our microscopic observations revealed intense autofluorescence at the male tail tip and copulatory spicules (Figure 3B,

Supplementary Figure S4B), an intriguing feature previously briefly mentioned in the literature (51–53).

To identify TENT-5 substrates in *C. elegans* males and validate the hypothesis that TENT-5 regulates male-specific mRNAs and other transcripts encoding secreted proteins, we performed DRS on *tent-5* mutant and wild-type males, analyzing gene expression and poly(A) patterns. We observed significant dysregulation of 33 genes upon TENT-5 deletion in males: the expression levels of 32 genes were downregulated, and one was upregulated (Figure 3C, Supplementary Table S3). Among the downregulated genes, 21 were consistent with our previous findings in *tent-5*-deficient hermaphrodites (Figure 3E, Supplementary Table S3) (13), showing similar functional enrichment in defense response processes (Figure 3D). This underscores TENT-5's role in *C. elegans* innate immunity, independently of nematode sex. However, poly(A) profiling of *tent-5* mutant males revealed more pronounced sex-related differences. Despite no global change in the poly(A) distribution between wild-type and *tent-5* mutant males (similar to our previous observations for hermaphrodites (13)) (Figure 4A), we detected 137 transcripts with altered poly(A) tail length (changed by at least 5 nt, FDR < 0.05), of which 119 had shorter and 18 longer poly(A) tails in *tent-5* mutant males (Figure 4B, Supplementary Figure S4C, Supplementary Table S4). Strikingly, among mRNAs with shortened poly(A) tails in mutants, 100 were male-specific, and 12 were male-enriched, accounting for 94% of mRNAs with shortened poly(A) tails in *tent-5* mutant males (Figure 4B, Supplementary Table S4). As expected, the majority of these transcripts are predicted to encode extracellular proteins (93%), a hallmark of TENT-5 substrates (13, 15, 16, 54, 55), which is even more pronounced in males than in hermaphrodites (Figure 4C and D, Supplementary Table S4).

Approximately 66% of TENT-5 targets encode putative components of seminal fluid expressed in male-specific tissues (Supplementary Table S4), suggesting the importance of TENT-5-mediated polyadenylation in regulating certain aspects of male reproduction. For instance, some of these substrates possess serine-type endopeptidase activity (Figure 4D), which has been previously linked with sperm activation in *C. elegans* (33, 56). Although the process of sperm activation is not well understood in nematodes, it is presumed that inactive spermatids stored in seminal vesicle are transported through the male's vas deferens and cloaca into the hermaphrodite's uterus. After ejaculation, spermatids are activated by seminal fluid proteins to form functional sperm (57). So far, only a few genes involved in this process have been reported (33, 56). Interestingly, one of them (*try-5*) is polyadenylated by TENT-5 based on our results (Supplementary Figure S4D, Supplementary Table S4).

Among the mRNAs with shortened poly(A) tails in *tent-5* mutant males, we also detected 22 members of the *clec* family (C-type lectins) (Supplementary Figure S4D, Supplementary Table S4), which are reported to be immune effectors (58, 59). The enrichment of C-type lectins in the seminal fluid suggests its possible function in the nematode's defense response. Recent studies have shown that *C. elegans* males are more resistant to fungal, viral, and bacterial infections, as well as to heat, osmotic, and oxidative stress (60). Also in other species, seminal fluid has been shown to have antimicrobial properties, helping to protect both female and male reproductive tracts against infections (61, 62). This raises the interesting hypothesis that *C. elegans* males utilize seminal fluid genes for both defense response and mating purposes (57). However, the exact mechanism of that possible parallel regulation remains to be established.

Additionally, we analyzed how differential gene expression correlates with poly(A) tail length. Although only six genes displayed both significant downregulation and shorter poly(A) tails in

mutant males (Figures 3E and 4E), we observed a trend where TENT-5-mediated poly(A) tail elongation tends to stabilize the corresponding mRNA (Figure 4E), a pattern also noted in hermaphrodites (13). This finding suggests that the correlation between mRNA poly(A) tail length and expression level might be largely tissue-specific or specific to certain transcript types or groups. Surprisingly, the overlap between transcripts with shortened poly(A) tails in *tent-5* mutant males and hermaphrodites was minimal (Figures 3E and 4F), with only three genes (*heh-1*, *far-2,* and *clec-209*) emerging as shared TENT-5 targets across sexes (Supplementary Figure S4D, Supplementary Table S4). Moreover, *heh-1* and *far-2* were also significantly downregulated in both sexes (Figure 3E). Although not well studied, these genes are predicted to encode proteins responsible for sterol or fatty acid binding (49). Interestingly, male-specific tissues, where TENT-5 activity in males is most pronounced, are known to be cholesterol-rich (63). Therefore, we speculate that TENT-5 might be involved in cholesterol metabolism in both sexes by stabilizing *heh-1* and *far-2* mRNAs.

Overall, TENT-5 activity in males focuses almost exclusively on polyadenylating male-specific transcripts, whereas in hermaphrodites, it appears much more versatile (Figure 4F). The functions of most male-specific TENT-5 substrates remain unknown, making it difficult to speculate about TENT-5's physiological role in males. However, we expect that the elongation of poly(A) tails of so many putative seminal fluid components in males is crucial for their reproduction.

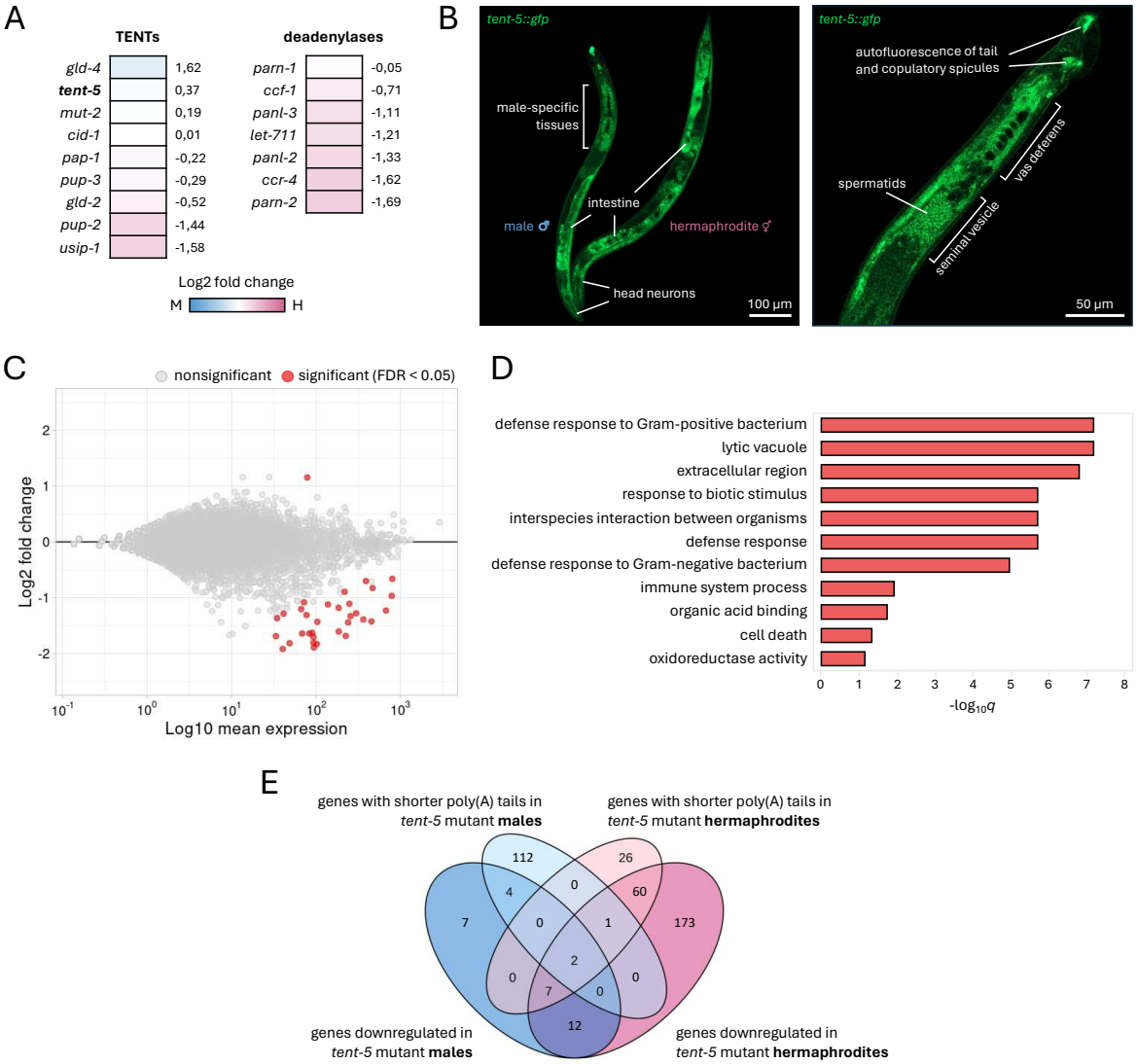


Figure 3. Differences in gene expression patterns between *tent-5* mutant and wild-type males.
(A) Heatmap showing expression levels of TENTs and deadenylases between males and hermaphrodites. The color of the heatmap specifies whether a gene is upregulated in males (blue) or hermaphrodites (pink). The numbers on the right represent the log2 fold change value derived from Supplementary Table S1. "M" in the heatmap legend represents males and "H" represents hermaphrodites.
(B) Fluorescence microscopic images of TENT-5-GFP expression in hermaphrodites and males (*tent-5(rtt6[tent-5::gfp::3xflag]* I). The left image represents differences and similarities in TENT-5 expression between sexes. It is composed of six separate shots to show the full body of a male and hermaphrodite. The right image shows the tail-end of a male body. Images were taken with 40x lens.
(C) MA plot illustrating differential gene expression between wild-type and *tent-5* mutant worms. Significantly changed genes (FDR < 0.05) are marked with red dots.
(D) Top GO terms for genes significantly downregulated in *tent-5* mutant males ordered by adjusted *p*-value (WormBase Enrichment Suite).
(E) Venn diagram showing overlaps between our DRS data and our previous study on *tent-5* mutant hermaphrodites (Liudkovska V. *et al.* 2022) (13). Four different conditions were compared: genes downregulated in *tent-5* mutant males (blue), genes with shorter poly(A) tails in *tent-5* mutant males (light blue), genes downregulated in *tent-5* mutant hermaphrodites (pink), and genes with shorter poly(A) tails in *tent-5* mutant hermaphrodites (light pink). Included were all significantly downregulated genes (FDR < 0.05) (for differential expression) and all transcripts with significantly shorter poly(A) tail by a minimum of 5 nucleotides (FDR < 0.05) (for differential polyadenylation).

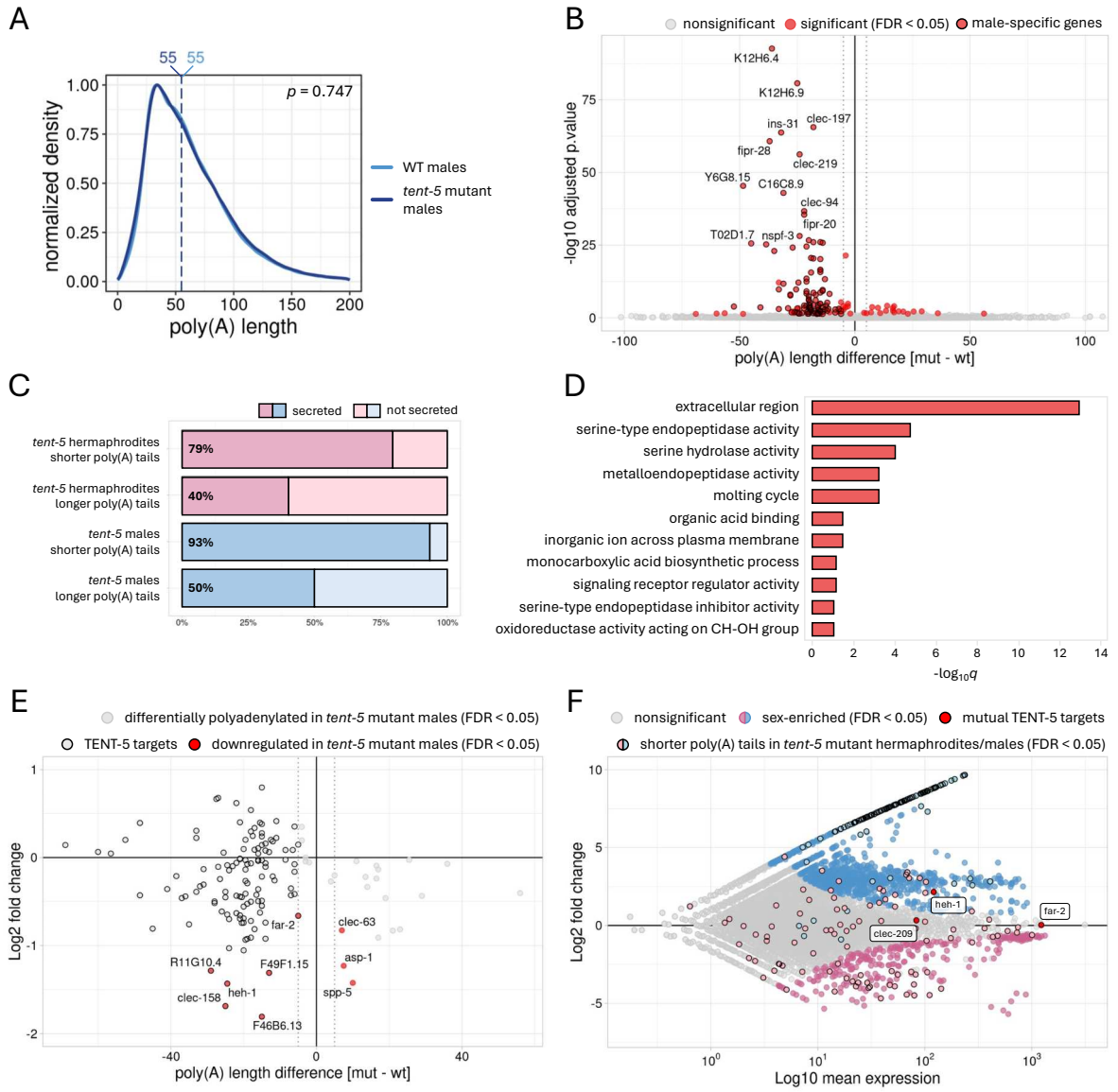


Figure 4. TENT-5 polyadenylates male-specific mRNAs encoding secreted proteins.
(A) Density plot showing global differences in the poly(A) tail distribution between wild-type (blue) and *tent-5* mutant males (dark blue). Vertical dashed lines represent the median poly(A) tail length for each condition (in nucleotides). The plot was generated for all identified transcripts and normalized to 1.
(B) Volcano plot showing differential polyadenylation between wild-type and *tent-5* mutant males. Transcripts with significantly changed median poly(A) tail length (FDR < 0.05) are marked with red dots. Black borderline is added to male-specific genes. Male-specific genes were all genes that had no expression in hermaphrodites based on our DRS data (Supplementary Table S1).
(C) Fractions of genes with predicted extracellular localization by DeepLoc 2.0 software (29). For each condition, a group of significantly shortened/elongated transcripts by a minimum of 5 nucleotides was included. Data for hermaphrodites comes from Liudkovska V. *et al.* 2022 (13).
(D) Top GO terms for genes with significantly shorter poly(A) tails in *tent-5* mutant males (by a minimum of 5 nucleotides) ordered by adjusted *p*-value (WormBase Enrichment Suite).
(E) The relationship between the difference in the expression level and the difference in the median poly(A) tail length of respective transcripts for *tent-5* mutant compared to wild-type males. TENT-5 targets (genes with significantly shorter poly(A) tail in *tent-5* mutant worms by a minimum of 5 nucleotides; FDR < 0.05) are highlighted with black borderline. Genes significantly downregulated in *tent-5* mutant males (FDR < 0.05) are marked with red dots.
(F) MA plot showing differential gene expression between males and hermaphrodites of N2 wild-type worms (Figure 1A). Significantly changed genes (FDR < 0.05) are marked with blue and pink dots for male-enriched and hermaphrodite-enriched genes, respectively. Additionally, marked are TENT-5 targets (genes with significantly shorter poly(A) tail in *tent-5* mutant worms by a minimum of 5 nucleotides; FDR < 0.05) in males (light blue,

black borderline) and hermaphrodites (light pink, black borderline; based on Liudkovska V. *et al.* 2022). Red dots with labels represent common TENT-5 targets for both sexes.

TENT-5-mediated poly(A) tail elongation does not affect male mating behavior

Following our observation that TENT-5 is responsible for the polyadenylation of male-specific transcripts in *C. elegans* males, we attempted to identify the resulting physiological phenotype. We hypothesized that since TENT-5 targets comprise presumed components of the seminal fluid, the phenotype of *tent-5* mutants might exhibit abnormal mating phenotypes. Firstly, we monitored the mating behaviors of wild-type and *tent-5* mutant males. We did not detect any significant differences in the frequency of male contact with hermaphrodites, characteristic backward movements, vulva searching, or spicule insertion (Figure 5A, Supplementary Movie S1). The significant difference between wild-type and mutant males was noticed only for the turning behavior. However, we cannot exclude that this observation results from a random variance between individual worm behavior rather than from the TENT-5 depletion. The absence of the obvious mating behavioral phenotype indicates that *tent-5* mutant males have no impaired neuronal or pheromonal signaling. Next, we assumed that post-transcriptional changes in male-specific transcripts might influence the potency of the seminal fluid in properly activating male sperm post-ejaculation. To investigate this, we assessed male fertility by counting the offspring produced by *fog-2* females crossed with either wild-type or *tent-5* mutant males. Progeny numbers were comparable between the two conditions (Figure 5B), which suggests that TENT-5 does not significantly influence male mating behavior or efficiency.

Knowing that some seminal fluid components are involved in the process of sperm activation (33, 56), we hypothesized that their TENT-5-mediated regulation might influence the morphology of spermatids stored in the male's seminal vesicle. However, our microscopic observations did not show any alteration in spermatids morphology (Figure 5C, Supplementary Figure S5) and size (Figure 5D). We also did not detect any incidents of matured sperm, that were activated before the ejaculation. Nevertheless, considering the strong modifications in seminal fluid components observed at the transcriptome level, we would expect that these changes should manifest as some physiological phenotype, that were not examined in this study. More complex *in vivo* studies would need to be performed to elucidate which physiological processes are regulated by TENT-5 poly(A) polymerase in males. Unfortunately, as mentioned above, most of the predicted TENT-5 targets in males do not have any functional annotation, which makes it harder to anticipate their possible function in male physiology.

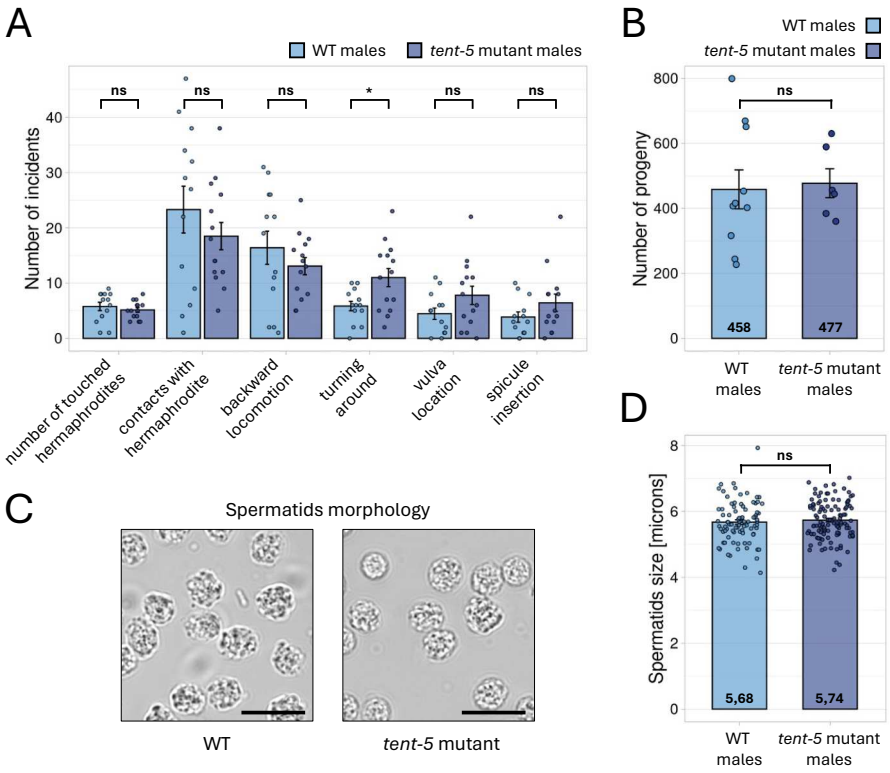


Figure 5. Comparison of mating behavior, fertility, and spermatids morphology between *tent-5* mutant and wild-type males.
(A) Differences between the number of mate-related incidents performed by wild-type (n=15) and *tent-5* mutant (n=14) males. Barplots represent mean values with SD. ns => not significant; * => *p*-value < 0.05 (two-tailed *t*-test).
(B) Differences in brood sizes between wild-type (n=10) and *tent-5* mutant (n=6) males crossed with *fog-2* females. Barplots represent mean values with SD. ns => not significant (two-tailed *t*-test).
(C) Microscopic images of spermatids morphology for wild-type and *tent-5* mutant males. Images were taken with 100x magnification. Scale bars => 10 microns.
(D) Differences in spermatids diameter between wild-type (n=86) and *tent-5* mutant (n=114) males. Barplots represent mean values with SD. ns => not significant (two-tailed *t*-test).

Discussion

Advancements in methods that allow precise genome-wide estimation of the poly(A) tail length and composition have significantly boosted our understanding of poly(A) tail metabolism and its role in gene regulation (8). However, it quickly became evident that this regulation is highly cell- and tissue-type specific, and extremely dynamic in response to environmental and developmental cues. Therefore, it should be explored in comprehensive ways that can capture these dynamic changes. Several excellent studies leveraged these novel methodologies to profile poly(A) tail length during oocyte maturation in *Drosophila melanogaster* (18) and *Xenopus laevis* (64), mammalian oocyte-to-embryo transition (65), and embryonic development of *Danio regio*, *X. laevis*, and *C. elegans* (10, 12, 50). However, sex-related differences in global poly(A) metabolism remain largely overlooked in most organisms. In this study, we employed Direct RNA Sequencing to investigate poly(A) tail distribution differences between hermaphrodite and male *C. elegans*, aiming to understand their underlying reasons and consequences.

We observed clear, previously reported differences in gene expression between males and hermaphrodites, reflecting the distinct cellular metabolism, growth, and physiology of adult hermaphrodites and males. Additionally, our poly(A) profiling revealed a distinct sex-

dependent distribution of poly(A) tail length. In general, our data align with previous observations that highly abundant housekeeping genes possess relatively short poly(A) tails (11, 19). We confirmed this negative correlation between mRNA poly(A) tail length and expression level for *C. elegans* males and hermaphrodites. Notably, highly expressed mRNAs have slightly shorter tails in males compared to hermaphrodites, whereas lowly expressed mRNAs have significantly longer tails in males. Interestingly, sex-enriched mRNAs (those expressed in both sexes but upregulated in one) tend to have longer tails in the respective sex. The mechanism behind this sex-dependent difference remains unclear. We also noted chromosomal differences in poly(A) distributions. Transcripts encoded by genes on the X chromosome have longer poly(A) tails in hermaphrodites, while those derived from chromosome IV have longer tails in males. The X chromosome is enriched with female transcripts and largely devoid of male-enriched ones (43, 66), likely leading to the observed global elongation of poly(A) tails for X chromosome transcripts in hermaphrodites. Conversely, the opposite trend for chromosome IV can be explained by the presence of many male-enriched, spermatogenesis-related genes (67). Another notable observation concerns male-specific transcripts. We found that 82% of these transcripts are predicted to contain signal peptide-encoding sequences targeting them to the secretory pathway through the ER. Interestingly, most of these transcripts have medium to low expression levels. When compared with transcripts of the same expression levels, male-specific mRNAs show no difference in their median poly(A) tail length. However, compared to all mRNAs expressed in males, male-specific transcripts tend to have longer poly(A) tails.

The mechanistic reasons for the sex-specific differences in poly(A) tail distributions are not yet fully understood. Multiple factors influence mRNA poly(A) tail length, with the enzymatic activities of deadenylases, noncanonical poly(A) polymerases, and terminal uridylyltransferases playing primary roles (4–6). The steady-state poly(A) distribution at the transcriptome scale is achieved through the cell- or tissue-specific concentrations of these enzymes and their varying affinities modulated by auxiliary factors toward particular transcripts. We discovered that for a specific group of mRNAs, which predominantly encode secreted components of seminal fluid, the poly(A) tail length is affected by the TENT-5 ncPAP. TENT-5 localizes to male reproductive tissues rich in ER, such as spermatids, seminal vesicle, and vas deferens, where it potentially stabilizes male-specific mRNAs. In hermaphrodites, the same enzyme is mainly expressed in the intestine and enhances the expression of genes encoding secreted innate immune effectors (13). Therefore, in both *C. elegans* sexes, TENT-5 polyadenylates transcripts encoding proteins with predicted extracellular localization, underscoring the notion that the TENT5 family of ncPAPs is essential for post-transcriptional regulation of mRNAs encoding proteins directed towards the ER secretory pathway (13, 15, 16, 54). As described in our previous work (13), the exact mechanism behind TENT-5-substrate specificity is not yet clear. However, as a fraction of TENT-5 resides in the ER, we proposed that TENT-5 specificity is directly or indirectly driven by its colocalization with ER-targeted mRNAs (13). We have previously predicted that cytoplasmic polyadenylation by TENT-5 would be most pronounced in tissues with high secretion capacity. Consequently, in males, it is not surprising that TENT-5 focuses on regulating male-specific genes expressed primarily in the ER-rich seminal vesicle and vas deferens. Furthermore, our data strongly suggests that TENT-5 affects the composition of seminal fluid by modulating its transcriptome, potentially playing a role in male reproduction. However, despite intensive efforts, we have not identified phenotypes related to fertility or mating behavior in TENT-5-deficient males, possibly because these phenotypes manifest only under specific environmental conditions. We also cannot exclude the possibility that other ncPAPs may redundantly function with TENT-5 specifically in male tissues. One potential candidate is GLD-4, which was slightly enriched in males

according to our DRS (Figure 3A, Supplementary Table S1). It has been shown that while GLD-4 alone partially reduces the fertility of hermaphrodites, it does not affect spermatogenesis and male fertility (21). In the future, it would be interesting to analyze the poly(A) distribution in males devoid of both TENT-5 and GLD-4.

From an evolutionary perspective, it is interesting to examine the role of noncanonical cytoplasmic poly(A) polymerases in reproduction and gametogenesis across different animal species. In *C. elegans*, GLD-2 is a crucial regulator of the germline mitosis-to-meiosis transition and subsequent gametogenesis (20). Consequently, GLD-2 depletion in nematodes results in defective oocytes and sperm production, leading to infertility in both hermaphrodites and males (22). Similarly, two homologs of GLD-2 in *D. melanogaster*, Gld2 and Wispy, are essential for proper gamete production in a sex-dependent manner (23, 24). Gld2 is expressed exclusively in the male germline and is required for the completion of spermatogenesis; its absence results in complete sterility in male flies (23). Wispy, on the other hand, regulates late oogenesis and is crucial for female fly fertility (24). Surprisingly, the knockout of the mouse homolog of GLD-2, TENT2, does not cause abnormalities in gamete production or affect reproduction (68). In mammals, the TENT5 family of ncPAPs, which includes four members (TENT5A-D), plays an essential role in spermatogenesis and oogenesis (15). In mice, TENT5C and TENT5D are expressed in spermatocytes and spermatids, regulating different aspects of spermiogenesis (15, 69, 70). Depletion of either TENT5C or TENT5D results in complete male infertility (15). Additionally, TENT5C and TENT5B regulate oogenesis, and the double knockout mutation of both proteins leads to infertility (15). In contrast, TENT-5 poly(A) polymerase in worms is not essential for hermaphrodite fertility (13) and only slightly impacts the male reproductive system. These findings indicate that the role of cytoplasmic polyadenylation is highly conserved across species. However, different species may rely on the activities of different TENTs to varying extents.

Further research is needed to uncover the complexity of poly(A) tail metabolism at the organismal level, not only in gametogenesis and reproduction but also in many other physiological processes. Our DRS analysis represents an important step toward expanding the current understanding of cytoplasmic polyadenylation. Additionally, we believe our results provide a foundation for further studies on *C. elegans* male physiology.

Data availability

The Nanopore Direct RNA Sequencing data have been deposited to the European Nucleotide Archive (ENA) with the following accession numbers: ERS20270802, ERS20271308, ERS20270803, ERS20271309, ERS20270804, ERS20270805, ERS20227048, and ERS20270807.

Acknowledgments

We are grateful to the members of the Dziembowski group for their support and insightful comments. We thank Paweł Krawczyk and Natalia Gumińska for the assistance with bioinformatic analyses, Aleksandra Brouze and Seweryn Mroczek for running Nanopore sequencing on a subset of samples, Bartosz Tarkowski for discussing ideas, Wojciech Pokrzywa and Anwesha Sarkar for sharing *unc-45(m94)* strain, and Tomasz Węgierski for the support with microscopic observations.

Funding

This work was supported by the National Science Center (OPUS 17 UMO-2019/33/B/NZ2/01773) and Horizon Europa (European Research (AdG nr 101097317). Some *C. elegans* strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440).

Author contributions

A.D. provided funding and directed the project. Z.M, V.L, and A.D designed the experiments. Z.M conducted all *C. elegans* experiments and analyzed the sequencing data. Z.M drafted the initial manuscript. V.L and A.D revised and edited the manuscript.

References

1. Jalkanen,A.L., Coleman,S.J. and Wilusz,J. (2014) Determinants and implications of mRNA poly(A) tail size - Does this protein make my tail look big? *Semin Cell Dev Biol*, **34**, 24–32.
2. Passmore,L.A. and Coller,J. (2022) Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. *Nat Rev Mol Cell Biol*, **23**, 93–106.
3. Wigington,C.P., Williams,K.R., Meers,M.P., Bassell,G.J. and Corbett,A.H. (2014) Poly(A) RNA-binding proteins and polyadenosine RNA: New members and novel functions. *Wiley Interdiscip Rev RNA*, **5**, 601–622.
4. Yan,Y. Bin (2014) Deadenylation: Enzymes, regulation, and functional implications. *Wiley Interdiscip Rev RNA*, **5**, 421–443.
5. Zhang,X., Virtanen,A. and Kleiman,F.E. (2010) To polyadenylate or to deadenylate: That is the question. *Cell Cycle*, **9**, 4437–4449.
6. Liudkovska,V. and Dziembowski,A. (2021) Functions and mechanisms of RNA tailing by metazoan terminal nucleotidyltransferases. *Wiley Interdiscip Rev RNA*, **12**.
7. Yu,S. and Kim,V.N. (2020) A tale of non-canonical tails: gene regulation by post-transcriptional RNA tailing. *Nat Rev Mol Cell Biol*, **21**, 542–556.
8. Brouze,A., Krawczyk,P.S., Dziembowski,A. and Mroczek,S. (2023) Measuring the tail: Methods for poly(A) tail profiling. *Wiley Interdiscip Rev RNA*, **14**.
9. Tudek,A., Krawczyk,P.S., Mroczek,S., Tomecki,R., Turtola,M., Matylla-Kulińska,K., Jensen,T.H. and Dziembowski,A. (2021) Global view on the metabolism of RNA poly(A) tails in yeast Saccharomyces cerevisiae. *Nat Commun*, **12**.
10. Subtelny,A.O., Eichhorn,S.W., Chen,G.R., Sive,H. and Bartel,D.P. (2014) Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature*, **508**, 66–71.
11. Lima,S.A., Chipman,L.B., Nicholson,A.L., Chen,Y.H., Yee,B.A., Yeo,G.W., Coller,J. and Pasquinelli,A.E. (2017) Short poly(A) tails are a conserved feature of highly expressed genes. *Nat Struct Mol Biol*, **24**, 1057–1063.
12. Legnini,I., Alles,J., Karaiskos,N., Ayoub,S. and Rajewsky,N. (2019) FLAM-seq: full-length mRNA sequencing reveals principles of poly(A) tail length control. *Nat Methods*, **16**, 879–886.
13. Liudkovska,V., Krawczyk,P.S., Brouze,A., Gumińska,N., Wegierski,T., Cysewski,D., Mackiewicz,Z., Ewbank,J.J., Drabikowski,K., Mroczek,S., *et al.* (2022) TENT5 cytoplasmic noncanonical poly(A) polymerases regulate the innate immune response in animals. *Sci. Adv*, **8**.
14. Begik,O., Diensthuber,G., Liu,H., Delgado-Tejedor,A., Kontur,C., Niazi,A.M., Valen,E., Giraldez,A.J., Beaudoin,J.D., Mattick,J.S., *et al.* (2023) Nano3P-seq: transcriptome-wide analysis of gene expression and tail dynamics using end-capture nanopore cDNA sequencing. *Nat Methods*, **20**, 75–85.

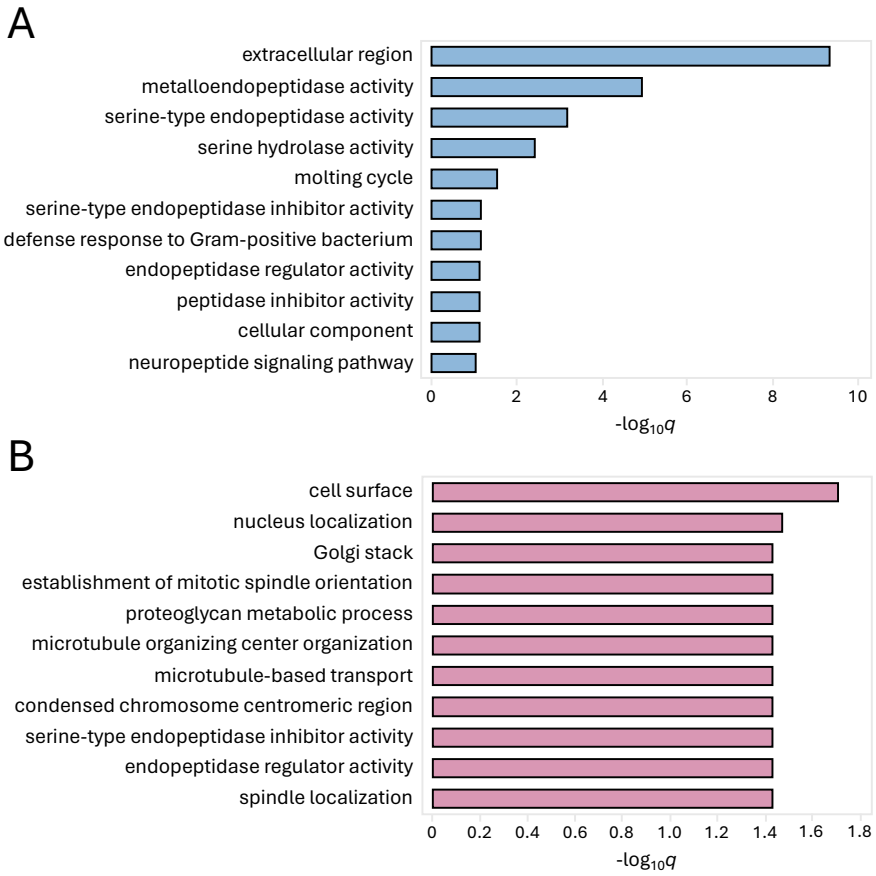
15. Brouze,M., Czarnocka-Cieciura,A., Gewartowska,O., Kusio-Kobiałka,M., Jachacy,K., Szpila,M., Tarkowski,B., Gruchota,J., Krawczyk,P.S., Mroczek,S., *et al.* (2023) Polyadenylation of mRNAs encoding secreted proteins by TENT5 family of enzymes is essential for gametogenesis in mice. *bioRxiv*, doi: https://doi.org/10.1101/2023.08.01.551432.
16. Bilska,A., Kusio-Kobiałka,M., Krawczyk,P.S., Gewartowska,O., Tarkowski,B., Kobyłecki,K., Nowis,D., Golab,J., Gruchota,J., Borsuk,E., *et al.* (2020) Immunoglobulin expression and the humoral immune response is regulated by the non-canonical poly(A) polymerase TENT5C. *Nat Commun*, **11**.
17. Chang,H., Lim,J., Ha,M. and Kim,V.N. (2014) TAIL-seq: Genome-wide determination of poly(A) tail length and 3' end modifications. *Mol Cell*, **53**, 1044–1052.
18. Lim,J., Lee,M., Son,A., Chang,H. and Kim,V.N. (2016) MTAIL-seq reveals dynamic poly(A) tail regulation in oocyte-to-embryo development. *Genes Dev*, **30**, 1671–1682.
19. Castellano,L.A. and Bazzini,A.A. (2017) Poly(A) tails: Longer is not always better. *Nat Struct Mol Biol*, **24**, 1010–1011.
20. Nousch,M., Minasaki,R. and Eckmann,C.R. (2017) Polyadenylation is the key aspect of GLD-2 function. *RNA*, **23**, 1180–1187.
21. Schmid,M., Küchler,B. and Eckmann,C.R. (2009) Two conserved regulatory cytoplasmic poly(A) polymerases, GLD-4 and GLD-2, regulate meiotic progression in C. elegans. *Genes Dev*, **23**, 824–836.
22. Boag,P.R., Harrison,P.F., Barugahare,A.A., Pattison,A.D., Swaminathan,A., Raymant,G., Monk,S., Tsyganov,K., Heinz,E., Davis,G.M., *et al.* (2018) Widespread cytoplasmic polyadenylation programs asymmetry in the germline and early embryo. *bioRxiv*, doi: https://doi.org/10.1101/428540.
23. Sartain,C. V., Cui,J., Meisel,R.P. and Wolfner,M.F. (2011) The poly(A) polymerase GLD2 is required for spermatogenesis in Drosophila melanogaster. *Development*, **138**, 1619–1629.
24. Benoit,P., Papin,C., Kwak,J.E., Wickens,M. and Simonelig,M. (2008) PAP- and GLD-2-type poly(A) polymerases are required sequentially in cytoplasmic polyadenylation and oogenesis in Drosophila. *Development*, **135**, 1969–1979.
25. Li,H. (2018) Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, **34**, 3094–3100.
26. Danecek,P., Bonfield,J.K., Liddle,J., Marshall,J., Ohan,V., Pollard,M.O., Whitwham,A., Keane,T., McCarthy,S.A. and Davies,R.M. (2021) Twelve years of SAMtools and BCFtools. *Gigascience*, **10**.
27. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, **15**.
28. Angeles-Albores,D., Raymond,R.Y., Chan,J. and Sternberg,P.W. (2016) Tissue enrichment analysis for C. elegans genomics. *BMC Bioinformatics*, **17**.
29. Thumuluri,V., Almagro Armenteros,J.J., Johansen,A.R., Nielsen,H. and Winther,O. (2022) DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. *Nucleic Acids Res*, **50**, W228–W234.
30. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
31. Liao,Y., Smyth,G.K. and Shi,W. (2014) FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.
32. Singaravelu,G., Chatterjee,I., Marcello,M.R. and Singson,A. (2010) Isolation and in vitro activation of Caenorhabditis elegans sperm. *Journal of Visualized Experiments*, **47**.

33. Stanfield,G.M. and Villeneuve,A.M. (2006) Regulation of sperm activation by SWM-1 is required for reproductive success of C. elegans males. *Current Biology*, **16**, 252–263.
34. Rueden,C.T., Schindelin,J., Hiner,M.C., DeZonia,B.E., Walter,A.E., Arena,E.T. and Eliceiri,K.W. (2017) ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*, **18**.
35. Schedl,T. and Kimble,J. (1988) fog-2, a Germ-Line-Specific Sex Determination Gene Required for Hermaphrodite Spermatogenesis in Caenorhabditis elegans. *Genetics*, **119**, 43–61.
36. Gower,N.J.D., Walker,D.S. and Baylis,H.A. (2005) Inositol 1,4,5-trisphosphate signaling regulates mating behavior in Caenorhabditis elegans males. *Mol Biol Cell*, **16**, 3978–3986.
37. Bever,B.W., Dietz,Z.P., Sullins,J.A., Montoya,A.M., Bergthorsson,U., Katju,V. and Estes,S. (2022) Mitonuclear Mismatch is Associated With Increased Male Frequency, Outcrossing, and Male Sperm Size in Experimentally-Evolved C. elegans. *Front Genet*, **13**.
38. Thoemke,K., Yi,W., Ross,J.M., Kim,S., Reinke,V. and Zarkower,D. (2005) Genome-wide analysis of sex-enriched gene expression during C. elegans larval development. *Dev Biol*, **284**, 500–508.
39. Miersch,C. and Döring,F. (2012) Sex Differences in Carbohydrate Metabolism Are Linked to Gene Expression in Caenorhabditis elegans. *PLoS One*, **7**.
40. Jiang,M., Ryu,J., Kiraly,M., Duke,K., Reinke,V. and Kim,S.K. (2001) Genome-wide analysis of developmental and sex-regulated gene expression profiles in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U.S.A*, **98**, 218–223.
41. Ebbing,A., Vértesy,Á., Betist,M.C., Spanjaard,B., Junker,J.P., Berezikov,E., van Oudenaarden,A. and Korswagen,H.C. (2018) Spatial Transcriptomics of C. elegans Males and Hermaphrodites Identifies Sex-Specific Differences in Gene Expression Patterns. *Dev Cell*, **47**, 801–813.
42. Kim,B., Suo,B. and Emmons,S.W. (2016) Gene Function Prediction Based on Developmental Transcriptomes of the Two Sexes in C. elegans. *Cell Rep*, **17**, 917–928.
43. Cutter,A.D., Morran,L.T. and Phillips,P.C. (2019) Males, outcrossing, and sexual selection in Caenorhabditis nematodes. *Genetics*, **213**, 27–57.
44. West,S.M., Mecenas,D., Gutwein,M., Aristizábal-Corrales,D., Piano,F. and Gunsalus,K.C. (2018) Developmental dynamics of gene expression and alternative polyadenylation in the Caenorhabditis elegans germline. *Genome Biol*, **19**.
45. Ma,X., Zhu,Y., Li,C., Xue,P., Zhao,Y., Chen,S., Yang,F. and Miao,L. (2014) Characterisation of Caenorhabditis elegans sperm transcriptome and proteome. *BMC Genomics*, **15**.
46. Singson,A. (2006) Sperm activation: Time and tide wait for no sperm. *Current Biology*, **16**.
47. Lints,R. and Hall,D.H. (2009) Male reproductive system, somatic gonad. In *WormAtlas*, doi: 10.3908/wormatlas.2.15.
48. Perez,M.F. and Lehner,B. (2019) Vitellogenins - Yolk Gene Function and Regulation in Caenorhabditis elegans. *Front Physiol*, **10**.
49. WormBase, https://wormbase.org
50. Roach,N.P., Sadowski,N., Alessi,A.F., Timp,W., Taylor,J. and Kim,J.K. (2020) The full-length transcriptome of C. elegans using direct RNA sequencing. *Genome Res*, **30**, 299–312.
51. Yi,W., Ross,J.M. and Zarkower,D. (2000) Mab-3 is a direct tra-1 target gene regulating diverse aspects of C. elegans male sexual development and behavior. *Development*, **127**, 4469–4480.

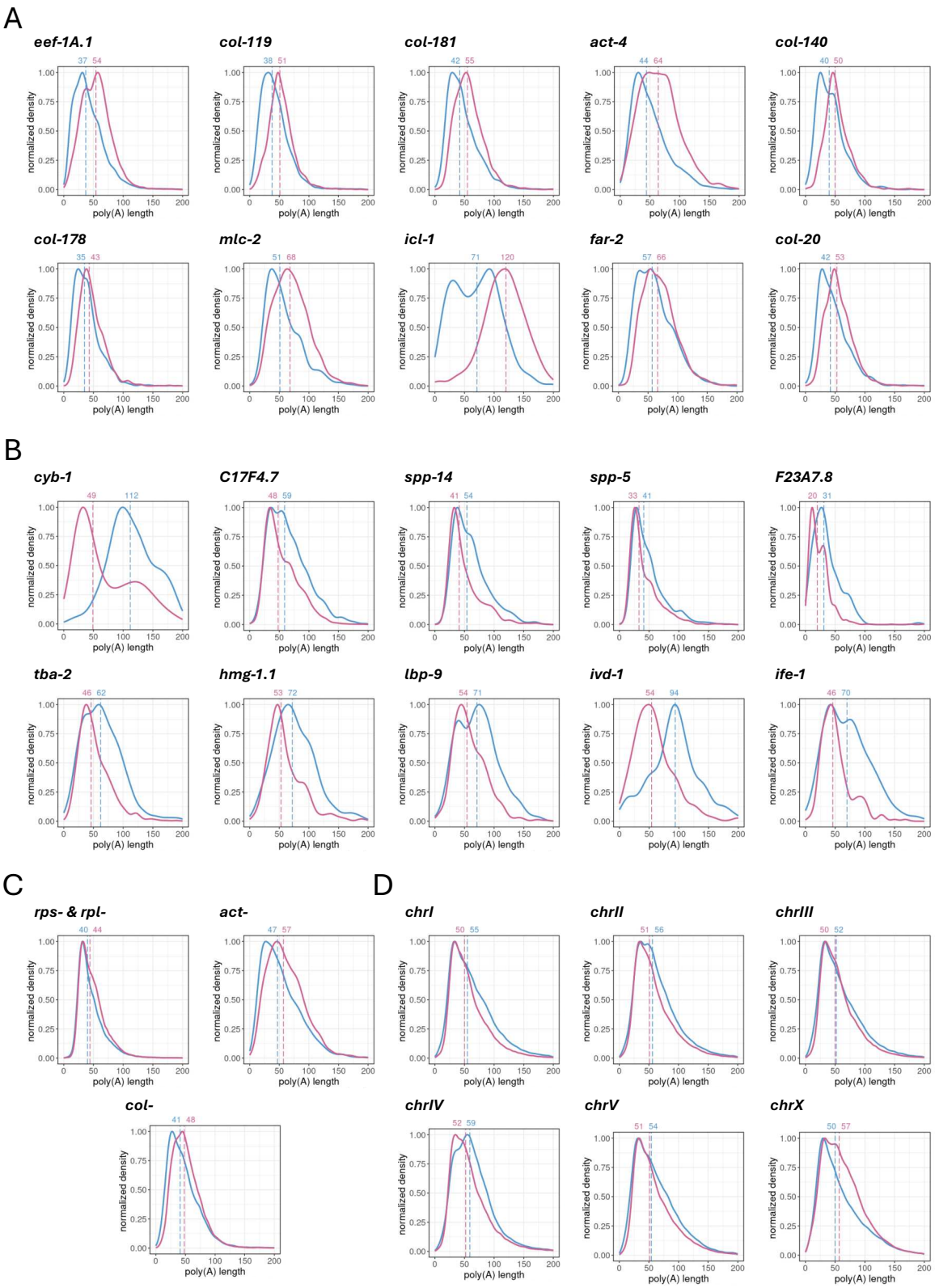
52. Schindelman,G., Whittaker,A.J., Thum,J.Y., Gharib,S. and Sternberg,P.W. (2006) Initiation of male sperm-transfer behavior in Caenorhabditis elegans requires input from the ventral nerve cord. *BMC Biol*, **4**.
53. Nikonorova,I.A., Wang,J., Cope,A.L., Tilton,P.E., Power,K.M., Walsh,J.D., Akella,J.S., Krauchunas,A.R., Shah,P. and Barr,M.M. (2022) Isolation, profiling, and tracking of extracellular vesicle cargo in Caenorhabditis elegans. *Current Biology*, **32**, 1924–1936.
54. Gewartowska,O., Aranaz-Novaliches,G., Krawczyk,P.S., Mroczek,S., Kusio-Kobiałka,M., Tarkowski,B., Spoutil,F., Benada,O., Kofroňová,O., Szwedziak,P., *et al.* (2021) Cytoplasmic polyadenylation by TENT5A is required for proper bone formation. *Cell Rep*, **35**.
55. Mroczek,S., Chlebowska,J., Kuliński,T.M., Gewartowska,O., Gruchota,J., Cysewski,D., Liudkovska,V., Borsuk,E., Nowis,D. and Dziembowski,A. (2017) The non-canonical poly(A) polymerase FAM46C acts as an onco-suppressor in multiple myeloma. *Nat Commun*, **8**.
56. Smith,J.R. and Stanfield,G.M. (2011) TRY-5 is a sperm-activating protease in Caenorhabditis elegans seminal fluid. *PLoS Genet*, **7**.
57. Barr,M.M., García,L.R. and Portman,D.S. (2018) Sexual dimorphism and sex differences in Caenorhabditis elegans neuronal development and behavior. *Genetics*, **208**, 909–935.
58. Pees,B., Yang,W., Zárate-Potes,A., Schulenburg,H. and Dierking,K. (2016) High Innate Immune Specificity through Diversified C-Type Lectin-Like Domain Proteins in Invertebrates. *J Innate Immun*, **8**, 129–142.
59. Pees,B., Yang,W., Kloock,A., Petersen,C., Peters,L., Fan,L., Friedrichsen,M., Butze,S., Zárate-Potes,A., Schulenburg,H., *et al.* (2021) Effector and regulator: Diverse functions of C. elegans C-type lectin-like domain proteins. *PLoS Pathog*, **17**.
60. Piloto,J.H., Rodriguez,M. and Choe,K.P. (2022) Sexual dimorphism in Caenorhabditis elegans stress resistance. *PLoS One*, **17**.
61. Avila,F.W., Sirot,L.K., Laflamme,B.A., Rubinstein,C.D. and Wolfner,M.F. (2011) Insect seminal fluid proteins: Identification and function. *Annu Rev Entomol*, **56**, 21–40.
62. Poiani,A. (2006) Complexity of seminal fluid: A review. *Behav Ecol Sociobiol*, **60**, 289–310.
63. Matyash,V., Geier,C., Henske,A., Mukherjee,S., Hirsh,D., Thiele,C., Grant,B., Maxfield,F.R. and Kurzchalia,T. V (2001) Distribution and Transport of Cholesterol in Caenorhabditis elegans. *Mol Biol Cell*, **12**, 1725–1736.
64. Yang,F., Wang,W., Cetinbas,M., Sadreyev,R.I. and Blower,M.D. (2020) Genome-wide analysis identifies cis-acting elements regulating mRNA polyadenylation and translation during vertebrate oocyte maturation. *RNA*, **26**, 324–344.
65. Xiong,Z., Xu,K., Lin,Z., Kong,F., Wang,Q., Quan,Y., Sha,Q. qian, Li,F., Zou,Z., Liu,L., *et al.* (2022) Ultrasensitive Ribo-seq reveals translational landscapes during mammalian oocyte-to-embryo transition and pre-implantation development. *Nat Cell Biol*, **24**, 968–980.
66. Albritton,S.E., Kranz,A.L., Rao,P., Kramer,M., Dieterich,C. and Ercan,S. (2014) Sex-biased gene expression and evolution of the X chromosome in nematodes. *Genetics*, **197**, 865–883.
67. Wang,G. and Reinke,V. (2008) A C. elegans Piwi, PRG-1, Regulates 21U-RNAs during Spermatogenesis. *Current Biology*, **18**, 861–867.
68. Nakanishi,T., Kumagai,S., Kimura,M., Watanabe,H., Sakurai,T., Kimura,M., Kashiwabara,S. ichi and Baba,T. (2007) Disruption of mouse poly(A) polymerase mGLD-2 does not alter polyadenylation status in oocytes and somatic cells. *Biochem Biophys Res Commun*, **364**, 14–19.

69. Cong,J., Yang,Y., Wang,X., Shen,Y., Qi,H.T., Liu,C., Tang,S., Wu,S., Tian,S., Zhou,Y., *et al.* (2022) Deficiency of X-linked TENT5D causes male infertility by disrupting the mRNA stability during spermatogenesis. *Cell Discov*, **8**.
70. Zheng,C., Ouyang,Y.C., Jiang,B., Lin,X., Chen,J., Dong,M.Z., Zhuang,X., Yuan,S., Sun,Q.Y. and Han,C. (2019) Non-canonical RNA polyadenylation polymerase FAM46C is essential for fastening sperm head and flagellum in mice. *Biol Reprod*, **100**, 1673–1685.
71. Suh,J. and Hutter,H. (2012) A survey of putative secreted and transmembrane proteins encoded in the C. elegans genome. *BMC Genomics*, **13**.

Supplementary Materials



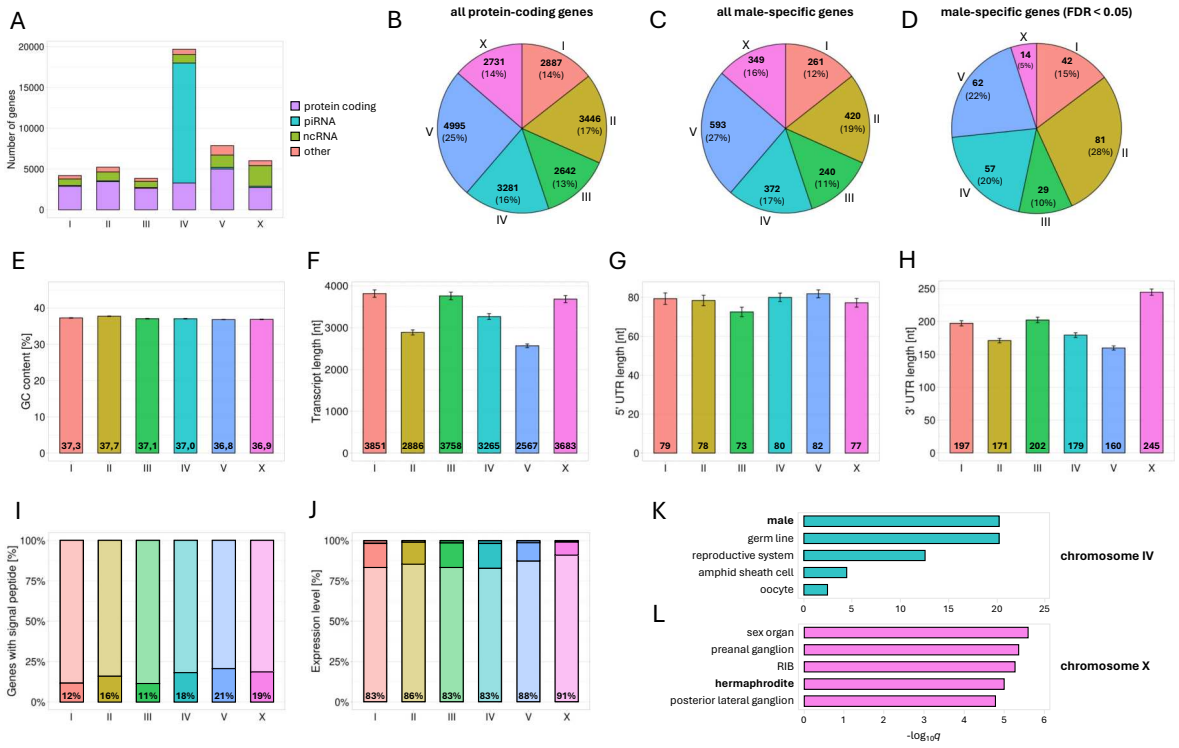
Supplementary Figure S1. Functional enrichment for groups of male- and hermaphrodite-specific genes.
(A) Top GO terms for male-specific genes with FDR < 0.05 ordered by adjusted *p*-value (WormBase Enrichment Suite).
(B) Top GO terms for hermaphrodite-specific genes with FDR < 0.05 ordered by adjusted *p*-value (WormBase Enrichment Suite).



Supplementary Figure S2. Differences in poly(A) distributions between wild-type males and hermaphrodites.

(A) Top ten protein-coding transcripts with the most significant poly(A) elongation in hermaphrodites according to adjusted *p*-value (Supplementary Table S2). Poly(A) distributions are shown as blue and pink lines for males and hermaphrodites, respectively. Vertical dashed lines represent the median poly(A) tail length for each condition (in nucleotides). The plots were normalized to 1.

(B) Top ten protein-coding transcripts with the most significant poly(A) elongation in males according to adjusted *p*-value (Supplementary Table S2). Poly(A) distributions are shown as blue and pink lines for males and hermaphrodites, respectively. Vertical dashed lines represent the median poly(A) tail length for each condition (in nucleotides). The plots were normalized to 1.
(C) Density plots showing differences in the global poly(A) tail distribution for ribosomal genes (*rpl* and *rps*), actin genes (*act*), and collagen genes (*col*) between males (blue line) and hermaphrodites (pink line). Vertical dashed lines represent the median poly(A) tail length for each condition (in nucleotides). The plots were generated for all transcripts identified from each group and normalized to 1.
(D) Density plots showing differences in the global poly(A) tail distribution for genes localized on all six chromosomes between males (blue line) and hermaphrodites (pink line). Vertical dashed lines represent the median poly(A) tail length for each condition (in nucleotides). Plots were generated for all transcripts assigned to a particular chromosome and normalized to 1.



Supplementary Figure S3. Differences in chromosome compositions.

(A) Composition of each chromosome regarding RNA type. Four different types of RNA were highlighted: protein-coding mRNAs (purple), piRNA (cyan), ncRNA (green), and all other types (orange). The dataset for analysis was downloaded from ParaSite BioMart (WS291).

(B-D) Pie charts representing distributions of all protein-coding genes **(B)**, all male-specific genes **(C),** and significantly upregulated male-specific genes **(D)** between different chromosomes. The dataset for analyses was downloaded from ParaSite BioMart (WS291) and only protein-coding genes were included. Numbers on charts represent the number and the percentage of genes on each chromosome.

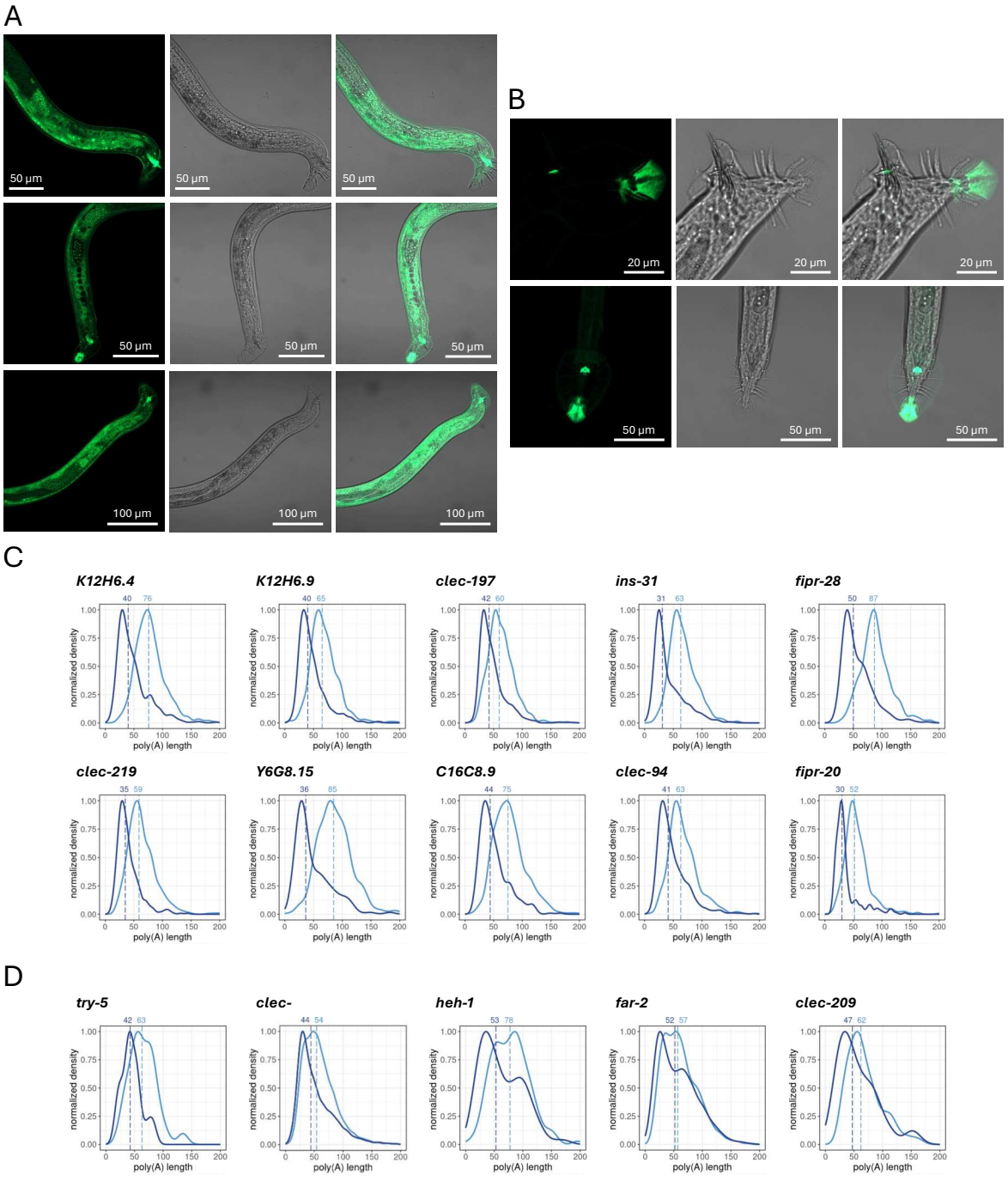
(E) Differences in GC content between genes localized on different chromosomes. The dataset for analysis was downloaded from ParaSite BioMart (WS291) and only protein-coding genes were included. Barplots represent mean values with SD.

(F-H) Differences in transcript length (unspliced transcript + UTRs) **(F)**, 5′ UTR length **(G),** and 3′ UTR length **(H)** between genes localized on different chromosomes. The dataset for analyses was downloaded from ParaSite BioMart (WS291) and only protein-coding genes were included. The average length of 5′ UTR and 3′ UTR is calculated considering the longer UTRs for each gene (the analyses were done for a narrowed-down subset of genes available in the dataset with necessary annotation). Barplots on each plot represent mean values with SD.

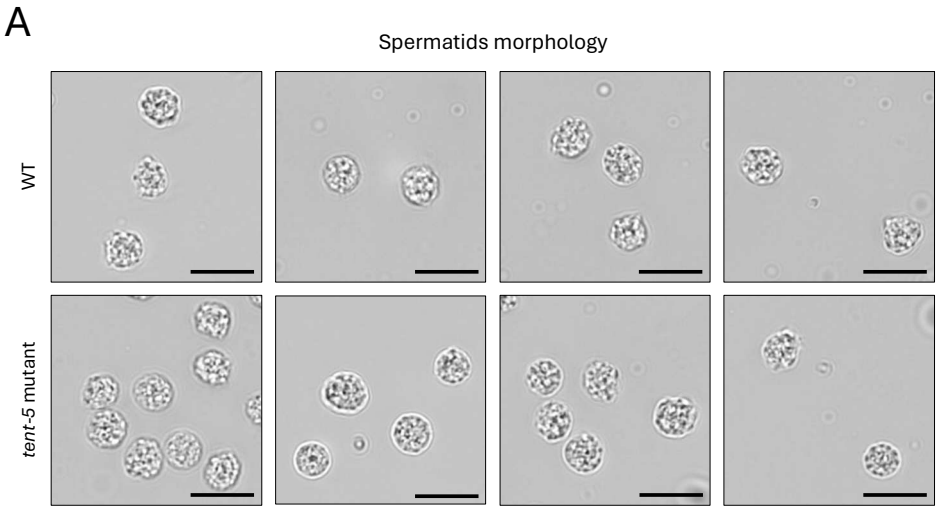
(I) Fractions of genes from each chromosome possessing signal peptide sequence based on studies by Suh J., Hutter H. (2012) (71). All genes detected in our DRS data were included. Numbers on the bars represent the percentage of genes with signal peptide sequence.

(J) Fractions of genes from each chromosome characterized by low, medium, or high expression levels. Expression levels are defined by baseMean values presented in Supplementary Table S1. Genes with baseMean < 20 are defined as lowly expressed, with 20 < baseMean < 500 as medium expressed, and with baseMean > 500 as highly expressed. All genes detected in our DRS data were included. Numbers on the bars represent the percentage of lowly expressed genes.

(K) and **(L)** Top five Tissue Enrichment terms for genes encoded on chromosome IV **(K)** or chromosome X **(L)** ordered by adjusted *p*-value (WormBase Enrichment Suite).



Supplementary Figure S4. TENT-5 regulates the expression of male-specific genes.
(A) Additional fluorescence microscopic images of TENT-5-GFP expression in *C. elegans* males (*tent-5(rtt6[tent-5::gfp::3xflag]* I). All images show the tail-end of a male body. Images were taken with a 40x lens and merged with bright-field images.
(B) Autofluorescence of the tail and copulatory spicules in wild-type (upper panel) and *tent-5:gfp* (lower panel) *C. elegans* males. Images were taken with a 40x lens and merged with bright-field images.
(C) Top ten protein-coding transcripts with the most significant poly(A) shortening in *tent-5(tm3504)* mutant males according to adjusted *p*-value (Supplementary Table S4). Poly(A) distributions are shown as light and dark blue for wild-type males and *tent-5(tm3504)* mutants, respectively. Vertical dashed lines represent the median poly(A) tail length for each condition (in nucleotides). The plots are normalized to 1.
(D) Density plots showing differences in the poly(A) tail distribution for genes or groups of genes mentioned individually in the chapter *"TENT-5 regulates the expression of male-specific genes"*. Vertical dashed lines represent the median poly(A) tail length for each condition (in nucleotides). The plots are normalized to 1.



Supplementary Figure S5. Comparison of spermatids morphology between wild-type and *tent-5* mutant males.
(A) Additional microscopic images of spermatids morphology for wild-type and *tent-5* mutant males. Images are taken with 100x magnification. Scale bars => 10 microns.

Supplementary Movie S1. Mating behavior experiment performed by crossing wild-type (N2) or *tent-5(tm3504)* mutant males with *unc-45* hermaphrodites. The movie is shown at 20x speed.

Appendix 3

Manuscript 3

Mackiewicz Z., Liudkovska V., Dziembowski A. Knockout of all nematode-specific NSPC genes expressed exclusively in the excretory gland cell results in transcriptomic signatures indicating an affected insulin signaling. *bioRxiv* (2024). DOI: 10.1101/2024.12.06.627198

Warsaw, 7.01.2025

Zuzanna Mackiewicz
Laboratory of RNA Biology
International Institute of Molecular and Cell Biology
4 Ks. Trojdena
02-109 Warsaw

PHD CANDIDATE'S
CONTRIBUTION STATEMENT

Article's title: Knockout of all nematode-specific NSPC genes expressed exclusively in the excretory gland cell results in transcriptomic signatures indicating an affected insulin signaling.

Authors: Zuzanna Mackiewicz, Vladyslava Liudkovska, Andrzej Dziembowski

Journal: bioRxiv

Date of publishing: 7th December 2024

DOI: 10.1101/2024.12.06.627198

I hereby declare that my contribution to the article titled *"Knockout of all nematode-specific NSPC genes expressed exclusively in the excretory gland cell results in transcriptomic signatures indicating an affected insulin signaling"* was as following:

- designing experiments
- preparing most of the *C. elegans* transgenic strains (without strains ADZ31-ADZ45)
- conducting most of the experiments (without studying localization of NSPC)
- analyzing all Illumina and Nanopore RNA sequencing data
- writing the initial draft of the manuscript and creating all figures

.....
PhD candidate signature

.....
Corresponding author signature

Knockout of all nematode-specific NSPC genes expressed exclusively in the excretory gland cell results in transcriptomic signatures indicating an affected insulin signaling

Zuzanna Mackiewicz[1], Vladyslava Liudkovska[2], Andrzej Dziembowski[1]

[1]Laboratory of RNA Biology, International Institute of Molecular and Cell Biology, Trojdena 4, 02-109 Warsaw, Poland
[2]Laboratory of Stem Cell RNA Metabolism, IMol Polish Academy of Sciences, Flisa 6, 02-247 Warsaw, Poland

Abstract

The nematode *Caenorhabditis elegans* is one of the best-studied model organisms in molecular biology; however, many aspects of its physiology and the functions of many genes remain poorly understood. In this study, we investigated the role of nematode-specific NSPC proteins, whose mRNAs were recently identified as primary targets of the poly(A) polymerase TENT-5. Surprisingly, we found that NSPCs are exclusively expressed in the excretory gland cell, a cell with still unclear functionality. Using an optogenetic approach, we precisely ablated the excretory gland cell and observed that nematodes exhibited no transcriptomic or physiological changes in its absence. Additionally, we generated and thoroughly studied a strain with a deletion of all 18 *nspc* genes, which revealed that, despite previous indications, NSPCs do not influence the worm's defense response. Instead, the transcriptomic analysis showed that the absence of NSPCs strongly impacts DAF-2/DAF-16 insulin signaling, suggesting that NSPCs may function as neuropeptides influencing key *C. elegans* signaling pathways. Although further studies are required to elucidate the physiological effects of this regulation, our findings provide new insights into this unexplored part of nematode physiology.

Introduction

The small nematode *Caenorhabditis elegans* was introduced to the scientific community in the 1970s and rapidly became one of the most important model organisms used in genetics and molecular biology [1]. Consequently, it is one of the best-studied multicellular organisms and was the first to have its entire genome sequenced [2], as well as its neuronal connectome [3,4] and cell lineage completely mapped [5–7]. Despite the extensive exploration of *C. elegans* over the years, some aspects of its anatomy and physiology remain poorly understood. Among these is the excretory gland cell, one of the 959 cells in the adult hermaphrodite worm, whose function remains a mystery, with only a few studies briefly mentioning its existence [8–11].

The excretory gland cell, along with the pore cell, duct cell, and excretory (canal) cell, constitutes the nematode's secretory system, positioned on the ventral side of the pharynx, adjacent to the anterior section of the intestine [12]. This four-cell organ develops from the early embryonic stem cell (AB) and becomes functional at or soon after hatching. The excretory gland cell is binucleate and A-shaped and forms by the fusion of two identical cells at the 1.5-fold embryo stage. In addition to its tight connection with other excretory cells, it is also joined to the nerve ring, suggesting possible neural regulation of its activity. Structurally, the excretory gland cell features an extensive endoplasmic reticulum (ER), numerous mitochondria, and a dense ribosomal population [12].

In the 1980s, laser ablation was used to investigate the function and interactions of individual cells within the excretory system, including the excretory gland cell [10,11]. By precisely targeting cell nuclei with a laser beam at different developmental stages (L1, L2, young adult,

or dauer), researchers observed that, while all other secretory cells are involved in osmoregulation, the excretory gland cell is not required for this process and also does not regulate *C. elegans* fertility or viability. Despite these findings made decades ago, no further studies have explored the function of the excretory gland cell.

In our previous research, we described the role of the non-canonical poly(A) polymerase TENT-5 in the innate immune response of *C. elegans* [13]. Using Nanopore Direct RNA sequencing, we demonstrated that TENT-5 polyadenylates and stabilizes a pool of mRNAs, many of which encode short, secreted proteins with well-defined or putative roles in the defense response. Intriguingly, our transcriptomic analysis revealed that the most prominent targets of TENT-5 were members of the NSPC family (Nematode Specific Peptide Family, group C), which encodes 18 small, highly similar proteins. Although some evidence suggests that NSPCs may be connected to immune or stress responses in worms [14–16], their exact function remains unknown. However, the observed relationship between TENT-5 and *nspc* mRNAs indicates that, like other TENT-5 targets, NSPC proteins may indeed be involved in immune regulation.

In this work, we show that all members of the NSPC family exhibit strong and highly specific expression exclusively in the excretory gland cell. This important finding prompted us to investigate the functions of both the excretory gland cell and NSPC proteins in *C. elegans* physiology. To uncover the role of the excretory gland cell, we employed an optogenetic approach to irreversibly damage the cell and disrupt its function. Since NSPCs are among the most abundant proteins in this cell, we also used a complementary approach by generating a CRISPR/Cas9 knockout strain lacking all *nspc* genes. Our extensive studies reveal that neither the ablation of the excretory gland cell nor the deletion of all NSPCs led to profound transcriptome changes or noticeable physiological phenotypes under tested laboratory conditions. However, here we propose that NSPCs might not possess antimicrobial properties, as predicted before, but might serve as neuropeptides that help regulate important signaling pathways such as the insulin DAF-2/DAF-16 pathway. Although our findings could not fully explain this mechanism, they provide valuable insights into this unexplored aspect of nematode biology, and we believe they will accelerate future research in the field.

Results and discussion

NSPCs are exclusively expressed in the excretory gland cell

The NSPC family comprises 18 genes clustered on chromosome X (Figure 1A), encoding small, secreted proteins with highly similar sequences and structures (Figure 1B and C) that are not related to any other known proteins outside the *Caenorhabditis* genus. Surprisingly, only one study has explored the possible functions of NSPC members, proposing that they may have antimicrobial activity based on their sequence and structure [14]. This prediction is consistent with high-throughput transcriptomic analyses showing dysregulation of *nspc* genes mostly upon various pathogen infections (Supplementary Table S1). For example, bacterial infection with *Enterococcus faecalis* (OG1RF), *Photorhabdus luminescens* (HB), or *Xenorhabdus nematophila* leads to the downregulation of multiple *nspc* genes [15,16]. In contrast, bacterial infection with *Pseudomonas aeruginosa* or *Serratia marcescens*, as well as fungal infection with *Drechmeria coniospora* or *Harposporium sp.,* results in increased expression levels of *nspc* genes [15–17]. Moreover, our previous studies on poly(A) polymerase TENT-5 function in *C. elegans* physiology revealed its pronounced activity toward NSPC-coding mRNAs, along with multiple known immune effectors [13]. However, despite many indications, NSPC activity has never been demonstrated either *in vitro* or *in vivo,* and their function remains unknown.

Therefore, we aimed to study the exact role of NSPCs and investigate the importance of their relationship with TENT-5. We began by studying the localization of NSPCs within the *C. elegans* body. We generated extrachromosomal reporters for three *nspc* genes fused with mCherry and examined their expression using confocal microscopy. Surprisingly, we observed that NSPCs are expressed exclusively in a single cell – the excretory gland cell (Figure 1D, Supplementary Figure S1), which agrees with high-throughput cell-specific expression profiling [18]. Moreover, the expression levels of NSPCs are extremely high, as they are detectable in RNA sequencing experiments, even from whole-animal samples. Although not all NSPCs exhibit the same expression levels throughout *C. elegans's* lifespan, they appear to be highly expressed throughout all developmental stages [18,19]. The excretory gland cell is a part of the nematode's excretory system, where four distinct cells collaborate to perform a specific task throughout the nematode's lifespan (Figure 1E). Notably, while the functions of the other excretory cells have been characterized, the role of the excretory gland cell, similarly to NSPC proteins, remains a mystery.

We also checked whether NSPCs and TENT-5 colocalize in the excretory gland cell using worms expressing both *tent-5*::GFP and *nspc-14p*::mCherry as extrachromosomal arrays. Indeed, we found that TENT-5 is also expressed in the excretory gland cell (Figure 1F), allowing for the direct regulation of *nspc* mRNAs by TENT-5. Following these discoveries, here we aimed to explain the function of the excretory gland cell, understand why it requires such a high and specific expression of NSPCs, and explore the role of TENT-5 in regulating this previously undescribed aspect of *C. elegans* physiology.

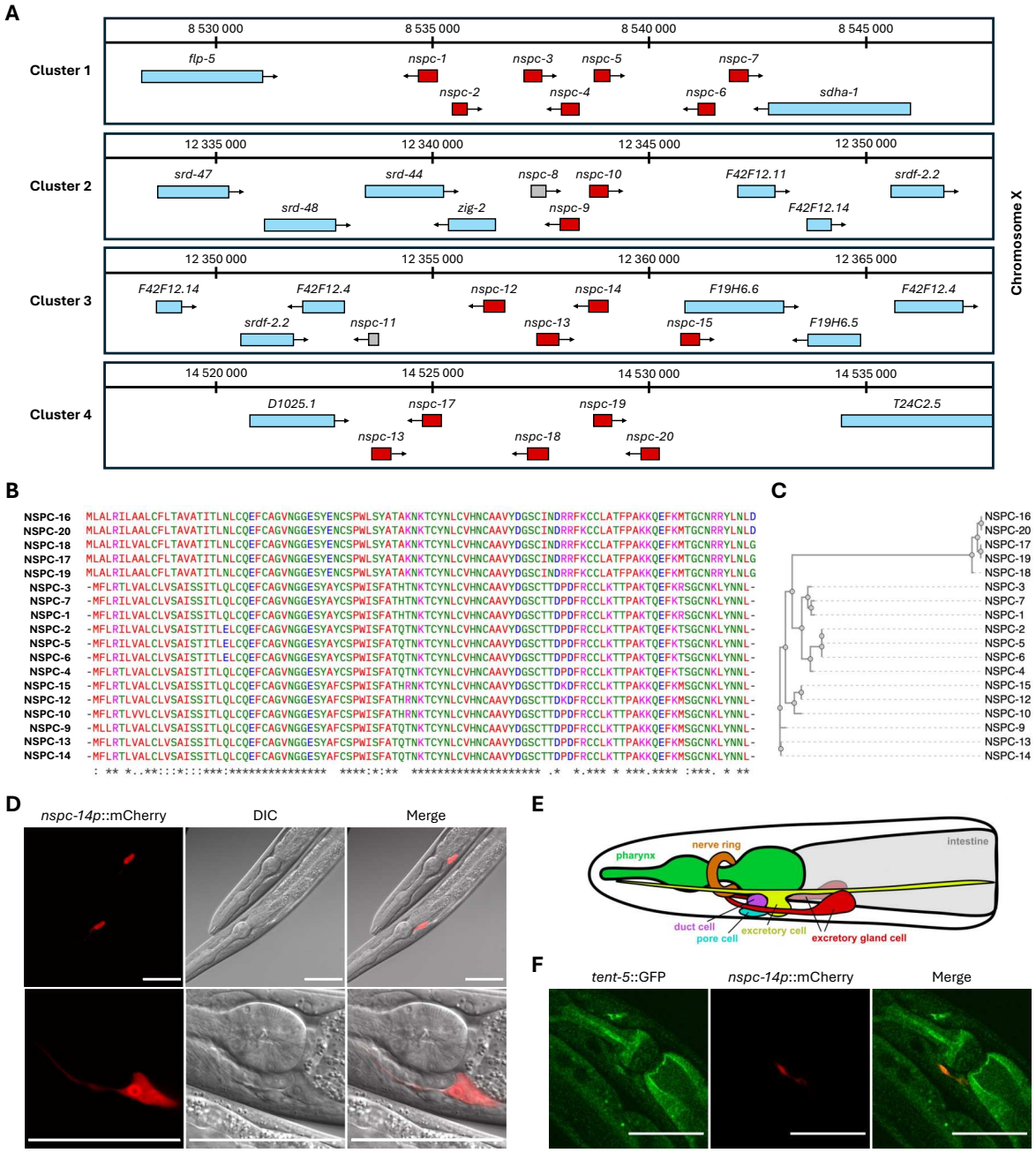


Figure 1. NSPCs are exclusively expressed in the excretory gland cell.

(A) Genomic locations of *nspc* genes on chromosome X. *Nspc* genes are marked in red, neighboring genes in blue, and pseudogenes *nspc-8* and *nspc-11* in gray. Boxes represent gene structures, including exons, introns, and UTRs, while arrows indicate the DNA strand. The genomic map was prepared based on the WormBase genome browser [19].

(B) Multiple sequence alignment of NSPC proteins. Asterisks (*) represent residues conserved across all NSPCs; colons (:) indicate moderately conserved residues; and dots (.) indicate non-conserved residues with similar chemical properties. Alignment was performed using the MUSCLE tool from EMBL-EBI [20].

(C) Phylogenetic tree of NSPC proteins. Analysis was performed using the MUSCLE tool from EMBL-EBI [20].

(D) Fluorescence and differential interference contrast (DIC) microscopy images of NSPC-mCherry expression in *C. elegans* excretory gland cell. Scale bars => 50 µm.

(E) Schematic representation of the excretory system in *C. elegans*. Prepared based on the scheme from Sundaram M.V. and Buechner M. (2016) [8]. The excretory system is located near the pharynx (green) and nerve ring (orange) and consists of the excretory cell (yellow), duct cell (purple), pore cell (blue), and excretory gland cell (red).

(F) Fluorescence microscopy images showing colocalization of TENT-5-GFP and NSPC-mCherry in the excretory gland cell. Scale bars => 50 µm.

Optogenetic ablation can be successfully implemented to destroy the excretory gland cell

To study the role of the excretory gland cell, we took advantage of the restricted NSPC expression, enabling the optogenetic cell ablation approach, which has been previously used for the precise ablation of neurons, muscles, and epidermal cells in *C. elegans* [21,22]. This method relies on the expression of the genetically encoded photosensitizer miniSOG (mini Singlet Oxygen Generator) specifically in the cell or tissue of interest [23]. Upon blue light illumination (maximum excitation at 448 nm), miniSOG is targeted to the outer mitochondrial membrane, where it generates toxic levels of singlet oxygen ($^1O_2$) and other reactive oxygen species (ROS), leading to cell damage and death (Figure 2A).

Using CRISPR/Cas9, we generated two transgenic strains expressing miniSOG under the *nspc-10* promoter: one with active miniSOG fused to the mitochondrial membrane protein TOMM-20 and another with miniSOG carrying a point mutation that renders the photosensitizer inactive, serving as a negative control (Figure 2B). To visualize the excretory gland cell and monitor its damage post-ablation, both miniSOG strains were crossed with a strain expressing the red-fluorescent protein mCherry under the *nspc-14* promoter on the extrachromosomal array (Figure 2C). Since miniSOG itself is a green fluorescent flavoprotein [23], we observed its localization to the excretory gland cell (Figure 2D) and also confirmed protein production by western blot analysis (Figure 2E), thus validating the correct strain preparation for optogenetic ablation.

To optimize the ablation protocol for the excretory gland cell, we tested various blue light sources, illumination durations, and developmental stages of the worms (Supplementary Figure S2A). Unlike other optogenetic studies, where photosensitizers were typically activated using microscope fluorescence lamps, we developed a high-throughput method using an LED advertising board emitting 460 nm blue light (Supplementary Figure S2B). This setup enabled the simultaneous processing of a large worm population for subsequent high-throughput analyses, such as RNA sequencing. Ultimately, we achieved optimal excretory gland cell ablation by illuminating L3 stage larvae with continuous blue light in five 10-minute intervals, allowing time to prevent worm overheating. This approach successfully and selectively destroyed the excretory gland cell in worms expressing active miniSOG, while worms expressing the mutated control showed no signs of cell damage. The ablation was confirmed by the loss of mCherry fluorescence approximately 40 hours post-illumination (Figure 2F, Supplementary Figure S2C) and by RT-qPCR, which demonstrated a nearly 90% reduction in NSPCs expression (Figure 2G). These observations indicate that the cell was either efficiently destroyed or significantly damaged, impairing its function.

In conclusion, we successfully employed a miniSOG-mediated optogenetic approach to ablate the excretory gland cell in *C. elegans*. This method provides a model for studying the function of the excretory gland cell and offers extensive possibilities for further investigation, surpassing the limitations of traditional laser ablation techniques.

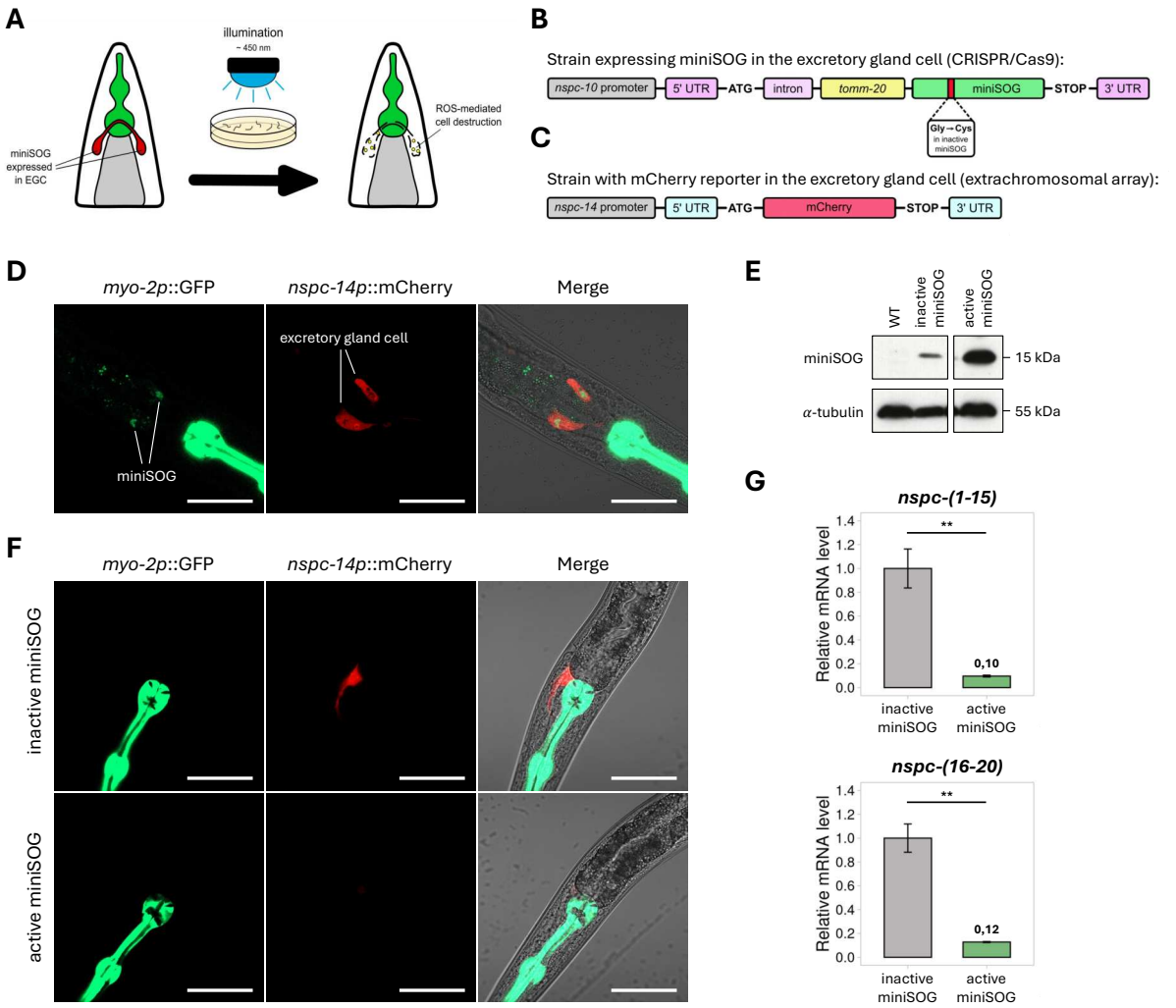


Figure 2. Optogenetic ablation can be implemented to destroy the excretory gland cell.
(A) Schematic representation of the optogenetic ablation of the excretory gland cell (EGC). Worms were illuminated on plates using the LED board emitting 460 nm blue light in five 10-minute intervals. Upon illumination, miniSOG generates reactive oxygen species (ROS), leading to cell-specific destruction.
(B) Scheme illustrating the design of *C. elegans* transgenic strains used for the optogenetic ablation of the excretory gland cell. miniSOG, together with mitochondria-directing TOMM-20, was expressed under the *nspc-10* promoter, with the *nspc-10* intron and 5/3′ UTRs. A single nucleotide mutation in the miniSOG sequence results in Gly->Cys amino acid substitution, blocking miniSOG activity.
(C) Scheme showing the design of *C. elegans* transgenic strain used for visualization of the excretory gland cell. Red fluorescence protein mCherry was expressed under the *nspc-14* promoter in the extrachromosomal array.
(D) Fluorescence microscopy images of the excretory gland cell expressing miniSOG. Green fluorescence of miniSOG is visible in the cytoplasm of the excretory gland cell, likely overlapping with mitochondria localization. Scale bars => 50 µm.
(E) Western blot showing successful generation of two miniSOG-expressing strains. α-tubulin was used as a loading control. The lower amount of miniSOG in the inactive miniSOG strain may result from suboptimal antibody binding to the mutated protein.
(F) Fluorescence microscopy images of the excretory gland cell after optogenetic ablation. Top images show a worm with inactive miniSOG (control), and the bottom – a worm with active miniSOG. The absence of red NSPC-mCherry fluorescence in the worm with active miniSOG indicates successful ablation. Scale bars => 50 µm.
(G) RT-qPCR showing a decrease in all *nspc* gene expression levels after ablation of the excretory gland cell. Relative *nspc* mRNA levels were normalized to *act-1*. Bar plots represent mean values with SD. ** indicates *p*-value < 0.01 (two-tailed *t*-test).

Ablation of the excretory gland cell leads to no transcriptome changes

We performed the optogenetic ablation of the excretory gland cell, as described above, and monitored worms for potential abnormal phenotypes. Interestingly, worms remained viable after ablation and displayed no visible phenotypes, which is consistent with results from previous laser ablation experiments [10,11]. A few possibilities might explain this surprising outcome. One possibility is that the excretory gland cell plays a more significant role during earlier developmental stages, which are challenging to study using the optogenetic setup due to the harmful effects of blue light exposure on younger worms (Supplementary Figure S2A). Another possibility is that defects in the excretory gland cell only become evident under some specific stress conditions. Furthermore, the loss of the excretory gland cell might lead to subtle effects that are not apparent in anatomical or physiological phenotypes but may instead be revealed through altered patterns of gene expression. To explore this latter possibility, we performed Illumina RNA sequencing (RNA-seq) on adult worms with both active and inactive (control) miniSOG two days after illumination. The RNA-seq results showed significant changes in gene expression, with 1210 genes downregulated and 1293 upregulated following cell ablation (Figure 3A, Supplementary Table S2). Notably, the downregulation of nearly all NSPC genes confirmed the successful removal and dysfunction of the excretory gland cell.

Interestingly, further analysis revealed that mainly sperm-enriched genes were downregulated, while oocyte-enriched genes were upregulated (Figure 3B and C). Although this pattern might suggest a role for the excretory gland cell in reproduction, we did not detect any impact of this gene dysregulation on worm brood sizes following ablation (Figure 3D). However, the observed transcriptome changes might also reflect a characteristic gene expression profile associated with age differences between worms, since sperm gene expression peaks in the L4 stage, while oocyte genes are expressed more in young adults [24,25]. To test this possibility, we used the R package RAPToR [26] to estimate the age of the worms in each RNA-seq sample (Figure 3E). Indeed, although all samples were processed simultaneously and appeared to be of similar age based on microscopic observations, we found an age difference of a few hours between worms with active and inactive miniSOG, which likely contributed to the overall gene expression pattern.

To minimize confusion caused by age-related differences, we repeated the RNA sequencing two more times and included wild-type worms post-illumination as an additional control. Unexpectedly, the transcriptome changes remained variable but consistently involved sperm- and oocyte-enriched genes (Supplementary Figure S3), indicating developmental stage variations between samples (Figure 3E). These differences were also strongly pronounced between two control samples in one of the experiments (Figure 3E and F, Supplementary Table S2), supporting the idea that age discrepancies influenced the results. The exact reason for such age variance between conditions in this particular experiment remains unclear, but we suspect that the illumination process may have impacted the worms' developmental timing [27], even though no changes in behavior or viability were observed afterward (Supplementary Figure S3).

Finally, to reduce the impact of the observed age differences, we combined the results from all three independent experiments. Surprisingly, the overall transcriptome changes following excretory gland cell ablation were minimal, with only 17 genes significantly downregulated and one upregulated (Figure 3G, Supplementary Table S2). Among the downregulated genes, ten belonged to the NSPC family, confirming the successful ablation of the excretory gland cell but also suggesting that the cell's function may primarily involve NSPC production.

The minimal impact on both nematode physiology and transcriptome following the ablation of the entire cell seems quite unusual. As mentioned above, the cell's role may be manifested under the environmental pressures found in the wild, but it is difficult to detect in standard laboratory conditions. More extensive studies will be required to further explore this hypothesis.

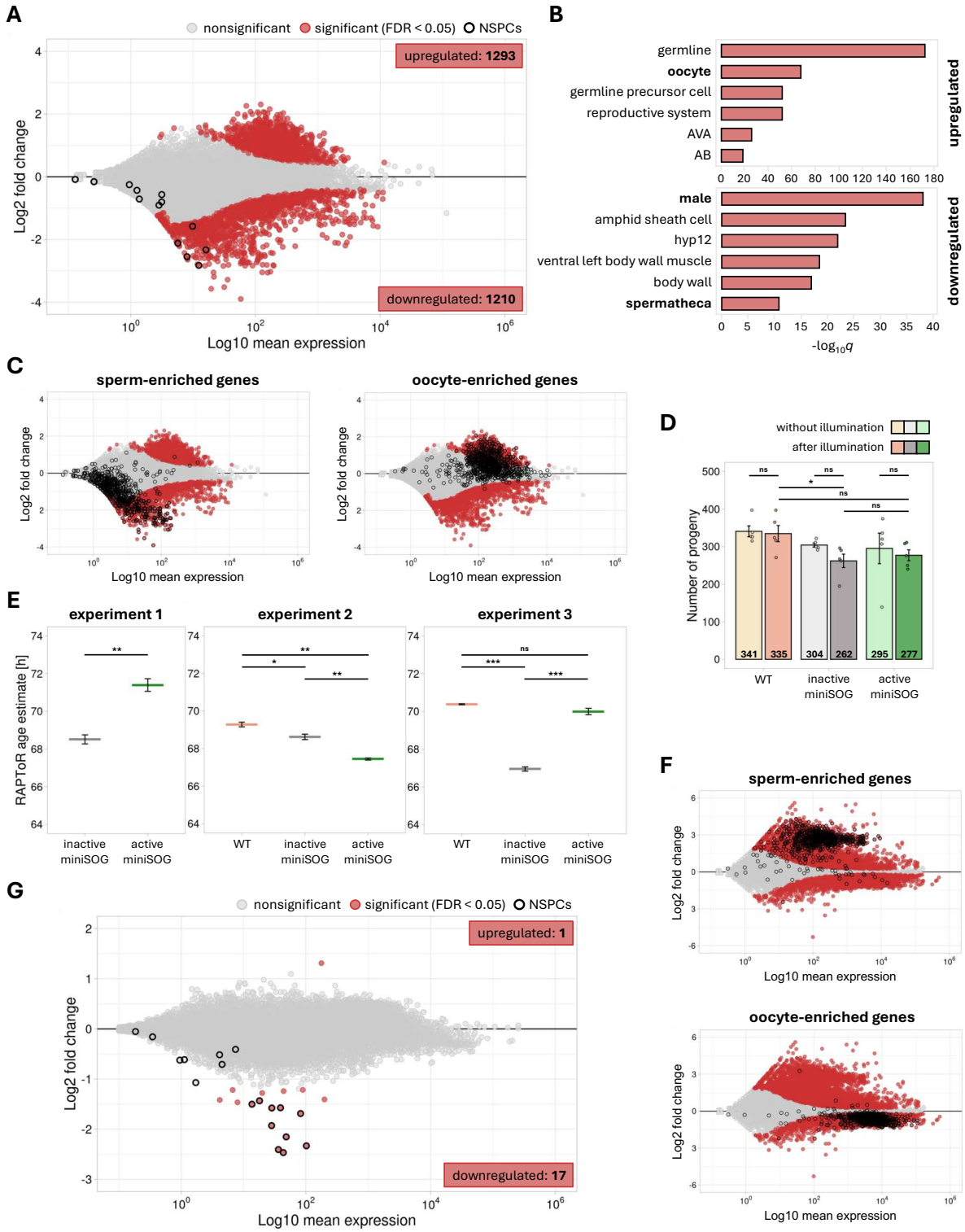


Figure 3. Ablation of the excretory gland cell leads to minimal transcriptome changes.
(A) MA plot illustrating differential gene expression between inactive and active miniSOG worms after the successful ablation of the excretory gland cell. Significantly changed genes (FDR < 0.05) are marked with red dots. The black borderline highlights *nspc* genes.

(B) Top GO terms for the upregulated (top) and downregulated (bottom) genes following the ablation of the excretory gland cell ordered by adjusted *p*-value. Terms related to the dysregulation of oocyte- and sperm-enriched genes are shown in bold.
(C) Duplicates of the MA plot in **(A)** showing the downregulation of sperm-enriched genes and the upregulation of oocyte-enriched genes. Marked genes are listed in Supplementary Table S2 and are derived from Reinke V. *et al.* (2004) [28].
(D) Differences in brood sizes between wild-type (orange), inactive (gray), and active miniSOG (green) worms without and after the illumination (lighter and darker shades, respectively). Bar plots represent mean values with SD (n=5). ns => not significant; * => *p*-value < 0.05 (two-tailed *t*-test).
(E) Age estimates for wild-type (orange), inactive (gray), and active miniSOG (green) worms after illumination, based on three independent RNA-seq replicates. Plots represent mean values with SD. ns => not significant; * => *p*-value < 0.05; ** => *p*-value < 0.01; *** => *p*-value < 0.001 (two-tailed *t*-test).
(F) MA plots showing differential gene expression between wild-type and inactive miniSOG worms in the third RNA-seq experiment (two control conditions). Significantly changed genes (FDR < 0.05) are marked with red dots. Black borderline is used to mark sperm- and oocyte-enriched genes, as in **(C)**.
(G) MA plots showing differential gene expression between inactive and active miniSOG worms after excretory gland cell ablation averaged across three independent RNA-seq experiments. Significantly changed genes (FDR < 0.05) are marked with red dots. Black borderline is used to mark *nspc* genes.

NSPC deletion results in no detectable physiological phenotypes

The averaged RNA sequencing results following the excretory gland cell ablation suggest that its primary role is likely connected with NSPC production. However, the functions of NSPC proteins remain largely unknown, complicating the interpretation of the optogenetic ablation results and making it challenging to assess the cell's overall significance. Therefore, we aimed to clarify the function of this poorly understood gene family, explain the strong TENT-5-mediated regulation of NSPCs, and provide new insights into the possible role of the excretory gland cell.

Given the similarity within NSPCs (Figure 1B and C), we initially assumed that RNA interference (RNAi), a method frequently and successfully employed in *C. elegans* research [29], would effectively silence all family members simultaneously. After extensive optimization, we achieved up to 90% silencing of NSPCs, followed by extensive transcriptomic and physiological experiments (Supplementary Figure S4A-D, Supplementary Table S3). However, due to the variability in silencing 18 genes simultaneously, RNAi results were often inconsistent, which complicated result interpretation. To ensure data reliability, we generated a CRISPR/Cas9 knockout strain lacking all NSPC genes. This process involved sequential deletions of four gene clusters: *nspc-1 – nspc-7*, *nspc-8 – nspc-10*, *nspc-11 – nspc-15,* and *nspc-16 – nspc-20* (Figure 4A).

First, we performed RNA-seq and identified significant changes in gene expression following NSPC deletion, with 1451 genes upregulated and 997 downregulated compared to wild-type worms (Figure 4B, Supplementary Table S3). Interestingly, many upregulated genes showed tissue enrichment in the worm's epidermis and intestine (Figure 4C), which are the first tissues to contact pathogens during infections and initiate immune response. Consistently with functional enrichment analysis, among the upregulated genes, we identified multiple regulators of innate immunity, such as infection response genes (4 *irg* family members) [30,31], C-type lectins (24 *clec* family members) [32], lysozymes (4 *lys* family members) [33], or aspartyl proteases (8 *asp* family members) [34] (Figure 4D). Similar GO terms were also observed for downregulated genes, suggesting a potential role for NSPCs in *C. elegans* defense responses. However, the observed expression changes, although significant, were generally small, with only a few genes exceeding a two-fold change. This modest gene expression change is consistent with the absence of visible phenotypes or significant lifespan alterations in the mutant worms (Figure 4E).

The potential involvement of NSPCs in defense processes, as indicated by our RNA-seq results, aligns with evolutionary predictions made for those genes [14] and the previously observed relationship between NSPCs and TENT-5 [13]. We previously demonstrated that TENT-5 stabilizes NSPC transcripts by elongating their poly(A) tails, although the role of this co-regulation in mediating the *C. elegans* response to pathogens remains unexplored. To better understand this mechanism, we first compared the pool of transcripts that were up- or downregulated following NSPC or TENT-5 deletion and found no clear overlaps between the groups, with some genes showing opposite regulation in the two mutants (Figure 4F, Supplementary Table S3). Next, using Nanopore Direct RNA Sequencing (DRS), we measured poly(A) tail lengths in *nspc* mutant worms and observed almost no changes in poly(A) metabolism compared to wild-type worms (Supplementary Figure S4E, Supplementary Table S4). This observation allowed us to exclude the possibility that TENT-5 is regulated or regulates other transcripts via NSPCs. Moreover, we generated a mutant strain lacking both TENT-5 and NSPCs and conducted additional RNA sequencing to analyze dependencies between these proteins (Figure 4G, Supplementary Figure S4F, Supplementary Table S3 and S4). Interestingly, NSPC deletion in the absence of TENT-5 led to smaller gene expression changes than when TENT-5 was present (Figure 4G, Supplementary Table S3), confirming that TENT-5 indeed regulates NSPCs. Among the few downregulated genes, we identified a unique operon containing the genes *islo-1, pmk-1, pmk-2,* and *pmk-3*, three of which are members of the PMK family, known for its crucial role in regulating *C. elegans* defense responses [35–37]. We showed that this result is due to the strong enrichment of PMK members in *tent-5* mutant worms, with levels returning to normal upon additional NSPC deletion (Supplementary Figure S4G), hinting at a possible complex interplay between TENT-5, NSPCs, and PMK pathway in regulating defense mechanisms.

Together, our results show that NSPCs are not essential for viability; however, they may play a role in facilitating defense responses. To further investigate this hypothesis, we conducted lifespan experiments with *P. aeruginosa* PAO1 infection, comparing four different *C. elegans* strains: wild-type N2, *tent-5* mutant, *nspc* mutant, and double *nspc*/*tent-5* mutant (Figure 4H). Although we confirmed previous findings of increased susceptibility of *tent-5* mutant worms to bacterial infection [13], we demonstrated that, despite previous indications, NSPCs are not required for worms' defense response to *P. aeruginosa*. Additionally, our results suggest that the excretory gland cell, which is the hub for the NSPCs production, is also unlikely to play a direct role in innate immunity. However, based on our results, we cannot rule out the possibility that the excretory gland cell and NSPCs may be involved in responses to other pathogens.

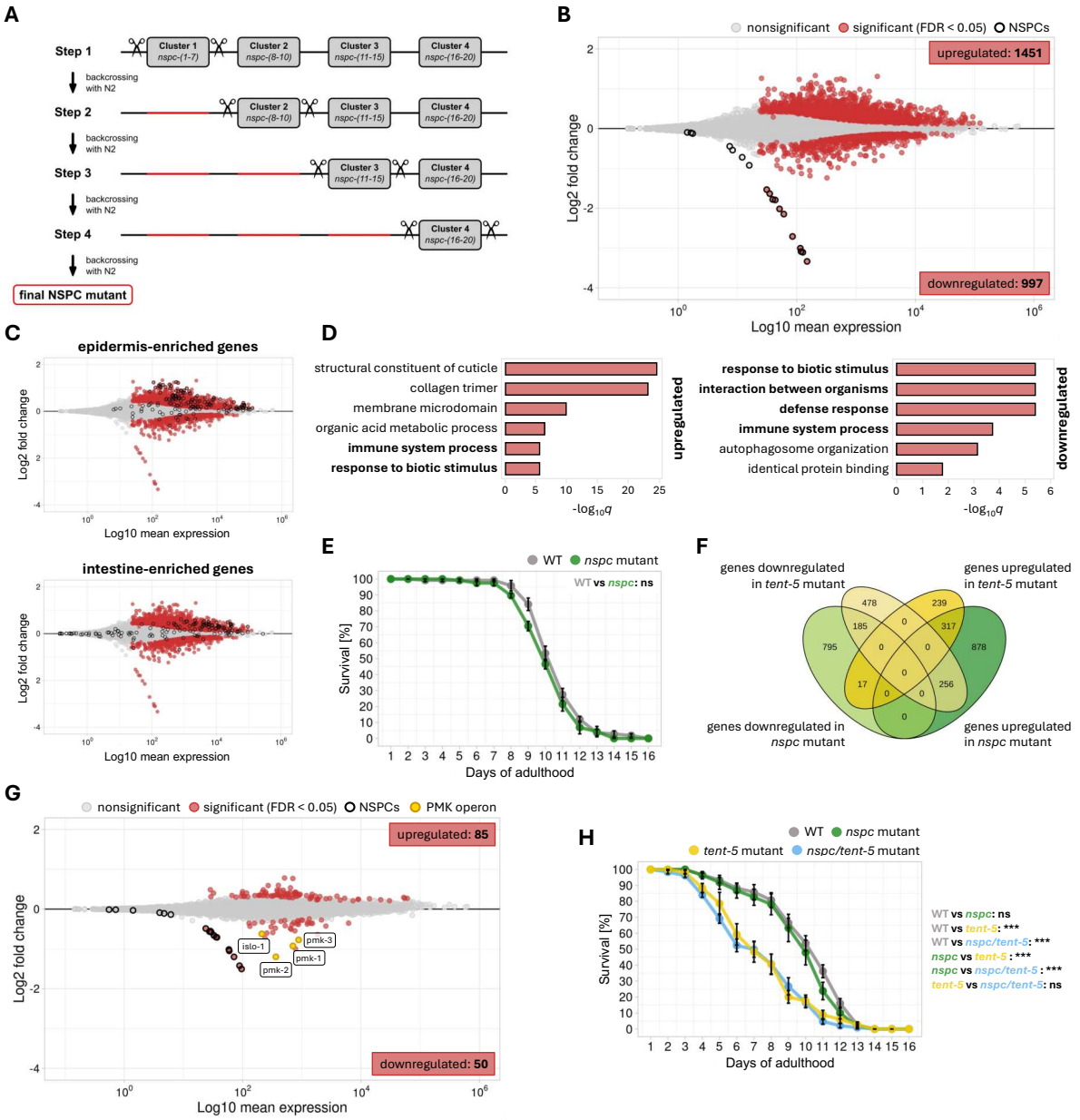


Figure 4. NSPC deletion does not result in any physiological phenotypes.

(A) Scheme illustrating the design and generation of the null *nspc* mutant. Clusters depicted in Figure 1A were subsequently deleted using CRISPR/Cas9 technology. After each deletion, worms were outcrossed with wild-type males.

(B) MA plots showing differential gene expression between wild-type and *nspc* mutant worms. Significantly changed genes (FDR < 0.05) are marked with red dots. Black borderline highlights *nspc* genes.

(C) Duplicates of MA plot in **(B)** showing the upregulation of epidermis- and intestine-enriched genes in *nspc* mutant worms. Marked genes are listed in Supplementary Table S3 and are derived from CeNGEN database [18,38].

(D) Top GO terms for the 200 most upregulated (left) and downregulated (right) genes in the *nspc* mutant, ordered by adjusted *p*-value. Terms related to the defense response are shown in bold.

(E) Survival of wild-type (gray) and *nspc* mutant (green) worms grown on non-pathogenic *E. coli* HB101. Points represent mean values from 6 separate plates (n≅20 each) with SD. ns => not significant (log-rank test).

(F) Venn diagram showing overlaps between upregulated (darker shade) and downregulated (lighter shade) genes in *tent-5* (yellow) and *nspc* mutant (green) worms. Included are all significantly changed genes (FDR < 0.05).

(G) MA plots showing differential gene expression between *tent-5* and *nscp/tent-5* mutant worms. Significantly changed genes (FDR < 0.05) are marked with red dots. Black borderline marks *nspc* genes. Yellow dots with an orange borderline show significantly downregulated genes from a single operon: *islo-1*, *pmk-1*, *pmk-2*, and *pmk-3*.

(H) Survival of wild-type (gray), *nspc* mutant (green), *tent-5* mutant (yellow), and *nspc/tent-5* mutant (blue) worms grown on pathogenic *P. aeruginosa* PAO1. Points represent mean values from 6 separate plates (n≅20 each) with SD. ns => not significant; *** => *p*-value < 0.001 (log-rank test).

NSPCs and TENT-5 might be involved in regulating cholesterol metabolism

Extensive polyadenylation of *nspc* mRNAs by TENT-5 and the strong, cell-specific expression of NSPCs in the excretory gland cell indicate a functional link between TENT-5 and NSPCs, suggesting their potential importance for excretory gland cell activity. Unfortunately, our research indicates that while TENT-5 plays a crucial role in the *C. elegans* defense response, NSPCs do not, challenging earlier assumptions about their antimicrobial activity.

However, in our previous work, we speculated that TENT-5 might also be involved in regulating cholesterol metabolism [39]. In *C. elegans* males, cholesterol levels are significantly higher in male-specific tissues [40], where TENT-5 activity is strongest. Interestingly, cholesterol is also abundant in the excretory gland cell [40], leading us to hypothesize that TENT-5 activity may be exaggerated in cholesterol-rich tissues, influencing cholesterol metabolism in worms. Additionally, cholesterol is known to influence the p38/PMK-1 MAPK pathway [41,42], which is critical for innate immune responses in worms, further linking this hypothesis to our previous observation that PMK genes are significantly upregulated in *tent-5* mutants. Given that TENT-5 positively regulates the expression of NSPCs, we hypothesized that NSPCs could also be involved in cholesterol regulation. To explore this possibility, we conducted lifespan experiments on non-pathogenic *E. coli* and pathogenic *P. aeruginosa* for four *C. elegans* strains (wild type, *tent-5*, *nspc*, and *nspc*/*tent-5* mutants) under standard conditions and with varying cholesterol concentrations (Figure 5A-D). As expected, lower cholesterol levels generally reduced worm lifespan under both conditions (Figure 5A and B). Interestingly, when fed on non-pathogenic bacteria, *nspc* mutants performed slightly better than other strains at low cholesterol levels (Figure 5C), suggesting that NSPCs might negatively regulate cholesterol uptake or metabolism. Interestingly, *tent-5* mutants showed a shortened lifespan on both low and normal cholesterol plates (Figure 5A), indicating that TENT-5 might be required to facilitate some cholesterol-related processes. Finally, when both *tent-5* and *nspc* genes were deleted, the opposing effects of these proteins mitigated the impact of cholesterol on nematode lifespan (Figure 5A). However, these trends were not observed under pathogenic conditions, possibly because the effects of infection outweighed cholesterol-related changes (Figure 5B and D). Nonetheless, the difference between wild-type and *nspc* mutants compared to *tent-5* and *nspc*/*tent-5* mutants was slightly smaller after exposure to high cholesterol, suggesting that cholesterol might contribute to pathogen resistance in *C. elegans* (Figure 5D).

In conclusion, while the lifespan experiments are not straightforward, they do not rule out a potential role for TENT-5 and NSPCs in regulating cholesterol metabolism. This regulation could involve cholesterol uptake from the environment, its intracellular transport, or further chemical modifications.

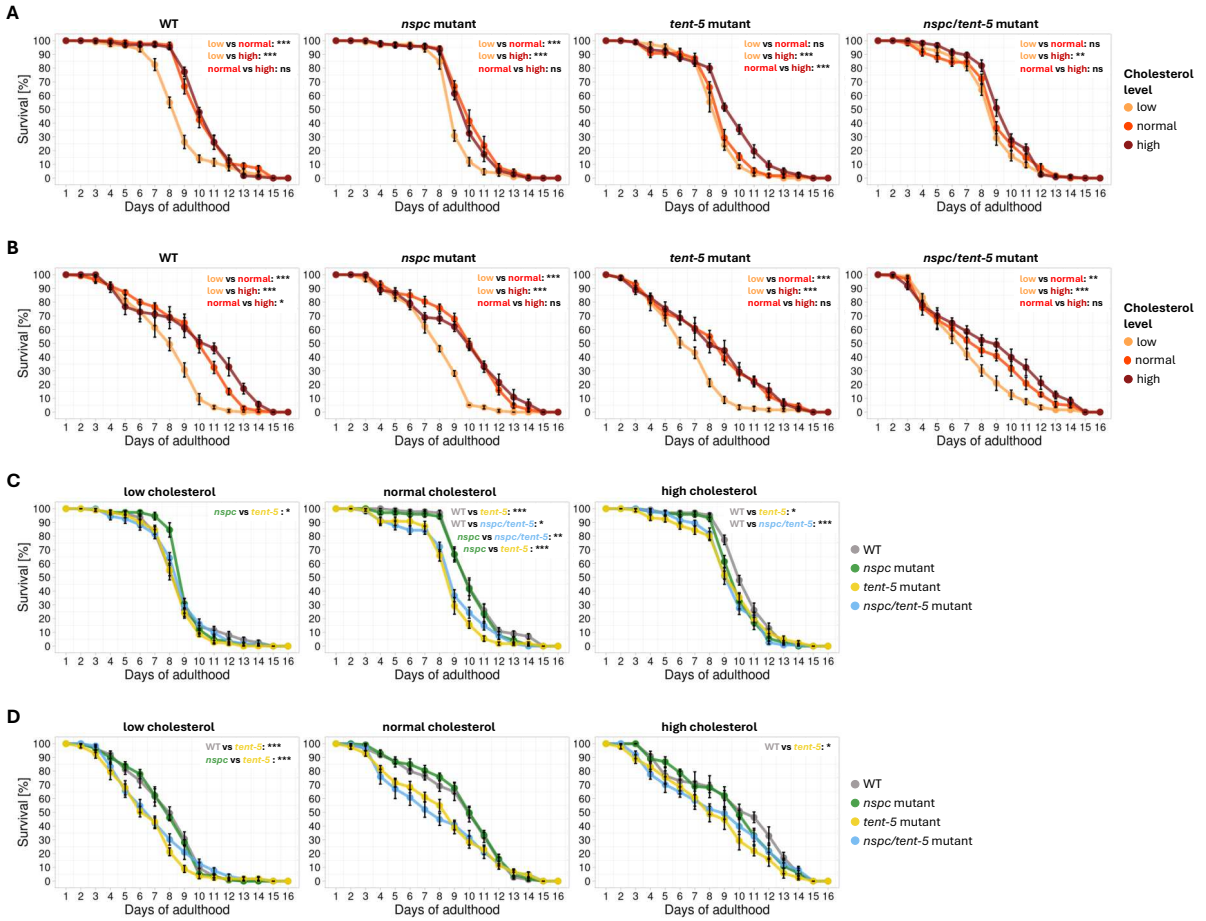


Figure 5. NSPCs might be involved in cholesterol metabolism.

(A) Survival plots for wild-type, *nspc* mutant, *tent-5* mutant, and *nspc/tent-5* mutant worms grown on non-pathogenic *E. coli* HB101 plates supplemented with three different cholesterol concentrations: low (0 mg/ml; orange), normal (20 mg/ml; red), and high (80 mg/ml; brown). Points represent mean values from 6 separate plates (n≅20 each) with SD. ns => not significant; ** => *p*-value < 0.01; *** => *p*-value < 0.001 (log-rank test).

(B) Survival plots for wild-type, *nspc* mutant, *tent-5* mutant, and *nspc/tent-5* mutant worms grown on pathogenic *P. aeruginosa* PAO1 plates supplemented with three different cholesterol concentrations: low (0 mg/ml; orange), normal (20 mg/ml; red), and high (80 mg/ml; brown). Points represent mean values from 6 separate plates (n≅20 each) with SD. ns => not significant; * => *p*-value < 0.05; ** => *p*-value < 0.01; *** => *p*-value < 0.001 (log-rank test).

(C) Survival plots for low (0 mg/ml), normal (20 mg/ml), and high (80 mg/ml) cholesterol plates with non-pathogenic *E. coli* HB101 bacteria. Survival is compared among four strains: wild-type (gray), *nspc* mutant (green), *tent-5* mutant (yellow), and *nspc/tent-5* mutant (blue) worms. Points represent mean values from 6 separate plates (n≅20 each) with SD. Only significant comparisons are shown on the plots. * => *p*-value < 0.05; ** => *p*-value < 0.01; *** => *p*-value < 0.001 (log-rank test).

(D) Survival plots for low (0 mg/ml), normal (20 mg/ml), and high (80 mg/ml) cholesterol plates with pathogenic *P. aeruginosa* PAO1 bacteria. Survival is compared among four strains: wild-type (gray), *nspc* mutant (green), *tent-5* mutant (yellow), and *nspc/tent-5* mutant (blue) worms. Points represent mean values from 6 separate plates (n≅20 each) with SD. Only significant comparisons are shown on the plots. * => *p*-value < 0.05; *** => *p*-value < 0.001 (log-rank test).

NSPC knockdown results in transcription signatures related to the DAF-2/DAF-16 insulin signaling

To further explore the potential functions of NSPCs, we analyzed high-throughput datasets from WormBase to identify conditions under which NSPCs become dysregulated (Supplementary Table S1). Our analysis showed that NSPCs are frequently dysregulated in studies involving the insulin DAF-2/DAF-16 pathway (Figure 6A). Following this lead, we examined the structures and sequences of NSPCs and noticed their similarity to insulin-like peptides from the INS family (Figure 6B and C). In *C. elegans,* these insulin peptides interact with the DAF-2 receptor, influencing the cytoplasmic or nuclear localization of the transcription factor DAF-16, which regulates key physiological processes such as lifespan, innate immunity, and metabolism [43–46] (Figure 6A). The similarity between NSPCs and insulin peptides led us to hypothesize that NSPCs might serve as neuropeptides that co-regulate the insulin signaling pathway, potentially in the excretory gland cell. Their potential redundancy with INS proteins could explain the lack of strong phenotypes following *nspc* cluster deletion. To test this hypothesis, we crossed a GFP-tagged DAF-16 strain [47] with both *tent-5* and *nspc* mutants and examined DAF-16-GFP localization. We observed no significant differences in DAF-16 localization across cell types in standard conditions, as well as after heat shock, pathogen infection, and varying cholesterol supply (Figure 6D, Supplementary Figure S5), suggesting that NSPCs and TENT-5 may not be essential for maintaining proper insulin signaling.

Next, we generated a strain lacking both *daf-16* and *nspc* genes and conducted RNA-seq to gain further insight into their possible co-regulation. Interestingly, we observed substantial overlap in genes upregulated upon deletion of either *nspc* genes or *daf-16* (Figure 6E, Supplementary Table S5), supporting the hypothesis of NSPC-mediated regulation of the insulin pathway and suggesting that NSPCs act antagonistically towards DAF-2. Consistently, our sequencing results showed that while DAF-16 depletion typically causes significant gene dysregulation, this effect is less pronounced in the absence of NSPCs (Figure 6F, Supplementary Table S5). In contrast, NSPC deletion in the absence of DAF-16 resulted in a transcriptomic shift similar to that seen with NSPC deletion alone, except for some genes that were upregulated when DAF-16 was present but downregulated in its absence (Figure 6G, Supplementary Table S5). These results suggest that NSPCs may act upstream of DAF-16, but their role is likely not limited to regulating insulin signaling. For example, defense response genes remained upregulated in the absence of NSPCs, regardless of DAF-16 presence (Figure 6H), indicating that these processes might be regulated independently of insulin signaling. Consequently, the deletion of NSPCs in *daf-16* mutants did not lead to any changes in worms' survival for both non-pathogenic *E. coli* and pathogenic *P. aeruginosa* conditions (Figure 6I). Although our RNA-seq results support the potential role of NSPCs as neuropeptides in the DAF-2/DAF-16 pathway, their possible redundancy with other small peptides and the complexity of the signaling machinery might mask the potential physiological effect and make it difficult to definitively prove this hypothesis. A comprehensive study would require generating a strain lacking both NSPCs and all INS peptides; however, this is beyond the scope of the current work.

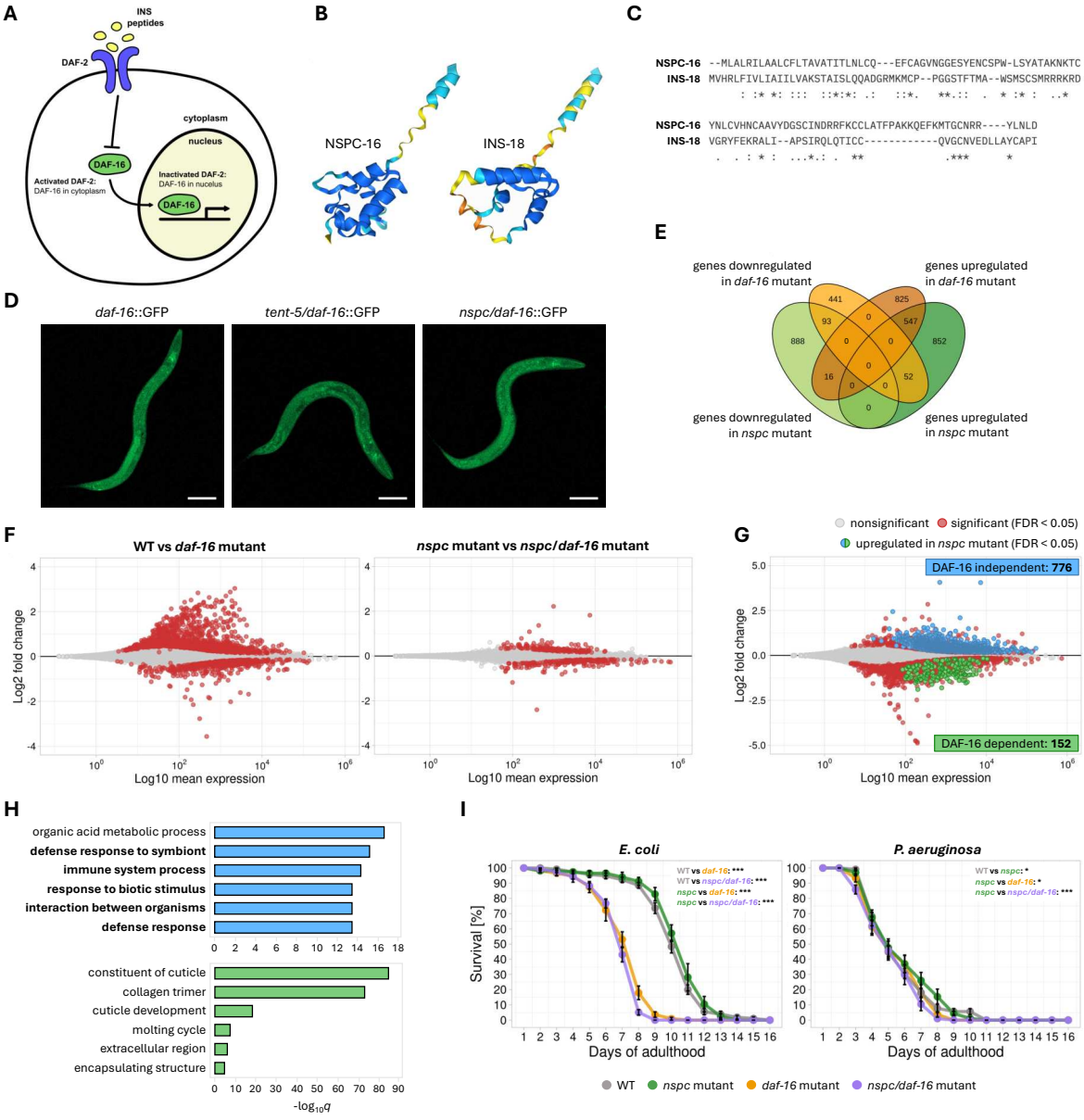


Figure 6. NSPCs might be involved in insulin signaling.

(A) Schematic representation of the DAF-2/DAF-16 insulin signaling pathway. Insulin peptides (INS) act as ligands for the DAF-2 receptor. Upon DAF-2 activation, DAF-16 is retained in the cytoplasm; when DAF-2 is inactivated, DAF-16 translocates to the nucleus to initiate DAF-16-mediated gene expression.

(B) Comparison of NSPC-16 and INS-18 protein structures predicted by AlphaFold 3 [48]. Representatives of NSPC and INS families were chosen arbitrarily. Protein similarity was assessed visually without the use of any quantitative parameters.

(C) Comparison of NSPC-16 and INS-18 protein sequences. Asterisks (*) represent residues that are conserved in all NSPCs; colons (:) represent moderately conserved residues; and dots (.) indicate non-conserved residues with similar chemical properties.

(D) Fluorescence microscopy images of DAF-16-GFP localization in wild-type, *tent-5*, and *nspc* mutant worms in standard growth conditions. Scale bars => 100 µm.

(E) Venn diagram showing overlaps between upregulated (darker shade) and downregulated (lighter shade) genes in *daf-16* (orange) and *nspc* mutant (green) worms. Included are all significantly changed genes (FDR < 0.05).

(F) MA plots showing differential gene expression after DAF-16 depletion in the presence of NSPCs (left) and in their absence (right). Significantly changed genes (FDR < 0.05) are marked with red dots.

(G) MA plot showing differential gene expression between *daf-16* and *nspc/daf-16* mutant worms. Significantly changed genes (FDR < 0.05) are marked with red dots. Genes upregulated in *nspc* mutant that are DAF-16 independent are marked in blue, and DAF-16-dependent are shown in green. Only genes significantly upregulated

in *nspc* mutant in both the presence (FDR < 0.05; Figure 4B; Supplementary Table S3) and absence of DAF-16 (Supplementary Table S5) were included.
(H) Top GO terms for DAF-16 independent (top; blue) and dependent (bottom; green) genes upregulated in *nspc* mutant ordered by adjusted *p*-value. Terms related to the defense response are shown in bold.
(I) Survival of wild-type (gray), *nspc* mutant (green), *daf-16* mutant (orange), and *nspc/daf-16* mutant (purple) worms grown on non-pathogenic *E. coli* HB101 (left) or pathogenic *P. aeruginosa* PAO1 (right). Points represent mean values from 6 separate plates (n≅20 each) with SD. Only significant comparisons are shown on the plots. * => *p*-value < 0.05; *** => *p*-value < 0.001 (log-rank test).

Conclusions

The research presented in this manuscript revealed that one of the main functions of the excretory gland cell is the production of secreted nematode-specific NSPC proteins. This family, which encompasses 18 highly similar genes, is unique to the *Caenorhabditis* genus and was previously suggested to function as an immune effectors. Here, we show that they might function rather as neuropeptides that are potentially involved in regulating signaling pathways, such as the insulin DAF-2/DAF-16 pathway. Although these findings do not entirely clarify the physiological role of the excretory gland cell, they lay a foundation for future research into how this cell and its associated NSPCs contribute to nematode biology, particularly under ecologically relevant conditions.

From the evolutionary perspective, it is interesting to consider the function of the excretory gland cell in nematode species that lack NSPCs. Although limited studies exist, the excretory gland cell has been suggested to secrete components of epicuticle in *Toxocara canis* [49], enzymes important for feeding in *Nippostrongylus brasiliensi* [50], or peptidases required for molting in *Phocanema decipiens* [51]. While the first two possibilities remain unexplored in *C. elegans*, the role of the excretory gland cell in molting has already been excluded [11]. It appears that in *C. elegans*, the excretory gland cell has evolved to perform a distinct role, possibly because of inhabiting a different natural environment in comparison to the mentioned parasitic nematodes. That change was accompanied by the generation of *Caenorhabditis*-specific NSPCs, further supporting their importance for the cell's function.

Methods

C. elegans culture and growth conditions

All *C. elegans* strains used in this study are summarized in Supplementary Table S6. Unless specified otherwise, worms were maintained at 20°C on nematode growth medium (NGM) plates seeded with *E. coli* HB101 as a food source.

Plasmids construction

All plasmids used in this study are presented in Supplementary Table S7. Plasmids pVL033 and pVL037-40 were constructed using sequence- and ligation-independent cloning (SLIC), following established protocols. The VL033 plasmid was built by introducing *tent-5* in fusion with GFP-3xflag (amplified from respective BAC) into the pCFJ151 backbone. Plasmids pVL037-40 were assembled from the following fragments: the pCFJ104 backbone, the *nspc-9, -14, or -20* promoter with or without gene CDS (amplified from genomic DNA isolated from wild-type worms), the mCherry reporter (from pCFJ104), and the *nspc-9, -14, or -20* 3′ UTR (also from wild-type genomic DNA). PCR amplification of each fragment was performed using Phusion polymerase with an initial denaturation at 98°C for 2 min, followed by 25 cycles of 98°C for 10 s, 60°C for 30 s, and 72°C for 10 s, and a final elongation step at 72°C for 5 min.
Plasmids pCE090 and pCE091 were prepared through classical restriction enzyme digestion and ligation methods. The *nspc-7* and *nspc-14* coding sequences were amplified from wild-type *C. elegans* cDNA and initially cloned into the pJET1.2 vector, following the manufacturer's

protocol. Both L4440 and pJET1.2 vectors containing *nspc-7* or *nspc-14* were digested with XhoI and XbaI and subsequently ligated with T4 DNA ligase. The pCE092 plasmid, containing *nspc-7*, *nspc-14*, and *nspc-20* sequences, was assembled with SLIC using an L4440 backbone containing the *nspc-14* sequence (from pCE091), *nspc-7* sequence (from pCE090), and *nspc-20* (from wild-type cDNA). All plasmids were validated by restriction digestion and Sanger sequencing.

NSPC localization

Localization of NSPCs and TENT-5 was assessed using an Olympus FV1000 confocal microscope with a 60x/1.2 water immersion lens. Images were processed using Fiji/ImageJ software [52].

Generation of transgenic C. elegans strains

All new *C. elegans* transgenic strains were generated by microinjection into the hermaphrodite gonad using either extrachromosomal arrays (plasmid-based) or CRISPR/Cas9-mediated genome editing (via guide RNA and repair template injection). Microinjections were conducted on Axio Observer 5 microscope.

Strains ADZ31, ADZ33, ADZ35, ADZ37, and ADZ44 were produced by injecting wild-type worms with a mix containing pVL037, pVL039, pVL038, pVL040, or pVL073 respectively (20 ng/µl), pRH269 (5 ng/µl), and GeneRuler DNA mix (75 ng/µl). Progeny were screened for green fluorescence in the pharynx and red fluorescence in the excretory gland cells, and transgenic worms were regularly enriched by isolating GFP-positive individuals. For the ADZ45 strain, pVL038 and pVL045 were injected without the co-injection marker.

Strains ADZ72, ADZ83, and ADZ111-114 were generated using a modified co-CRISPR approach, based on established protocols. Injection mixes contained 4 µM Cas9, 10 µM specific guide/s, 4 µM *dpy-10* guide RNA, 0.4 µM repair template, and 0.2 µM *dpy-10* repair template. To create strain ADZ72, the *nspc-10* coding sequence was removed using two guide RNAs and repaired by a template containing the first *nspc-10* intron, *tomm-20*, and miniSOG(426Cys) sequences. The ADZ82 strain was generated by introducing a single nucleotide substitution in miniSOG present in ADZ72, converting Cys426 to Gly and thus activating the miniSOG protein. For the ADZ111 strain, the first NSPC cluster was removed, and successive deletions were conducted until the final strain lacking all NSPC clusters was obtained. In all cases, worms were co-injected with *dpy-10* guide and repair template introducing a visible roller or dumpy phenotype in affected worms, enabling easier selection of mutants for genotyping. After homozygote selection, worms were outcrossed with wild-type males to eliminate any potential off-target mutations. All CRISPR guide RNAs and repair templates are listed in Supplementary Table S8. The repair template for strain ADZ72 was generated by PCR from pCZGY1703 plasmid, using primers with 35-50 nucleotide homology arms. The PCR reaction was performed using Phusion polymerase with an initial denaturation at 98°C for 2 min, followed by 25 cycles of 98°C for 10 s, 60°C for 30 s, and 72°C for 10 s, and a final elongation step at 72°C for 5 min. PCR products were purified using Gel-out and KAPA Pure magnetic beads, and prepared for microinjection.

Additional strains, including ADZ74, ADZ83, ADZ116, ADZ122, ADZ130, and ADZ131, were generated by crossing previously established strains as outlined in Supplementary Table S6.

Optogenetic ablation

To induce optogenetic cell ablation, age-synchronized populations of wild-type and active/inactive miniSOG worms were cultured on NGM plates at 15°C until the L3 stage. Plates were then exposed to blue light (460 nm) from an LED advertising panel. The plates were

inverted and placed without lids on the panel, with a 1 mm glass separator to minimize the transfer of heat. The illumination process was performed at ambient temperatures of 18-20°C and included five intervals of 10 minutes of light, separated by 10-minute breaks. After illumination, worms were transferred back to 15°C and grown for two days to allow complete miniSOG-mediated cell destruction. Ablation success was verified by detecting the loss of mCherry fluorescence in the excretory gland cell and measuring the expression of *nspc* genes by RT-qPCR. All imaging was conducted on a Zeiss LSM800 confocal microscope with a 40×/oil immersion lens.

Western blotting

Similarly dense population of worms expressing active or inactive miniSOG were harvested and washed three times with 50 mM NaCl. Worm pellets were mixed with zirconia-silica beads and 3x SDS sample buffer (187.5 mM Tris-HCl (pH 6.8), 6% SDS, 150 mM DTT, 0.02% bromophenol blue, 30% glycerol, and 3% 2-mercaptoethanol), vortexed for 5 minutes, and incubated at 95°C for 5 minutes. Samples were centrifuged, and the supernatant was run on 15% SDS-PAGE gels. Proteins were transferred onto Protran nitrocellulose membranes through semi-dry transfer at 15V for 1 hour. Following the transfer, membranes were blocked with 5% non-fat milk in 1x TBST (20 mM Tris-HCl, 150 mM NaCl, 0.01% Tween 20) for 1 hour at room temperature. Membranes were then incubated with primary antibodies for miniSOG (1:1000; Kerafast; EFH004) and α-tubulin (1:10 000; Sigma-Aldrich; CP06) overnight at 4°C. After washing with 1x TBST, membranes were incubated with secondary anti-mouse antibodies (1:5000 for miniSOG, 1:10000 for α-tubulin; Sigma-Aldrich; 401215) for 1 hour at room temperature, activated with Clarity Western ECL Substrate (Bio-Rad), and visualized using CL-Exposure film (Thermo Fisher Scientific) developed with an AGFA Curix CP-1000 device.

RNA interference (RNAi)

RNA interference (RNAi) was used to silence NSPC genes in *C. elegans*. Plates for RNAi were additionally supplemented with 1 mM IPTG (Sigma-Aldrich), 25 µg/ml carbenicillin (Sigma-Aldrich), and 25 U/ml nystatin (Sigma-Aldrich). *E. coli* HT115 bacteria transformed with RNAi plasmids were prepared by overnight culture, centrifuged, concentrated 15x times, and seeded onto RNAi plates, which were left at room temperature overnight. Age-synchronized L4 worms (approximately 30 per 100-mm plate) were transferred to the RNAi plates, and their progeny were used for subsequent analyses. Silencing efficiency was verified using RT-qPCR.

RNA isolation

Populations of worms were collected and washed three times with 50 mM NaCl. Worm pellets were resuspended in 1 ml TRI Reagent (Sigma-Aldrich), vortexed for 15 minutes at room temperature, and stored at -80°C until further processing. RNA extraction was performed according to the manufacturer's protocol for TRIzol-based isolation, with additional purification using KAPA Pure magnetic beads (RNA to beads ratio: 1:3 v/v). RNA quality and integrity were confirmed using the Agilent TapeStation system. RNA samples were prepared in three independent biological replicates for downstream RT-qPCR and Illumina RNA-seq, and two replicates were used for direct RNA sequencing.

RT-qPCR

Expression levels of NSPCs following excretory gland cell ablation and RNAi silencing were evaluated using RT-qPCR. Total RNA was treated with TURBO DNase (Thermo Fisher Scientific) to eliminate genomic DNA. cDNA synthesis was conducted following the manufacturer's protocol with the use of SuperScript III (Thermo Fisher Scientific) and a mix

of oligo(dT)$_{20}$ and random primers. RT-qPCR reactions were carried out using Platinum SYBR Green Mix (Thermo Fisher Scientific) on a QuantStudio 5 PCR system (Thermo Fisher Scientific) with primers targeting NSPCs (Supplementary Table S8). Due to high similarity in NSPC sequences, primers were designed to detect groups of genes: one primer pair for *nspc-(1-15),* and another for *nspc-(16-20).* Relative gene expression was normalized to *act-1* or *rps-23* and quantified using the $2^{-\Delta\Delta C(t)}$ method. Statistical significance was determined by two-tailed unpaired Student's t-tests.

Illumina RNA sequencing and data analysis

RNA samples were treated with TURBO DNase (Thermo Fisher Scientific) before sequencing. Libraries were prepared from 2 µg of poly(A)-enriched RNA and sequenced on an Illumina platform in 150-bp paired-end mode (performed by the Center for New Technology, University of Warsaw or GENEWIZ, Germany). Reads were mapped to the *C. elegans* WBCel235 reference genome using STAR aligner v2.7.10a [53] and processed with samtools v1.961 [54]. Read counts were obtained with featureCounts from the Subread package v2.0.6 [55] (options -Q 10 -p -C -B). Differential expression analysis was conducted using DESeq2 Bioconductor package v1.28 [56] with default settings. Gene ontology (GO) enrichment analysis was performed using WormBase Enrichment Suite [57]. Raw sequencing data were deposited in the Gene Expression Omnibus (GEO) database under accession number GSE282286.

Nanopore Direct RNA sequencing and data analysis

Libraries were generated using the Direct RNA Sequencing Kit (SQK-RNA002, Oxford Nanopore Technologies), using 1.5 µg of total RNA with added in vitro transcribed poly(A) standards (2 ng). Sequencing was conducted on a MinION device and followed by basecalling by Guppy v6.0.0. Reads were aligned to the WBCel235 reference using MiniMap v2.17 [58] (options -k 14 -ax map-ont –secondary=no) and processed with samtools v1.9 [54]. Poly(A) tail lengths were estimated using Nanopolish v0.13.2 as described before [13,59]. Distribution comparisons across conditions were assessed by the Wilcoxon test, with p-values adjusted via the Benjamini-Hochberg method. Raw DRS data are available in the GEO database under accession number GSE282286.

Worms age estimations

Worm age was estimated for samples used in RNA-seq experiments using the R packages RAPToR v1.2 and wormRef v0.5 [26]. Given the worms' stages (L4 for NSPC mutants and young adults for optogenetic ablations), the *"Cel_YA_1"* reference was selected from wormRef as appropriate for both experiments. Age estimations were conducted using *ae* function with normalized gene counts as an input. Statistical differences were tested by a two-tailed unpaired Student's t-test.

Lifespan experiments

Lifespan assays were conducted on 35-mm NGM plates supplemented with 50 µM 5-fluoro-2′-deoxyuridine (FUdR) to inhibit reproduction and maintained at 25°C. To assess NSPC function in cholesterol metabolism, plates were prepared with varying cholesterol concentrations of 0 mg/ml (low), 20 mg/ml (standard), and 80 mg/ml (high). All plates were seeded with either non-pathogenic *E. coli* HB101 or pathogenic *P. aeruginosa* PAO1 (100 µl/plate) and incubated overnight at 37°C. Age-synchronized *C. elegans* populations were prepared by bleaching and seeding embryos on plates. Once worms reached the L4 stage, they were transferred to the experimental plates (20 worms per plate, six replicates per condition). The number of live and dead worms was recorded daily until all worms had died. Survival data were analyzed by the Kaplan-Meier method, with condition comparisons assessed via log-rank significance tests.

Brood-size analysis

To evaluate reproductive potential post-ablation, progeny counts were done for wild-type, inactive miniSOG (ADZ74), and active miniSOG (ADZ83) worms, with and without illumination. Worms were synchronized by bleaching, grown to L3 at 15°C, and half of the population was illuminated as described. After 24 hours, five worms per condition were placed on fresh plates seeded with *E. coli* HB101 and later grown at 20ºC. Worms were transferred to new plates daily throughout their entire lifespan. Progeny counts were compared across conditions by a two-tailed unpaired Student's t-test. For RNAi-silenced worms, assays followed a similar approach but with HT115 RNAi bacterial plates and worms' preparation steps used in RNAi protocol.

DAF-16 localization

Localization of DAF-16 was visualized in strains OH16024 (*daf-16*::*GFP)* [47], ADZ116 (*tent-5/daf-16::GFP*), and ADZ122 (*nspc/daf-16::GFP*) under standard conditions upon various stimuli. For standard imaging, synchronized L4 worms were immobilized with 25 µM levamisole on freshly prepared 2% agarose pads and imaged immediately. For studying DAF-16 nuclear translocation after heat shock, synchronized L4 worms were exposed to 35ºC for 1 hour before imaging. For studying DAF-16 delocalization from the nucleus to the cytoplasm upon infection, worms after heat shock were either maintained on standard *E. coli* HB101 or switched to *P. aeruginosa* PAO1 for 16 hours before imaging as described before [60]. To investigate the effects of cholesterol concentration on DAF-16 localization, worms were grown on NGM plates containing low (0 mg/ml) or high (80 mg/ml) cholesterol before standard imaging. All imaging was conducted on a Zeiss LSM800 confocal microscope with a 20x objective.

Sequence alignment and phylogenetic analysis

Amino acid sequences of all NPSC and INS peptides were retrieved from WormBase [19]. Alignments and phylogenetic analysis were performed using the MUSCLE tool from EMBL-EBI [20].

Data availability

The raw data from Illumina RNA Sequencing and Nanopore Direct RNA Sequencing have been deposited in the Gene Expression Omnibus (GEO) database under the accession number GSE282286.

Acknowledgments

We are grateful to the members of the Dziembowski group and Krzysztof Drabikowski for the discussion and valuable feedback. We especially thank Paweł Krawczyk for helping with bioinformatic analyses, Aleksandra Brouze and Seweryn Mroczek for running Nanopore DRS sequencing, Jonathan J. Ewbank and Krzysztof Drabikowski for sharing plasmids, and Tomasz Węgierski for supporting microscopic observations.

Funding

This work was supported by the National Science Center (OPUS 17 UMO-2019/33/B/NZ2/01773) and Horizon Europe (European Research (AdG nr 101097317). Some *C. elegans* strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440).

Author contributions

A.D. provided funding and directed the project. Z.M., V.L., and A.D. designed the experiments. V.L. generated the strains ADZ31-ADZ45 and performed visualizations of NSPC localization and its colocalization with TENT-5. Z.M. prepared the remaining strains, carried out the remaining experiments, and analyzed all sequencing data. Z.M. prepared the initial draft of the manuscript and created all figures. V.L. and A.D. revised and edited the manuscript.

References

1. Brenner S. The Genetics of Caenorhabditis Elegans. Genetics. 1974;77: 71–94. doi: 10.1093/genetics/77.1.71
2. The C. elegans Sequencing Consortium. Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. Science. 1998;282: 2012–2018. doi: 10.1126/science.282.5396.2012
3. Cook SJ, Jarrell TA, Brittin CA, Wang Y, Bloniarz AE, Yakovlev MA, et al. Whole-animal connectomes of both Caenorhabditis elegans sexes. Nature. 2019;571: 63–71. doi:10.1038/s41586-019-1352-7
4. White JG, Southgate E, Thomson JN, Brenner S. The Structure Of The Nervous System Of The Nematode Caenorhabditis Elegans. Phil Trans R Soc Lond. 1986;314: 1–340. doi: 10.1098/rstb.1986.0056
5. Kimble J, Hirsh D. The Postembryonic Cell Lineages of the Hermaphrodite and Male Gonads in Caenorhabditis elegans. Dev Biol. 1979;70: 396–417. doi: 10.1016/0012-1606(79)90035-6
6. Sulston JE, Horvitz HR. Post-embryonic Cell Lineages of the Nematode, Caenorhabditis elegans. Dev Biol. 1977;56: 110–156. doi: 10.1016/0012-1606(77)90158-0
7. Sulston JE, Schierenberg E, White JG, Thomson JN. The Embryonic Cell Lineage of the Nematode Caenorhabditis elegans. Dev Biol. 1983;100: 64–119. doi: 10.1016/0012-1606(83)90201-4
8. Sundaram M V., Buechner M. The Caenorhabditis elegans excretory system: A model for tubulogenesis, cell fate specification, and plasticity. Genetics. 2016;203: 35–63. doi:10.1534/genetics.116.189357
9. Nelson FK, Albert PS, Riddle DL. Fine Structure of the Caenorhabditis elegans Secretory-Excretory System. J Ultrastruct Res. 1983;82: 156–171. doi: 10.1016/s0022-5320(83)90050-3
10. Nelson FK, Riddle DL. Functional Study of the Caenorhabditis elegans Secretory-Excretory System Using Laser Microsurgery. J Exp Zool. 1984;231: 45–56. doi: 10.1002/jez.1402310107
11. Singh RN, Sulston J.E. Some Observations On Moulting In Caenorhabditis Elegans. Nematologica. 1978; 63–71. doi: 10.1163/187529278X00074
12. Altun ZF, Hall DH. Excretory system. WormAtlas. 2009. doi:10.3908/wormatlas.1.17
13. Liudkovska V, Krawczyk PS, Brouze A, Gumińska N, Wegierski T, Cysewski D, et al. TENT5 cytoplasmic noncanonical poly(A) polymerases regulate the innate immune response in animals. Sci Adv. 2022;8. doi:10.1126/sciadv.add9468
14. Thomas JH. Concerted evolution of two novel protein families in Caenorhabditis species. Genetics. 2006;172: 2269–2281. doi:10.1534/genetics.105.052746
15. Engelmann I, Griffon A, Tichit L, Montañana-Sanchis F, Wang G, Reinke V, et al. A comprehensive analysis of gene expression changes provoked by bacterial and fungal infection in C. elegans. PLoS One. 2011;6. doi:10.1371/journal.pone.0019055
16. Sinha A, Rae R, Iatsenko I, Sommer RJ. System Wide Analysis of the Evolution of Innate Immunity in the Nematode Model Species Caenorhabditis elegans and Pristionchus pacificus. PLoS One. 2012;7. doi:10.1371/journal.pone.0044255

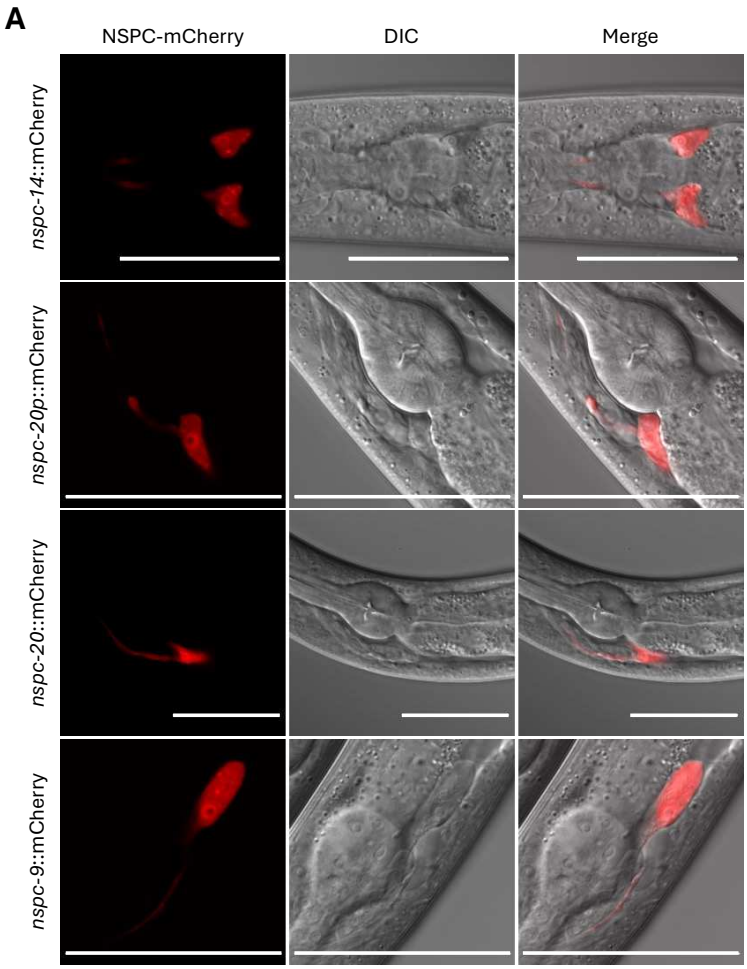
17. Sellegounder D, Liu Y, Wibisono P, Chen C-H, Leap D, Sun J. Neuronal GPCR NPR-8 regulates C. elegans defense against pathogen infection. Sci Adv. 2019;5. doi:10.1126/sciadv.aaw4717
18. Taylor SR, Santpere G, Weinreb A, Barrett A, Reilly MB, Xu C, et al. Molecular topography of an entire nervous system. Cell. 2021;184: 4329-4347.e23. doi:10.1016/j.cell.2021.06.023
19. WormBase. [cited 12 Nov 2024]. Available: https://wormbase.org/
20. Madeira F, Madhusoodanan N, Lee J, Eusebi A, Niewielska A, Tivey ARN, et al. The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. Nucleic Acids Res. 2024;52: W521–W525. doi:10.1093/nar/gkae241
21. Qi YB, Garren EJ, Shu X, Tsien RY, Jin Y. Photo-inducible cell ablation in Caenorhabditis elegans using the genetically encoded singlet oxygen generating protein miniSOG. Proc Natl Acad Sci U S A. 2012;109: 7499–7504. doi:10.1073/pnas.1204096109
22. Xu S, Chisholm AD. Highly efficient optogenetic cell ablation in C. Elegans using membrane-targeted miniSOG. Sci Rep. 2016;6. doi:10.1038/srep21271
23. Shu X, Lev-Ram V, Deerinck TJ, Qi Y, Ramko EB, Davidson MW, et al. A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms. PLoS Biol. 2011;9. doi:10.1371/journal.pbio.1001041
24. Hubbard EJA, Greenstein D. Introduction to the germ line. WormBook. 2005. doi:10.1895/wormbook.1.18.1
25. Reinke V, Smith HE, Nance J, Wang J, Van Doren C, Begley R, et al. A Global Profile of Germline Gene Expression in C. elegans. Mol Cell. 2000;6: 605–616. doi:10.1016/S1097-2765(00)00059-9
26. Bulteau R, Francesconi M. Real age prediction from the transcriptome with RAPToR. Nat Methods. 2022;19(8): 969–975. doi:10.1038/s41592-022-01540-0
27. De Magalhaes Filho CD, Henriquez B, Seah NE, Evans RM, Lapierre LR, Dillin A. Visible light reduces C. elegans longevity. Nat Commun. 2018;9. doi:10.1038/s41467-018-02934-5
28. Reinke V, Gil IS, Ward S, Kazmer K. Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. Development. 2004;131: 311–323. doi:10.1242/dev.00914
29. Conte D, MacNei LT, Walhout AJM, Mello CC. RNA Interference in Caenorhabditis elegans. Curr Protoc Mol Biol. 2015;2015: 26.3.1-26.3.30. doi:10.1002/0471142727.mb2603s109
30. Peterson ND, Cheesman HK, Liu P, Anderson SM, Foster KJ, Chhaya R, et al. The nuclear hormone receptor NHR-86 controls anti-pathogen responses in C. elegans. PLoS Genet. 2019;15. doi:10.1371/journal.pgen.1007935
31. Yunger E, Safra M, Levi-Ferber M, Haviv-Chesner A, Henis-Korenblit S. Innate immunity mediated longevity and longevity induced by germ cell removal converge on the C-type lectin domain protein IRG-7. PLoS Genet. 2017;13. doi:10.1371/journal.pgen.1006577
32. Schulenburg H, Hoeppner MP, Weiner J, Bornberg-Bauer E. Specificity of the innate immune system and diversity of C-type lectin domain (CTLD) proteins in the nematode Caenorhabditis elegans. Immunobiology. 2008;213: 237–250. doi:10.1016/j.imbio.2007.12.004
33. Mallo G V, Lé Opold Kurz C, Couillault C, Pujol N, Granjeaud S, Kohara Y, et al. Inducible Antibacterial Defense System in C. elegans. Current Biology. 2002;12: 1209–1214. doi:10.1016/S0960-9822(02)00928-4

34. Wong D, Bazopoulou D, Pujol N, Tavernarakis N, Ewbank JJ. Genome-wide investigation reveals pathogen-specific and shared signatures in the response of Caenorhabditis elegans to infection. Genome Biol. 2007;8. doi:10.1186/gb-2007-8-9-r194
35. Kim DH, Feinbaum R, Alloing G, Emerson FE, Garsin DA, Inoue H, et al. A Conserved p38 MAP Kinase Pathway in Caenorhabditis elegans Innate Immunity. Science. 2002;297: 623–626. doi:10.1126/science.1073759
36. Shivers RP, Pagano DJ, Kooistra T, Richardson CE, Reddy KC, Whitney JK, et al. Phosphorylation of the conserved transcription factor ATF-7 by PMK-1 p38 MAPK regulates innate immunity in Caenorhabditis elegans. PLoS Genet. 2010;6. doi:10.1371/journal.pgen.1000892
37. Troemel ER, Chu SW, Reinke V, Lee SS, Ausubel FM, Kim DH. p38 MAPK regulates expression of immune response genes and contributes to longevity in C. elegans. PLoS Genet. 2006;2: 1725–1739. doi:10.1371/journal.pgen.0020183
38. CeNGEN. [cited 12 Nov 2024]. Available: https://cengen.shinyapps.io/CengenApp/
39. Mackiewicz Z, Liudkovska V, Dziembowski A. TENT-5 regulates the expression of male-specific genes in Caenorhabditis elegans. bioRxiv. 2024. doi:10.1101/2024.06.18.599341
40. Matyash V, Geier C, Henske A, Mukherjee S, Hirsh D, Thiele C, et al. Distribution and Transport of Cholesterol in Caenorhabditis elegans. Mol Biol Cell. 2001;12: 1725–1736. doi: 10.1091/mbc.12.6.1725
41. Otarigho B, Aballay A. Cholesterol Regulates Innate Immunity via Nuclear Hormone Receptor NHR-8. iScience. 2020;23. doi:10.1016/j.isci.2020.101068
42. Peterson ND, Icso JD, Salisbury JE, Rodríguez T, Thompson PR, Pukkila-Worley R. Pathogen infection and cholesterol deficiency activate the C. elegans p38 immune pathway through a TIR-1/SARM1 phase transition. Elife. 2022;11. doi:10.7554/eLife.74206
43. Lee SS, Kennedy S, Tolonen AC, Ruvkun G. DAF-16 Target Genes That Control C. elegans Life-Span and Metabolism. Science. 2003;300: 644–647. doi:10.1126/science.1083614
44. Libina N, Berman JR, Kenyon C. Tissue-Specific Activities of C. elegans DAF-16 in the Regulation of Lifespan. Cell. 2003;115: 489–502. doi:10.1016/s0092-8674(03)00889-4
45. Garsin DA, Villanueva JM, Begun J, Kim DH, Sifri CD, Calderwood SB, et al. Long-Lived C. elegans daf-2 Mutants Are Resistant to Bacterial Pathogens. Science. 2003;300: 1921. doi:10.1126/science.1080147
46. Lin K, Hsin H, Libina N, Kenyon C. Regulation of the Caenorhabditis elegans longevity protein DAF-16 by insulin/IGF-1 and germline signaling. Nat Genet. 2001;28: 139–145. doi:10.1038/88850
47. Aghayeva U, Bhattacharya A, Hobert O. A panel of fluorophore-tagged daf-16 alleles. MicroPubl Biol. 2020;2020. doi:10.17912/micropub.biology.000210
48. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024;630: 493–500. doi:10.1038/s41586-024-07487-w
49. Page AP, Hamilton AJ, Maizels RM. Toxocara canis: Monoclonal Antibodies to Carbohydrate Epitopes of Secreted (TES) Antigens Localize to Different Secretion-Related Structures in Infective Larvae. Exp Parasitol. 1992;75: 56–71. doi: 10.1016/0014-4894(92)90122-q
50. Lee DL. The Fine Structure Of The Excretory System In Adult Nippostrongylus Brasiliensis (Nematoda) And A Suggested Function For The “Excretory Glands.” Tissue Cell. 1970;2: 225–231. doi:10.1016/S0040-8166(70)80017-9

51. Davey KG, Kan SP. Molting in a parasitic nematode, Phocanemu decipiens. IV. Ecdysis and its control. Can J Zool. 1968;46: 893–898. doi: 10.1139/z68-125
52. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: An open-source platform for biological-image analysis. Nature Methods. 2012. pp. 676–682. doi:10.1038/nmeth.2019
53. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29: 15–21. doi:10.1093/bioinformatics/bts635
54. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10. doi:10.1093/gigascience/giab008
55. Liao Y, Smyth GK, Shi W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30: 923–930. doi:10.1093/bioinformatics/btt656
56. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15. doi:10.1186/s13059-014-0550-8
57. Angeles-Albores D, Raymond RY, Chan J, Sternberg PW. Tissue enrichment analysis for C. elegans genomics. BMC Bioinformatics. 2016;17. doi:10.1186/s12859-016-1229-9
58. Li H. Minimap2: Pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34: 3094–3100. doi:10.1093/bioinformatics/bty191
59. Bilska A, Kusio-Kobiałka M, Krawczyk PS, Gewartowska O, Tarkowski B, Kobyłecki K, et al. Immunoglobulin expression and the humoral immune response is regulated by the non-canonical poly(A) polymerase TENT5C. Nat Commun. 2020;11. doi:10.1038/s41467-020-15835-3
60. Evans EA, Kawli T, Tan MW. Pseudomonas aeruginosa suppresses host immunity by activating the DAF-2 insulin-like signaling pathway in Caenorhabditis elegans. PLoS Pathog. 2008;4. doi:10.1371/journal.ppat.1000175

Supplementary Materials

(Supplementary Figures S1-S5, Supplementary Tables S6-S8)



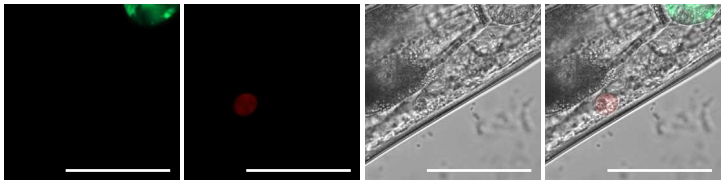
Supplementary Figure S1. NSPCs are exclusively expressed in the excretory gland cell.
(A) Additional fluorescence and differential interference contrast (DIC) microscopy images of NSPC localization in the excretory gland cell for different transcriptional and translational NSPC-mCherry reporters. Scale bars => 50 µm.

A

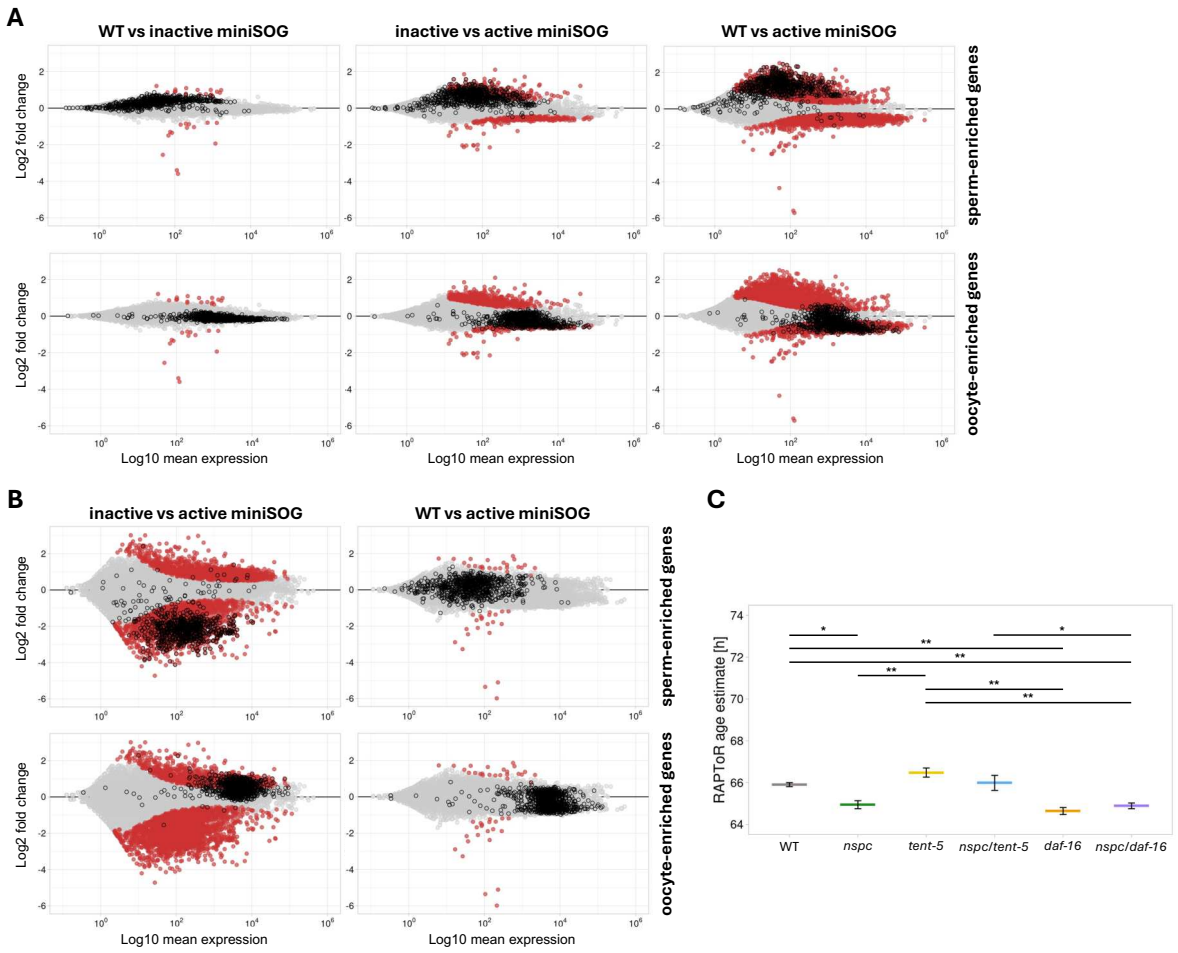
Blue light source	Mode of light	Worms' stage	Illumination time	Temperature	Ablation efficiency	Additional comments
stereo microscope fluorescence adapter (440-460 nm)	continuous light	multiple variants	multiple variants	20°C	depending on condition	- illumination of one plate at a time - uneven illumination of plate
military torches (460 nm and 470 nm)	continuous light					
WormLab's optogenetic module (470 nm)	flushing light					
LED board (460 nm)	continuous light; flushing light	L4	30 min; 10 min	20°C	medium	- better effect for continuous light - better effect for longer illumination
LED board (460 nm)	continuous light	L4	1 h	20°C	medium	- many dead worms
LED board (460 nm)	continuous light	L4	45 min	15°C	low	
LED board (460 nm)	continuous light	L2	30 min	20°C	low	
LED board (460 nm)	continuous light	embryo-L4	5 min every 2 h	RT	low	- worms' developmental delay
LED board (460 nm)	continuous light	L4	30 min (repeated next day); 2x15 min (repeated next day)	20°C	high	
LED board (460 nm)	**continuous light**	**L3**	**5x10 min**	**15°C**	**high**	- **complete disappearance of the excretory gland cell**

B

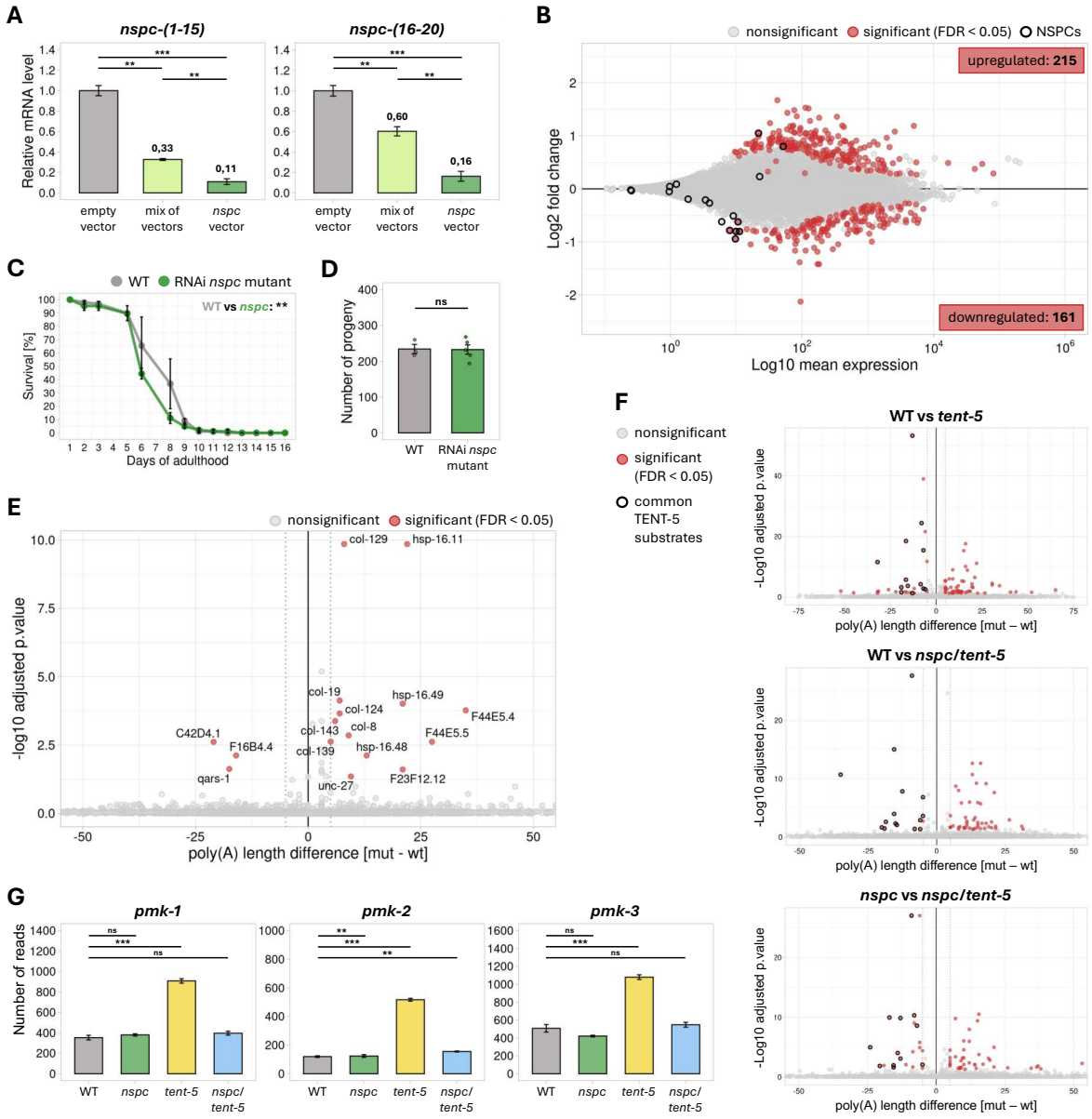
C



Supplementary Figure S2. Optogenetic ablation can be implemented to destroy the excretory gland cell.
(A) Table summarizing the optimization process for optogenetic ablation of the excretory gland cell. Various blue light sources, illumination durations, modes of light, and developmental stages of the worms were tested.
(B) LED advertising board used for illumination of miniSOG-expressing worms. The LED board emits 460 nm blue light with adjustable frequency, allowing simultaneous illumination of multiple plates containing worms.
(C) Fluorescence microscopy image of the excretory gland cell after optogenetic ablation. In some cases, the excretory gland cell was not destroyed completely but was significantly shrunken. However, such a drastic change in cell size and morphology is likely sufficient to impair its functions. Scale bars => 50 µm.

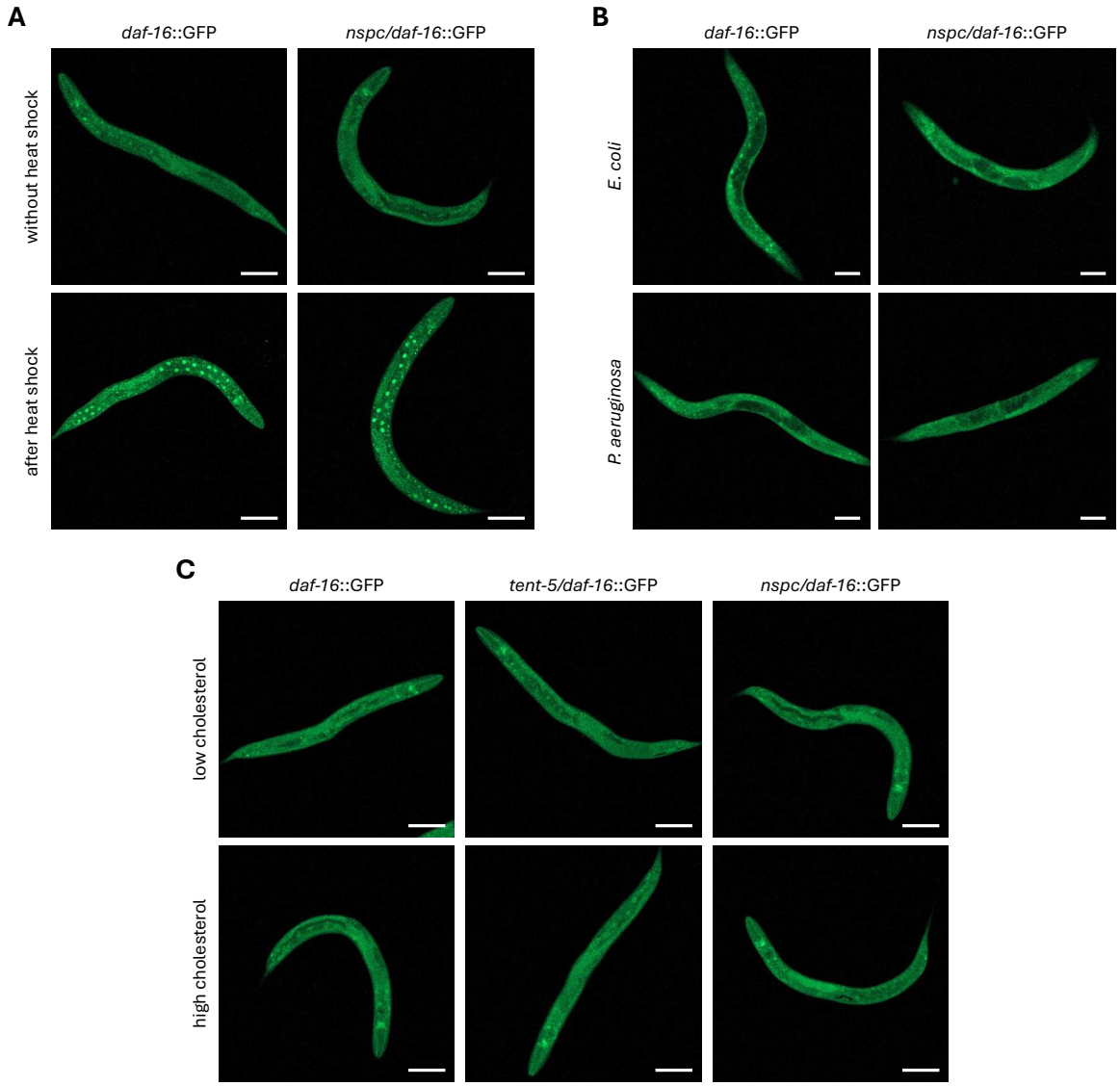


Supplementary Figure S3. Ablation of the excretory gland cell leads to no transcriptome changes.
(A) MA plots illustrating all possible combinations of differential gene expression between wild-type, inactive, and active miniSOG worms from the second replicate of the excretory gland cell ablation. Significantly changed genes (FDR < 0.05) are marked with red dots. Black borderline is used to mark sperm- and oocyte-enriched genes, as in Figure 3C.
(B) MA plots illustrating all remaining combinations of differential gene expression between wild-type, inactive, and active miniSOG worms from the third replicate of the excretory gland cell ablation. Significantly changed genes (FDR < 0.05) are marked with red dots. Black borderline is used to mark sperm- and oocyte-enriched genes, as in Figure 3C.
(C) Age estimates for wild-type (gray), *nspc* mutant (green), *tent-5* mutant (yellow), *nspc/tent-5* mutant (blue), *daf-16* mutant (orange), and *nspc/daf-16* mutant (purple) worms from other RNA-seq experiments described in this study. Although age differences are significant between multiple conditions, they are much smaller than those observed in RNA-seq samples after the ablation. Plots represent mean values with SD. Only significant comparisons are shown on the plots. ns => not significant; * => *p*-value < 0.05; ** => *p*-value < 0.01 (two-tailed *t*-test).



Supplementary Figure S4. NSPC deletion does not result in any physiological phenotypes.
(A) RT-qPCR showing decreased expression levels of all *nspc* genes following RNAi silencing. Worms were grown on *E. coli* HT115 expressing one of the following: an empty RNAi vector (control; gray), a mix of bacteria each expressing a vector targeting *nspc-7*, *-14*, or *-20* individually (light green), or a single silencing vector simultaneously targeting *nspc-7, -14,* and *-20* (*nspc-7/14/20*) (green). Relative *nspc* mRNA levels were normalized to *act-1*. Bar plots represent mean values with SD. ** => p-value < 0.01; *** => p-value < 0.001 (two-tailed t-test). Due to higher efficiency, the *nscp-7/14/20* vector was used for further RNAi silencing experiments.
(B) MA plot illustrating differential gene expression between wild-type and RNAi *nspc* mutant worms grown on *E. coli* HT115 expressing empty RNAi vector or *nspc-7/14/20* silencing vector, respectively. Significantly changed genes (FDR < 0.05) are marked with red dots (Supplementary Table S3). Black borderline is used to mark *nspc* genes. *Nspc-7*, *-14*, and *-20* were upregulated despite RNAi silencing due to the detection of bacterial dsRNA during RNA-seq.
(C) Survival of wild-type (gray) and RNAi *nspc* mutant (green) worms grown on *E. coli* HT115 expressing empty RNAi vector or *nspc-7/14/20* silencing vector, respectively. Points represent mean values from 3 separate plates (n≅40 each) with SD. ** => *p*-value < 0.01 (log-rank test).
(D) Differences in brood sizes between wild-type (gray) and RNAi *nspc* mutant (green) worms grown on *E. coli* HT115 expressing empty RNAi vector or the *nspc-7/14/20* silencing vector, respectively. Bar plots represent mean values with SD (n=3 for WT, n=5 for RNAi *nspc* mutant). ns => not significant (two-tailed t-test).

(E) Volcano plot showing differential polyadenylation between wild-type and *nspc* mutant worms (Supplementary Table S4). Transcripts with significantly changed median poly(A) tail length (FDR < 0.05) by a minimum of 5 nucleotides (dotted lines) are marked with red dots and labeled.
(F) Volcano plots showing a lack of NSPC influence on differential polyadenylation between wild-type and *tent-5* mutant worms (Supplementary Table S4). Transcripts with significantly changed median poly(A) tail length (FDR < 0.05) by a minimum of 5 nucleotides (dotted lines) are marked with red dots. Black borderline is added to transcripts commonly shortened in all three comparisons.
(G) Plots showing the number of RNA-seq reads for *pmk* genes in wild-type (gray), *nspc* mutant (green), *tent-5* mutant (yellow), and *nspc/tent-5* mutant (blue) worms. Bar plots represent mean values with SD. ns => not significant; ** => p-value < 0.01; *** => p-value < 0.001 (two-tailed t-test).



Supplementary Figure S5. Studying the influence of TENT-5 and NSPCs on DAF-16 localization.
(A) Fluorescence microscopy images of DAF-16-GFP localization in wild-type and *nspc* mutant worms after heat shock (1h; 35ºC), which is the best-known stimuli for relocating DAF-16 from the cytoplasm to the nucleus. Scale bars => 100 µm.
(B) Fluorescence microscopy images of DAF-16-GFP localization in wild-type and *nspc* mutant worms after heat shock (1h; 35ºC) followed by exposure to either non-pathogenic *E. coli* HB101 or pathogenic *P. aeruginosa* PAO1 as described in Evans E.A. *et al.* (2008) [60]. Scale bars => 100 µm.
(C) Fluorescence microscopy images of DAF-16-GFP localization in wild-type, *tent-5*, and *nspc* mutant worms grown on plates with low (0 mg/ml) or high (80 mg/ml) cholesterol concentration. Scale bars => 100 µm.

Supplementary Table S6. List of *C. elegans* strains used in this study.

Strain	Genotype	Source	Comments
N2 Bristol	wild type	CGC*	-
CF1038	*daf-16(mu86) I*	CGC*	-
OH16024	*daf-16(ot971[daf-16::GFP]) I*	CGC*	-
TM3504	*tent-5(tm3504) I*	NBRP**	-
ADZ31	*rttEx12[nspc-14p::mCherry::nspc-14 3′ UTR; myo-2p::gfp::unc-54 3' UTR] in N2*	this study	extrachromosomal array; transcriptional reporter
ADZ33	*rttEx13[nspc-20p::mCherry::nspc-20 3′ UTR; myo-2p::gfp::unc-54 3' UTR] in N2*	this study	extrachromosomal array; transcriptional reporter
ADZ35	*rttEx15[nspc-14p::nspc-14::mCherry::nspc-14 3′ UTR; myo-2p::gfp::unc-54 3' UTR] in N2*	this study	extrachromosomal array; translational reporter
ADZ37	*rttEx17[nspc-20p::nspc-20::mCherry::nspc-20 3′ UTR; myo-2p::gfp::unc-54 3' UTR] in N2*	this study	extrachromosomal array; translational reporter
ADZ44	*rttEx24[nspc-9p::nspc-9::mCherry::nspc-9 3′ UTR; myo-2p::gfp::unc-54 3' UTR] in N2*	this study	extrachromosomal array; translational reporter
ADZ45	*rttEx25[tent-5p::tent-5abd::egfp::3xflag::tent-5 3' UTR; nspc-14p::nspc-14::mCherry::nspc-14 3' UTR] in N2*	this study	extrachromosomal array
ADZ72	*nspc-10(rtt31[nspc-10p::tomm-20::miniSOG(426Cys)::nspc-10 3′ UTR]) X*	this study	CRISPR/Cas9
ADZ74	*rttEx12[nspc-14p::mCherry::nspc-14 3′ UTR; myo-2p::gfp::unc-54 3' UTR] in ADZ72*	this study	cross between ADZ72 and ADZ31; **final inactive miniSOG strain**
ADZ82	*nspc-10(rtt33[nspc-10p::tomm-20::miniSOG::nspc-10 3′ UTR]) X*	this study	CRISPR/Cas9
ADZ83	*rttEx28[nspc-14p::mCherry::nspc-14 3′ UTR; myo-2p::gfp::unc-54 3' UTR] in ADZ82*	this study	cross between ADZ82 and ADZ31; **final active miniSOG strain**
ADZ111	*nspc-(1-7)(rtt34) X*	this study	CRISPR/Cas9
ADZ112	*nspc-(1-10)(rtt35) X*	this study	CRISPR/Cas9
ADZ113	*nspc-(1-15)(rtt36) X*	this study	CRISPR/Cas9
ADZ114	*nspc-(1-20)(rtt37) X*	this study	CRISPR/Cas9; **final *nspc* mutant**
ADZ116	*tent-5(tm3504) I; daf-16(ot971[daf-16::GFP]) I*	this study	cross between TM3504 and OH16024
ADZ122	*nspc-(1-20)(rtt37) X; daf-16(ot971[daf-16::GFP]) I*	this study	cross between ADZ114 and OH16024
ADZ130	*nspc-(1-20)(rtt37) X; tent-5(tm3504) I*	this study	cross between ADZ114 and TM3504
ADZ131	*nspc-(1-20)(rtt37) X; daf-16(mu86) I*	this study	cross between ADZ114 and CF1038

***CGC** - Caenorhabditis Genetics Center

****NBPJ** - National Bioresource Project of Japan

Supplementary Table S7. List of plasmids used in this study.

ID	Sequence	Source
pCFJ104	*myo-3p::mCherry:unc-54 3′ UTR*	J. J. Ewbank
pRH269	*myo2p::gfp::unc-54 3′ UTR*	K. Drabikowski
pCZGY1703	*tomm-20N::miniSOG(426Cys)* in pENTR	Addgene #66781
L4440	-	K. Drabikowski
pCFJ151	-	J. J. Ewbank
pJET1.2	-	Thermo Fisher Scientific
WRM069A	BAC containing *tent-5abcd::egfp::3xflag*	TransgeneOme
pVL033	*tent-5pshort::tent-5abd::egfp::3xflag::tent-5 3' UTR* in pCFJ151	this study
pVL037	*nspc-14p::mCherry::nspc-14 3′ UTR* in pCFJ104	this study
pVL038	*nspc-14p::nspc-14::mCherry::nspc-14 3′ UTR* in pCFJ104	this study
pVL039	*nspc-20p::mCherry::nspc-20 3′ UTR* in pCFJ104	this study
pVL040	*nspc-20p::nspc-20::mCherry::nspc-20 3′UTR* in pCFJ104	this study
pVL073	*nspc-9p::nspc-9::mCherry::nspc-9 3′ UTR* in pCFJ104	this study
pCE090	*nspc-7* in L4440	this study
pCE091	*nspc-14* in L4440	this study
pCE092	*nspc-7::nspc-14::nspc-20* in L4440	this study

Supplementary Table S8. List of guide RNAs, repair templates and primers used in this study.

ID	Sequence	Comments
guide RNAs		
sgRNA01	GCUACCAUAGGCACCACGAG	*dpy-10*; for co-CRISPR
sgRNA08	AAAGGACAACGAGGGTGCGA	for ADZ72 generation
sgRNA09	AGCTTTACAACAATCTTTAA	
sgRNA17	TCTCTGGTCCCTGCAGAAAG	for ADZ82 generation
sgRNA18	TATAGATGTGATATAATGAC	for ADZ111 generation
sgRNA19	AGGAAGTGGATGTTATCTCC	
sgRNA20	ATTGTTCTAGATGCGCAAAC	for ADZ112 generation
sgRNA21	ACAGCTCCTGTTATCAATAC	
sgRNA22	TTGTGATCGCATTTTAAATG	for ADZ113 generation
sgRNA23	CATCACTTGAACAGTAATCC	
sgRNA24	TGTAACCAGGCCGTGTTTGA	for ADZ114 generation
sgRNA25	GTGAATACCACCCACCAGTT	
repair templates		
VL314	CACTTGAACTTCAATACGGCAAGATGAGAATGACTGGAA ACCGTACCGCATGCGGTGCCTATGGTAGCGGAGCTT CACATGGCTTCAGACCAACAGCCTAT	*dpy-10*; for co-CRISPR
-	C**ATG**GTAAGTAGTTTCAGTTTTAAATGGAACAATTGAAT ATCCTTGCAGTCGGACACAATTCTTGGTTTCAACAAATCA AACGTCGTTTTGGCTGCTGGAATTGCTGGAGCCGCTTTCC TCGGCTACTGCATTTACTTCGATCATAAGAGAATCAACGC TCCAGACTACAAGGACAAGATTAGGCAAAAGAGACGTGC CCAGGCTGGAGCATCCGGAGAGAAAAGTTTCGTGATAAC TGATCCACGGCTGCCAGACAATCCCATCATCTTCGCATCC GATGGCTTCCTGGAGCTGACCGAGTATTCCAGAGAGGAG ATCCTGGGCCGCAATTGCCGCTTTCTGCAGGGACCAGAG ACAGACCAGGCCACAGTGCAGAAGATTCGCGATGCCATT AGAGATCAGCGCGAGATTACCGTGCAGCTGATAAACTAC ACAAAAAGCGGGAAGAAATTCTGGAACCTCCTGCACCTC CAGCCCATGAGGGACCAGAAGGGTGAGCTCCAGTATTTC ATCGGAGTGCAGCTGGATGGAGGATCTTCCGGATAAAGG AGTTTTCTGCACATTCAGTTGTTGCACATTCATTC	for ADZ72 generation **ATG** – start codon XXX – *nspc-10* intron XXX – *tomm-20* XXX – *miniSOG(426Cys)* **T** – place of point mutation (G in active *miniSOG*)
prCE580	GTATTCCAGAGAGGAGATCCTGGGCCGCAATGGTCGCTT TCTGCAGGGACCAGAGACAGACC	for ADZ82 generation
prCE660	TAAACCAGATTTGTAAATTATAGATGTGATATAATTCCTG GTTTGCTGAAAATTATTTGCAATATGTGCC	for ADZ111 generation
prCE661	TCCCGCAAACAAAAATACATTGTTCTAGATGCGCATTGAT AACAGGAGCTGTAAAAATGGCCCTATGCCC	for ADZ112 generation
prCE662	AACAAATTTTACACATATTTGTGATCGCATTTTAATCCTG GAGTCTGAAACATCATTTTTTGGATTATAT	for ADZ113 generation
prCE663	TGCAGAAGTTGAACTCACTGTAACCAGGCCGTGTTGTTAG GAAATGTATTTTATTAAATATTGCGGACAA	for ADZ114 generation
primers		
prCE509	CATGGTAAGTAGTTTCAGTTTTAAATGGAACAATTGAATA TCCTTGCAGTCGGACACAATTCTTGGTTTCAACAAATC	for repair template for ADZ72 generation
prCE510	GAATGAATGTGCAACAACTGAATGTGCAGAAAACTCCTT TATCCGGAAGATCCTCCATCCAGC	
VL047	GATATCTGGATCCACGAAGC	for cloning of pVL033
VL048	CTGCAGGAATTCCTCGAGAC	
VL049	CGCACCGTACGTCTCGAGGAATTCCTGCAGGCATTCGAC GAGATGAAGAA	
VL050	ACCATGGGAAGCTTCGTGGATCCAGATATCCTATGTTCAG TGCATTACGTTTTTG	
VL182	ATAGCTTGGCGTAATCATGG	for cloning of pVL037, pVL038, pVL039, pVL040 and pVL073
VL183	ATAATTCACTGGCCGTCGTT	
VL183	TATGACCATGATTACGCCAAGCTATCATCTCATGATGGGC TACTTTC	

VL184	CCTTTGAGACCATGGTGTTGTATCAGATAAAACCTG	
VL185	CTGATACAACACCATGGTCTCAAAGGGTGAAGAAG	
VL186	AACATACACAAATCTACTTATACAATTCATCCATGCC	
VL187	ATTGTATAAGTAGATTTGTGTATGTTGCTTCAGTC	
VL188	TGTAAAACGACGGCCAGTGAATTATAACAGAAAAACTGTAGAATAGTTTTTTTC	
VL189	ATAGCTTGGCGTAATCATGG	
VL190	CCTTTGAGACCATAAGATTATTGTAGAGTTTGTTGCAAC	
VL191	CTACAATAATCTTATGGTCTCAAAGGGTGAAGAAG	
VL192	TATGACCATGATTACGCCAAGCTATAGCCCACGCTGATCTTCTTC	
VL193	CCTTTGAGACCATGTTTAGAATAGCCAGTTGTTGGAC	
VL194	GGCTATTCTAAACATGGTCTCAAAGGGTGAAGAAG	
VL195	ATCATATAGAGTACTACTTATACAATTCATCCATGCC	
VL196	ATTGTATAAGTAGTACTCTATATGATCAACTCGAATAAAAAC	
VL197	TGTAAAACGACGGCCAGTGAATTATGTTTAGACAAAAATTATGTACTGTAAAAG	
VL198	CCTTTGAGACCATATCGAGATTGAGGTAACGACG	
VL199	CCTCAATCTCGATATGGTCTCAAAGGGTGAAGAAG	
VL378	CTATGACCATGATTACGCCAAGCTATGATATTTTATGTGTCTCCATGATGATTATTTATAGT	
VL379	TCACCCTTTGAGACCATAAGATTATTGTAGAGTTTGTTGCAACCG	
VL380	TGCAACAAACTCTACAATAATCTTATGGTCTCAAAGGGTGAAGAAGA	
VL381	ACTGAAGCAACATATAGAACTCTACTTATACAATTCATCCATGCCACC	
VL382	GGATGAATTGTATAAGTAGAGTTCTATATGTTGCTTCAGTCGTTG	
VL383	AACGACGGCCAGTGAATTATAAGTACATGGTTCTCCCATGT	
prCE576	CTCGAGATTCAACTGATACCTCAACATG	for pCE090 cloning
prCE577	TCTAGATGATGAAAACCTTCATTATATC	
prCE578	CTCGAGTCTGATACAACACCATGTTC	for pCE091 cloning
prCE579	TCTAGAGGAGTGCAACGACTGAAGCAAC	
prCE588	GCTTCAGTCGTTGCACTCCTATTCAACTGATACCTCAAC	for pCE092 cloning
prCE589	GTTTAGAATAGCCAGTTGTTGTGATGAAAACCTTCATTATATC	
prCE590	GATATAATGAAGGTTTTCATCACAACAACTGGCTATTCTAAAC	
prCE591	GGGAGACCGGCAGATCTGATCGAGTTGATCATATAGAG	
VL001	TCAGGTTTCCACTGACAATG	*tent-5*; for genotyping
VL002	TGATCTCGACCTGATATTCC	
VL003	GTTCACACTCGTCCAACTC	
VL076	ACGACAAGACAGGCGGTATC	*daf-16*; for genotyping
VL077	TGAAGGGAGCCCATCAATGC	
VL078	GGCAATCTGAGGTGATGATG	
prCE676	GCATACTGTCGCTTCTTCATC	*daf-16::GFP*; for genotyping
prCE678	TATTGCTGCTTACCTCACTCCTC	
prCE511	GAATCTGATGTACGTGAATAGCGCAGTAG	in/active miniSOG; for genotyping
prCE512	GCTCGACTACAGTGCATAAGATTTCTCAC	in/active miniSOG, *nspc-(8-10)*; for genotyping
prCE596	GAGATCCTCGGCCGCAATGGT	active miniSOG; for genotyping
prCE597	GAGATCCTCGGCCGCAATTGC	inactive miniSOG; for genotyping

prCE649	AAAACAACGAGTCGCTTAACAAAT	*nspc-(1-7)*; for genotyping
prCE650	CTTCACCGGCCTTGAATTGC	
prCE609	GAGTTTTGGGGCGTAGGATGA	
prCE651	GACAACGTTACGACTCTGACC	*nspc-(8-10)*; for genotyping
prCE652	GTAATGGACTTTGCGCCCTCT	
prCE654	TGGCAAAAAGTCAGTTCAGGA	*nspc-(11-15)*; for genotyping
prCE655	AAATTCTGAAAACAAAGTTCTGTGC	
prCE657	GTAATATTTGTCTCGCTCAACTA	
prCE658	ACGTGTTTCCACGTATTTTGTAG	*nspc-(16-20)*; for genotyping
prCE659	TGAATTGAATGGTCATCCGGC	
prCE501	TCACCAGTTGCTCACTCATTACAG	
prCE574	ATGTTCCTTCGCACCCTCGTTG	*nspc-(1-15)*; for RT-qPCR
prCE575	CTGATCCATGGAGAGCAGAATG	
prCE573	ATGCTGGCTCTTCGCATCCTTG	*nspc-(16-20)*; for RT-qPCR
prCE547	CTTGACAGAGGTTCAAGGTTATGG	

Appendix 4

User manual for data analysis tools

User manual for data analysis tools

Introduction

This manual describes the functionality of data analysis tools designed for quicker and easier management of large-scale datasets obtained through RNA sequencing. The series of data analysis tools contains five R Shiny applications with following purposes:

1. ***RNA-seq data analysis***

   This tool provides a robust and clear visualization of gene expression or poly(A) tail length changes and allows to capture similarities and differences in transcriptomic signatures across multiple conditions for a single gene or group of genes.

2. ***Visualization of gene groups***

   This extension of the first tool allows to screen the differential expression or polyadenylation results for changes in characteristic gene groups, for example genes enriched in individual cells or associated to specific physiological processes.

3. ***qPCR analysis***

   Third tool uses the built-in script to analyze RT-qPCR results using $2^{-\Delta\Delta Ct}$ method using only an output file from the thermocycler.

4. ***Comparisons of gene groups***

   This tool identifies and visualize genes overlapping between two gene groups.

5. ***Translation of gene names***

   The last tool allows the quick translation between alternative gene identifiers – gene name, transcript name, and WormBase ID.

How to start

The applications can be accessed in two different ways:

1. Web Access via URL

The easiest method is to access the apps online through the URL provided below. However, in this format, the applications will load only with default datasets described in the PhD thesis titled **"Investigating cytoplasmic polyadenylation and its role in gene regulation and physiology in *Caenorhabditis elegans*"**.

URL: *http://zmackiewicz-rstudio.iimcb.gov.pl:3838/*

2. R Studio Access

For users who wish to customize the applications or analyze their own datasets, the raw code is available in my GitHub repository (link provided below). Each application is organized in its own folder. By downloading the folder and opening it in R Studio, users can run the corresponding R Shiny file locally and later modify the app according to their needs.

GitHub: *https://github.com/zuzanna-mackiewicz/PhD-data-analysis*

Detailed description of data analysis tools functionality

Tool nr 1 – RNA-seq data analysis

The first tool is designed to visualize datasets obtained using Illumina or Nanopore RNA sequencing and compare changes in gene expression levels or differences in poly(A) tail lengths across multiple conditions.

The tool for **RNA-seq data analysis** is depicted below and consists of the following:

(1) Experiment list for choosing the dataset that user wants to explore. The list of datasets is built into each R Shiny app and can be changed only in the application's code. The input should be structured as a table containing analyzed sequencing data, obligatorily with averaged normalized counts and log2 fold change of gene expression or poly(A) length difference between control and tested condition.

(2) Gene window for searching a singular gene or family of genes. As a result, all genes containing the typed-in phrase will show up on the **(5) MA plot**. Sometimes, a few gene families have the same phrase in their names, causing some unwanted genes to be marked. The user can limit the search by typing ^ symbol before or $ symbol after the desired name.

(3) Group of genes window for exploring not only genes from the same family but also genes interlinked, for example, by their function or site of expression. Gene names can be entered manually or copied into the window, preferably from an Excel sheet. Although, there is no limit to the number of genes that can by visualized simultaneously, the app works smoothly for up to 1000 genes. Both this window and the **(2) Gene** window are case-sensitive, and gene names need to be entered exactly as they appear in the *C. elegans* reference.

(4) Table representing all the statistics calculated for chosen comparison and can be sorted by any column. Additionally, by clicking on a gene symbol, the user can select genes to be visualized on the **(5) MA plot**.

(5) MA plot showing a differential expression or differential polyadenylation between conditions in the chosen experiment. The "x" axis corresponds to log10 mean expression level, and "y" to log2 fold change expression change. Significantly changed genes appear as red dots, and not significantly as gray. All selected or typed genes are marked with black borderlines and additional labels, as long as these labels do not disrupt plot's clarity.

(6) Selected gene summary presenting how the expression of a chosen gene changes across all experiments listed in the **(1) Experiment** list. This allows for a comprehensive analysis of a single gene without the need to jump between different plots.

(7) Clear selection button for restarting the gene search by emptying both the **(2) Gene** and the **(3) Group of genes** windows, and unselecting genes chosen by clicking in the **(4) Table**.

For clarity, the datasets from the PhD thesis titled **"Investigating cytoplasmic polyadenylation and its role in gene regulation and physiology in *Caenorhabditis elegans*"** – available at the provided URL and GitHub repository – have been organized into three separate applications, each focusing on different part of the research. The first app focuses on differential gene expression in worms after the excretory gland cell ablation, second app on differential expression in *nspc* mutant worms, and the third app on differential polyadenylation driven by various poly(A) polymerases.

Tool nr 2 – Visualization of gene groups

The second tool enables robust screening of differential expression or polyadenylation results for changes in specific gene groups, which could suggest potential functions of the studied tissue or protein. Its design is similar to the first tool and uses the same input files containing analyzed RNA sequencing data.

The tool for **Visualization of gene groups** is depicted below and consists of the following:

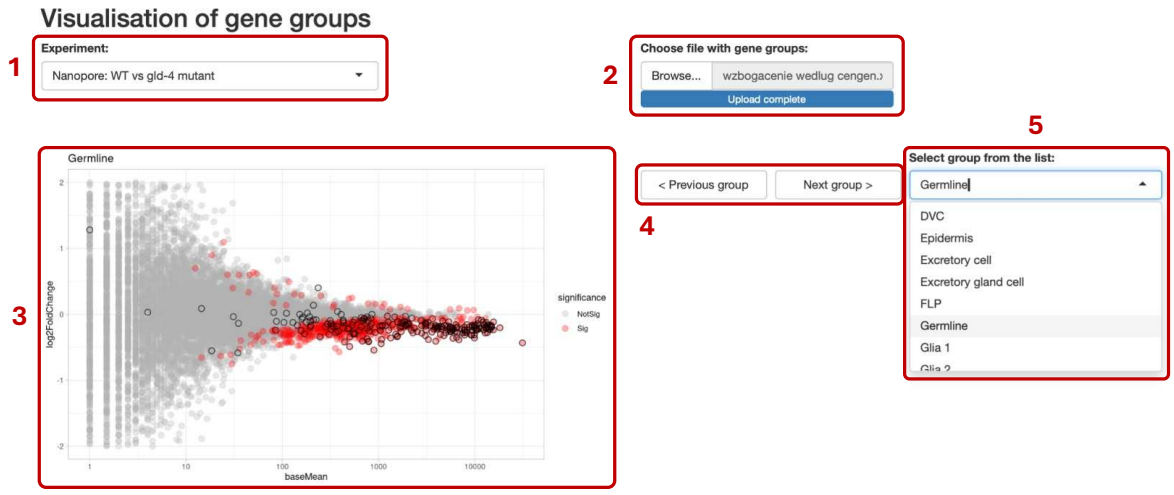
(1) Experiment list for choosing the conditions for which the user wants to perform a screening. In this version of the app, the user needs to select only one experiment for analysis at the beginning, as changing experiments during the screening significantly slow down the app. To change the experiment, the app must be restarted.

(2) Choose file with gene groups browser for uploading an Excel file with gene groups for screening. The table should provide groups of genes segregated into columns with headers defining each group. Groups can, for example, represent genes enriched in individual cells.

(3) MA plot showing differential expression or polyadenylation, similarly to the previous tool. Black borderlines are added for the genes included in the selected gene group.

(4) Previous and next group buttons for easy screening through all groups defined in the uploaded table.

(5) Select group from the list for visualizing genes for a specific group without the need to process through all the groups in the table.



Tool nr 3 – qPCR analysis

The third tool allows to automate the process of the RT-qPCR results analysis using $2^{-\Delta\Delta Ct}$ method. Currently, the tool is adapted to work with output files from the QuantStudio 5 thermocycler, but with minor modifications to the code, it can easily be adapted for use with data from other devices.

The tool for **qPCR analysis** is depicted below and consists of the following:

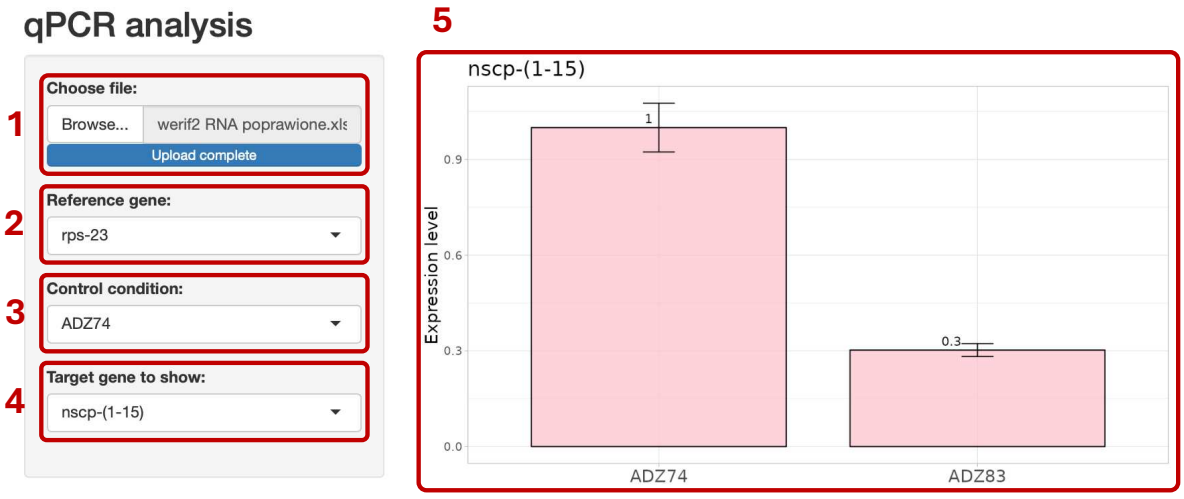
(1) Choose file browser for selecting the file containing the RT-qPCR results. After a successful upload, a **(5) Bar plot** will appear on the right side of the screen showing the analyzed data. For the file to work correctly, gene names need to be defined already in the QuantStudio software, and biological replicates must be defined as a number separated from the condition name with a space.

(2) Reference gene list for choosing which gene should be treated as the reference gene (ideally a housekeeping gene whose expression remains constant across conditions).

(3) Control condition list summarizing different conditions run on the RT-qPCR plate and allowing to select the internal control for analysis (ideally a sample from wild-type or untreated worms).

(4) Target gene to show list allowing to screen through the various target genes for which reactions were run together on the plate.

(5) Bar plot representing the final results, including the calculated mean Ct value with standard deviation error bars for each condition.



Tool nr 4 – Comparisons of gene groups

The fourth tool can be used to identify genes overlapping between two gene groups.

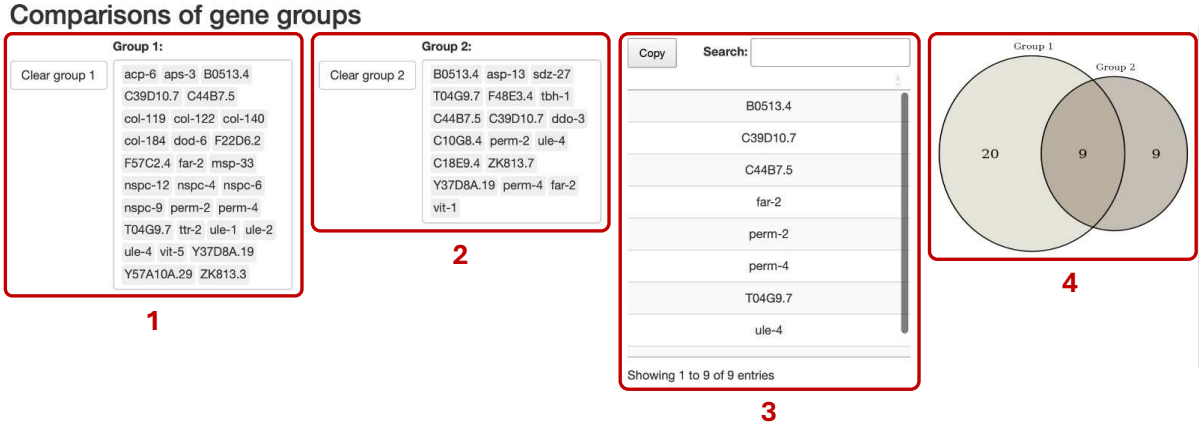
The tool for **Comparisons of gene groups** is depicted below and consists of the following:

(1) Group 1 window for entering the first group of genes. Gene names can be entered manually or copied into the window, preferably from an Excel sheet. It can be later cleared with the button on the left side.

(2) Group 2 window for entering the second group of genes. As long as both groups are entered consistently, the app accepts any type of gene identifiers. For optimal performance, it is recommended to use groups containing fewer than 1 000 genes.

(3) Table listing all mutual genes present in both groups, that can be copied for further analyses.

(4) Venn diagram providing a visual representation of the overlap.



Tool nr 5 – Translation of gene names

The fifth tool aims at simplifying the translation between three alternative gene identifiers: gene name, transcript name and WormBase ID.

The tool for **Translation of gene names** is depicted below and consists of the following:

(1) Genes to translate window for entering names of genes for translation. Gene names in any form can be entered manually or copied into the window, preferably from an Excel sheet.

(2) Translate into WormBase ID and **Translate into symbol** buttons for converting inserted genes into the respective format, **Whole translation** button for showing all three types of identifiers for each gene, and **Clear** button for restarting the analysis.

(3) Table displaying list of translated genes, which can be copied for subsequent analyses.