

Silencing of Repetitive DNA Is Controlled by a Member of an Unusual *Caenorhabditis elegans* Gene Family

Eduardo Leyva-Díaz,^{*,1} Nikolaos Stefanakis,^{*,2} Inés Carrera,^{*,3} Lori Glenwinkel,^{*} Guoqiang Wang,[†] Monica Driscoll,[†] and Oliver Hobert^{*,1}

^{*}Departments of Biological Sciences and Biochemistry and Molecular Biophysics, Howard Hughes Medical Institute, Columbia University, New York, New York 10027 and [†]Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854

ORCID IDs: 0000-0001-6750-9168 (E.L.-D.); 0000000010719203142 (N.S.); 0000-0002-7634-2854 (O.H.)

ABSTRACT Repetitive DNA sequences are subject to gene silencing in various animal species. Under specific circumstances repetitive DNA sequences can escape such silencing. For example, exogenously added, extrachromosomal DNA sequences that are stably inherited in multicopy repetitive arrays in the nematode *Caenorhabditis elegans* are frequently silenced in the germline, whereas such silencing often does not occur in the soma. This indicates that somatic cells might utilize factors that prevent repetitive DNA silencing. Indeed, such “antisilencing” factors have been revealed through genetic screens that identified mutant loci in which repetitive transgenic arrays are aberrantly silenced in the soma. We describe here a novel locus, *pals-22* (for protein containing ALS2CR12 signature), required to prevent silencing of repetitive transgenes in neurons and other somatic tissue types. *pals-22* deficiency also severely impacts animal vigor and confers phenotypes reminiscent of accelerated aging. We find that *pals-22* is a member of a large family of divergent genes (39 members), defined by homology to the ALS2CR12 protein family. While gene family members are highly divergent, they show striking patterns of chromosomal clustering. The family expansion appears *C. elegans*-specific and has not occurred to the same extent in other nematode species for which genome sequences are available. The transgene-silencing phenotype observed upon loss of *PALS-22* protein depends on the biogenesis of small RNAs. We speculate that the *pals* gene family may be part of a species-specific cellular defense mechanism.

KEYWORDS *Caenorhabditis elegans*; transgene silencing; RNA interference

Over half the human genome consists of repetitive DNA elements (Lander *et al.* 2001; de Koning *et al.* 2011). The view of the role of repetitive DNA has evolved over the last few decades from considering it as “junk DNA” to the recognition of repetitive DNA as being essential for genome function (Doolittle and Sapienza 1980; Orgel and Crick 1980; Lynch and Conery 2003; Shapiro and von Sternberg

2005). The main constituents of these repetitive DNA elements are retrotransposons, a large family of transposable elements capable of copying themselves and reinserting into the host genome (Kazazian 2004; Goodier and Kazazian 2008; Cordaux and Batzer 2009). Retrotransposons and other elements with the ability to copy themselves pose a threat to genome integrity due to the potential deleterious effects of landing in coding or regulatory regions (Friedli and Trono 2015). The activation of proto-oncogenes in some leukemias represent an example of such harmful consequences (Hacein-Bey-Abina *et al.* 2003). However, repetitive DNA elements have also been found to play beneficial roles in a number of processes, ranging from the regulation of gene expression to interaction with nuclear structures for genome packaging, to DNA repair and restructuring (Shapiro and von Sternberg 2005; Goke and Ng 2016). Hence, transposable elements have been postulated as a powerful genetic force involved in

Copyright © 2017 by the Genetics Society of America

doi: <https://doi.org/10.1534/genetics.117.300134>

Manuscript received June 1, 2017; accepted for publication August 10, 2017; published Early Online August 11, 2017.

Supplemental material is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.117.300134/-/DC1.

¹Corresponding authors: Department of Biological Sciences, Department of Biochemistry and Molecular Biophysics, Howard Hughes Medical Institute, Columbia University, 1212 Amsterdam Ave., New York, NY 10027. E-mail: eld2154@columbia.edu; and or38@columbia.edu

²Present address: Laboratory of Developmental Genetics, The Rockefeller University, New York, NY 10065

³Present address: Worm Biology Laboratory, Institut Pasteur de Montevideo, 11400 Montevideo, Uruguay

the evolution of organismal complexity (Kazazian 2004; Friedli and Trono 2015).

To balance the deleterious and beneficial features of repetitive sequences, organisms have evolved ways to finely tune the regulation of repetitive DNA elements (Schlesinger and Goff 2015; Chuong *et al.* 2017). For example, endogenous silencing mechanisms have evolved to prevent genome damage by the spread of mobile repetitive DNA elements. Silencing can be directed by sequence-specific transcription factors and by small RNAs. For example, retrotransposons are extensively recognized by the Krueppel-associated box-zinc finger (KRAB-ZFP) proteins (Rowe *et al.* 2010; Quenneville *et al.* 2012; Jacobs *et al.* 2014), which form a large family of repressive transcription factors. These transcription factors are among the fastest evolving group of genes in the human genome and their diversity facilitates their ability to recognize a large number of retrotransposons (Nowick *et al.* 2010). Sequence-specific binding to retrotransposons by KRAB-ZFP factors triggers a cascade leading to chromatin-based silencing mechanisms (Wolf and Goff 2009).

RNAs of retrotransposons that escape transcriptional silencing are targeted and destroyed by the small RNA pathways in the cytoplasm (Toth *et al.* 2016). RNA-based mechanisms represent the most ancient defense against the genomic spread of repetitive DNA elements (Friedli and Trono 2015). These mechanisms comprise the action of small RNA molecules, including small interfering RNAs (siRNAs), PIWI-interacting RNAs (piRNAs), and microRNAs (miRNAs), which guide repressor protein complexes to particular targets in a sequence-specific manner. In addition to intervening at the post-transcriptional level, small RNAs can also intervene at the transcriptional level by directing deposition of repressive histone marks and DNA methylation to copies of retrotransposons and other elements (Le Thomas *et al.* 2013). piRNAs and siRNAs can be produced from repetitive DNA elements, which they silence in return (Law and Jacobsen 2010). Thus, in many organisms, repeated sequences can be particularly sensitive to gene silencing.

Repetitive DNA elements are not only abundant in vertebrate genomes. Repetitive DNA elements also abound in model organisms with smaller genome sizes; repetitive DNA accounts for 34–57% of the total genome in *Drosophila melanogaster* (Celniker *et al.* 2002), and at least 17% of the *Caenorhabditis elegans* genome (Stein *et al.* 2003). Repetitive DNA elements can also be generated experimentally. DNA transformation techniques in *C. elegans* produce repetitive extrachromosomal DNA arrays (“simple” arrays) (Stinchcomb *et al.* 1985). Several studies have shown that expression of transgenes organized in these repetitive arrays is silenced both in somatic cells and in the germline through heterochromatin formation, involving several chromatin factors [reviewed in Cui and Han (2007)]. Somatic and especially germline transgene expression can be improved when the transgenic DNAs are cotransformed with an excess of carrier DNA, producing a less repetitive, more “complex” array (Kelly *et al.* 1997).

Importantly, gene expression from repetitive genomic regions can still be observed, suggesting that there are mechanisms that can antagonize silencing effects (Tseng *et al.* 2007). Multiple forward genetic screens in *C. elegans* have indeed identified factors that act to counter silencing of genes contained in repetitive sequences, based on screens for mutations that alter gene expression from transgenes present in tandemly repeated arrays (Hsieh *et al.* 1999; Grishok *et al.* 2005; Tseng *et al.* 2007; Fischer *et al.* 2013). In a classic study, mutations in *tam-1* (a RING finger/B-box factor) were found to reduce the expression of transgenes organized in simple but not complex repetitive arrays (Hsieh *et al.* 1999). Therefore, *tam-1* is an “antisilencing factor” that attenuates the context-dependent silencing mechanism affecting multicopy tandem-array transgenes in *C. elegans*. A subsequent study identified mutations in another gene important for expression of repetitive sequences, *lex-1*, which genetically interacts with *tam-1* (Tseng *et al.* 2007). *LEX-1* encodes a protein containing an ATPase domain and a bromodomain, both of which suggest that *LEX-1* associates with acetylated histones and modulates chromatin structure. Hence, *TAM-1* and *LEX-1* are antisilencing factors that function together to influence chromatin structure and to promote expression from repetitive sequences.

Further studies found that tandem-array transgenes become silenced in most mutants that cause enhanced exogenous RNA interference (RNAi) (Simmer *et al.* 2002; Kennedy *et al.* 2004; Fischer *et al.* 2013). Examples of gene inactivation known to cause increased transgene silencing and enhanced RNAi include the retinoblastoma-like gene *lin-35* (Hsieh *et al.* 1999; Wang *et al.* 2005; Lehner *et al.* 2006), the RNA-dependent RNA polymerase *rnf-3* (Simmer *et al.* 2002), and the helicase gene *eri-6/7* (Fischer *et al.* 2008). Some of the silencing acting on repetitive DNA elements (multicopy transgenes) depends on a complex interaction between different small RNA pathways (Fischer *et al.* 2013).

Here, we identify a novel locus, *pals-22* (protein containing ALS2CR12 signature), whose loss confers a transgene-silencing phenotype. *pals-22* mutants display context-dependent array silencing, affecting the expression of highly repetitive transgenes but not single-copy reporters. Animals lacking *pals-22* show locomotory defects and premature aging. *pals-22* is a member of a large family of divergent genes defined by homology to the ALS2CR12 protein family (Interpro, IPR026674; Panther, PTHR21707). The ALS2CR12 protein family is specifically expanded in *C. elegans*, and *pals* gene family members are clustered in the genome. We found that transgene silencing on *pals-22* mutants depends on a component of the RNAi pathway, indicating that *pals-22* might act as regulator of small RNA-dependent gene silencing.

Materials and Methods

Mutant strains

Strains were maintained by standard methods (Brenner 1974). The *C. elegans* mutant alleles used in this study were:

pals-22(ot723), *pals-22*(ot810), *pals-22*(ot811), *rde-4*(ne301) (Tabara *et al.* 1999), and *tam-1*(cc567) (Hsieh *et al.* 1999).

Reporter and transgenic strains

The *C. elegans* transgenic strains used in this study were *otIs381*[*ric-19*^{prom6}::NLS::gfp], *otIs380*[*ric-19*^{prom6}::NLS::gfp], *ccls4251*[*myo-3*^{prom}::gfp], *otIs251*[*cat-2*^{prom}::gfp], *otIs355*[*rab-3*^{prom1}::NLS::tagrfp], *otIs447*[*unc-3*^{prom}::mChOpti], *otEx6944*[*ric-4*^{prom26}::NLS::yfp], *otIs620*[*unc-11*^{prom8}::NLS::gfp], *otIs353*[*ric-4*^{fsmid}::SL2::NLS-YFP-H2B], *otIs534*[*cho-1*^{fsmid}::SL2::NLS-YFP-H2B], *otTi32*[*lin-4*^{prom}::yfp], *ieSi60*[*myo-2*^{prom}::TIR1::mRuby], *otEx7036*[*pals-22*^{prom}::gfp], *otEx7037*[*pals-22*::gfp], and *jsi37*[*pals-22*::gfp]. *pals-22*::gfp reporters (*otEx7036* and *otEx7037*) were generated using a PCR fusion approach (Hobert 2002). Genomic fragments were fused to the GFP coding sequence, which was followed by the *unc-54* 3'-UTR. See Supplemental Material, Table S1 in File S1 for transgenic strain names and microinjection details.

Forward genetic screens

Standard ethyl methanesulfonate (EMS) mutagenesis was performed on the fluorescent transgenic reporter strain, *otIs381*[*ric-19*^{prom6}::NLS::gfp], and ~60,000 haploid genomes were screened for expression defects with an automated screening procedure (Doitsidou *et al.* 2008) using the Union Biometrica COPAS FP-250 system. The mutant allele *pals-22*(ot723) was identified in an independent manual clonal screen for changes in reporter expression in neurons of *otIs381*[*ric-19*^{prom6}::NLS::gfp] after EMS mutagenesis. *ot723*, *ot810*, and *ot811* were the only three mutations derived from screens with *otIs381*[*ric-19*^{prom6}::NLS::gfp]. To identify the causal genes of the mutants obtained, we performed Hawaiian single-nucleotide polymorphism mapping and a whole-genome sequencing pipeline (Doitsidou *et al.* 2010; Minevich *et al.* 2012).

Microscopy

Worms were anesthetized using 100 mM sodium azide (NaN₃) and mounted on 5% agarose pads on glass slides. All images (except Figure 3, B, C, and E and Figure S1C in File S1) were acquired as Z-stacks of ~1 μm-thick slices with the Micro-Manager software (Edelstein *et al.* 2010) using the Zeiss Axio Imager.Z1 automated fluorescence microscope (Zeiss [Carl Zeiss], Thornwood, NY). Images were reconstructed via maximum intensity Z-projection of 2–10 μm Z-stacks using the ImageJ software (Schneider *et al.* 2012). Images shown in Figure 3, B, C, and E and Figure S1C in File S1 were acquired using a Zeiss confocal microscope (LSM880). Several Z-stack images (each ~0.4 μm thick) were acquired with the ZEN software. Representative images are shown following orthogonal projection of 2–10 μm z-stacks.

Single-molecule FISH (smFISH)

smFISH was done as previously described (Ji and van Oudenaarden 2012). Samples were incubated overnight

at 37° during the hybridization step. The *ric-19* and *gfp* probes were designed using the Stellaris RNA FISH probe designer and were obtained conjugated to Quasar 670 from Biosearch Technologies.

Fluorescence quantification

Synchronized day 1 adult worms were grown on NGM plates seeded with OP50 and incubated at 20°. The COPAS FP-250 system (Union Biometrica) was used to measure the fluorescence of 200–1000 worms for each strain. In Figures S2 and S3 in File S1, ImageJ software was used to measure the fluorescence intensity in the head of L4 worms (images obtained in the Zeiss Axio Imager.Z1 automated fluorescence microscope as described in Microscopy).

Bioinformatic analysis

The ALS2CR12 family protein phylogenetic tree was generated using MrBayes (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). Default setting for MrBayes were as follows: lset nst = 6, rates = invgamma, and ngen increased until the standard deviation of split frequencies was < 0.05. Input protein coding sequences for ALS2CR12 protein family proteins were aligned with M-Coffee (Wallace *et al.* 2006) using default settings. The MrBayes tree figure was rendered with FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>). The ALS2CR12 signature logo was generated using Skylign.org (Wheeler *et al.* 2014) from the PANTHER database (<http://www.pantherdb.org/>) (Mi *et al.* 2013, 2017) HMM profile PTHR21707.SF42.

RNAi by feeding

RNAi was performed as previously described with minor adaptations (Kamath and Ahringer 2003). L4-stage hermaphrodite worms were placed onto NGM plates containing seeded bacteria expressing dsRNA for each assayed gene. After 24 hr at 20°, adults were removed. After a further 36–40 hr at 20°, phenotypes were scored blindly.

Swim analysis

The swimming assay was performed using the CeleST program as previously described (Restif *et al.* 2014). In brief, we transferred five (day 1) adult hermaphrodites into 50 μl M9 buffer located in a 10-mm staggered ring on a glass slide. A 30-sec dark-field video (18 frames per second) was immediately recorded via StreamPix 7. Multiple features of the swim behavior were then analyzed using CeleST (Restif *et al.* 2014). Graphpad Prism 6 was used for data plotting and statistics.

Crawling assay

We washed ~30 day 1 adult hermaphrodites into M9 buffer (containing 0.2% BSA to prevent worms sticking to plastic tips) via low-speed centrifuging. We transferred these worms (in 20 μl volume) to a NGM agar plate (60 mm). After the liquid was completely absorbed and most animals were separated from each other, we started a 30-sec video recording

(20 frames per second). The video was processed on ImageJ and analyzed via wrMTrack plugin (Nussbaum-Krammer *et al.* 2015). The crawling paths were generated in ImageJ.

Age pigment assay

Age pigments of day 5 adult hermaphrodite (Gerstbrein *et al.* 2005) were captured via Zeiss LSM510 Meta Confocal Laser Scanning Microscope (excitation: Water cooled Argon laser at 364 nm; emission: 380–420 nm). The auto-fluorescence intensity was quantified in ImageJ.

Life span

Synchronized worms were picked at the L4 stage, and fed with OP50-1 bacteria on a 35 mm NGM agar plate (12 worms per plate, ~100 animals initiating each trial). Before the end of reproductive phase, animals were transferred into a new plate every 2 days to keep adults separated from progeny. Immobile animals without any response to touch were counted as dead; bagged worms were also counted as deaths; and animals crawling off the NGM agar were counted as lost and were excluded from analysis.

Data availability

Strains are available upon request. File S1 contains four supplemental figures, supplemental figure legends, and one supplemental table.

Results

pals-22 mutants show a transgene-silencing phenotype

Based on our long-standing interest in studying the regulation of pan-neuronal gene expression (Stefanakis *et al.* 2015), we sought to use genetic mutant screens to isolate factors that control the expression of pan-neuronally-expressed reporter transgenes. One screen that we undertook used a regulatory element from the pan-neuronally-expressed *ric-19* locus, fused to *gfp* (*otIs381[ric-19^{prom6}::NLS::gfp]*) (Stefanakis *et al.* 2015). We identified three independent mutant alleles, *ot723*, *ot810*, and *ot811*, in which expression of *ric-19^{prom6}::NLS::gfp* was reduced throughout the nervous system in all animals examined (Figure 1, A and B). Using our previously described whole-genome sequencing and mapping pipeline (Doitsidou *et al.* 2010; Minevich *et al.* 2012) we found that all three mutations affect the same locus, *C29F9.1* (Figure 2, A and B), which we named *pals-22* for reasons that we explain further below. The *ric-19^{prom6}::NLS::gfp* expression defect of *pals-22(ot811)* can be rescued by a fosmid (WRM0616DC09) encompassing the *pals-22* locus plus neighboring genes, as well as a genomic fragment that only contains the *pals-22* locus (791 bp upstream of the start codon to the stop codon and its 3'-UTR) (Figure 2C). Both *ot810* and *ot811* alleles carry early nonsense mutations and are therefore predicted to be null alleles (Figure 2B).

pals-22 mutants display reduced GFP expression of two separate *ric-19^{prom6}::NLS::gfp* integrated reporter transgenes (Table 1). However, smFISH against endogenous *ric-19*

transcripts failed to detect effects on the endogenous *ric-19* expression (Figure 1C). Since the two *ric-19* reporter transgenes that are affected by *pals-22* are repetitive, “simple” arrays, we considered the possibility that *pals-22* may encode a transgene-silencing activity. To test this notion, we examined the expression of a wide range of reporter transgenes in *pals-22*-deficient mutants (summarized in Table 1). Six additional simple arrays with widely different cellular specificities of expression (pan-neuronal, dopaminergic neurons, ventral cord motoneurons, and muscle) are also silenced in *pals-22* mutants (Figure 3 and Table 1). Two of these arrays, *myo-3::gfp* (*ccls4251*) and *cat-2^{prom}::gfp* (*otIs251*), were previously shown to be silenced by loss of *tam-1*, a “classic” transgene silencer mutation (Hsieh *et al.* 1999) (M. Doitsidou and O. Hobert, unpublished data) (Figure 3). We quantified the magnitude of the *pals-22(ot811)* effect on expression of simple array reporters by acquiring fluorescence intensity information from a synchronized worm population using a COPAS FP-250 system (Union Biometrica; “worm sorter”). At the L4 larval stage, we observed a 76% reduction in green fluorescence intensity for *ric-19^{prom6}::NLS::gfp*, 66% reduction for *myo-3::gfp*, 32% reduction for *cat-2^{prom}::gfp*, and 42% reduction in red fluorescence intensity for *rab-3^{prom1}::NLS::rfp* (Figure 1B and Figure 3, B, D, and E).

We also analyzed the expression of complex arrays (tandemly repeated transgenes with a less repetitive structure) and single-copy reporter transgenes in *pals-22* mutants (summarized in Table 1). One complex array transgene was silenced (*unc-11^{P8}::gfp*, Table 1) but others were not (*ric-4^{fosmid}::yfp* and *cho-1^{fosmid}::yfp*, Table 1 and Figure S1A in File S1). Perhaps the much smaller reporter fragment *unc-11^{P8}::gfp* generated more repetitive structures even in the context of “complex” arrays than the larger fosmid-based reporters. Two single-copy insertions and one endogenously-tagged gene were not affected by mutations in *pals-22* (Figure S1, B and C in File S1).

Similar to previously characterized transgene-silencing mutants, such as *tam-1* (Hsieh *et al.* 1999), the reduction of reporter expression is temperature sensitive. However, the direction of the sensitivity is inverted as compared to the *tam-1* case: the decrease in *ric-19^{prom6}::NLS::gfp* expression is most pronounced at 15°, while the effect is milder at 25° (Figure S2 in File S1). We also found stage-dependent variability: the decrease in transgene reporter expression is stronger as the animals develop, from mild differences in expression in early stages of development to more obvious defects at later stages (Figure S3 in File S1).

PALS-22 is a broadly expressed, cytoplasmically located protein

To analyze the expression pattern of *pals-22*, we fused the entire locus (including 791 bp of all intergenic 5' sequences and all exons and introns) to *gfp* (Figure 4A). This reporter construct fully rescues the transgene-silencing phenotype of *pals-22(ot811)* mutants (Figure 4B). The *pals-22* gene is expressed in a broad set of tissues: nervous system (pan-neuronal),

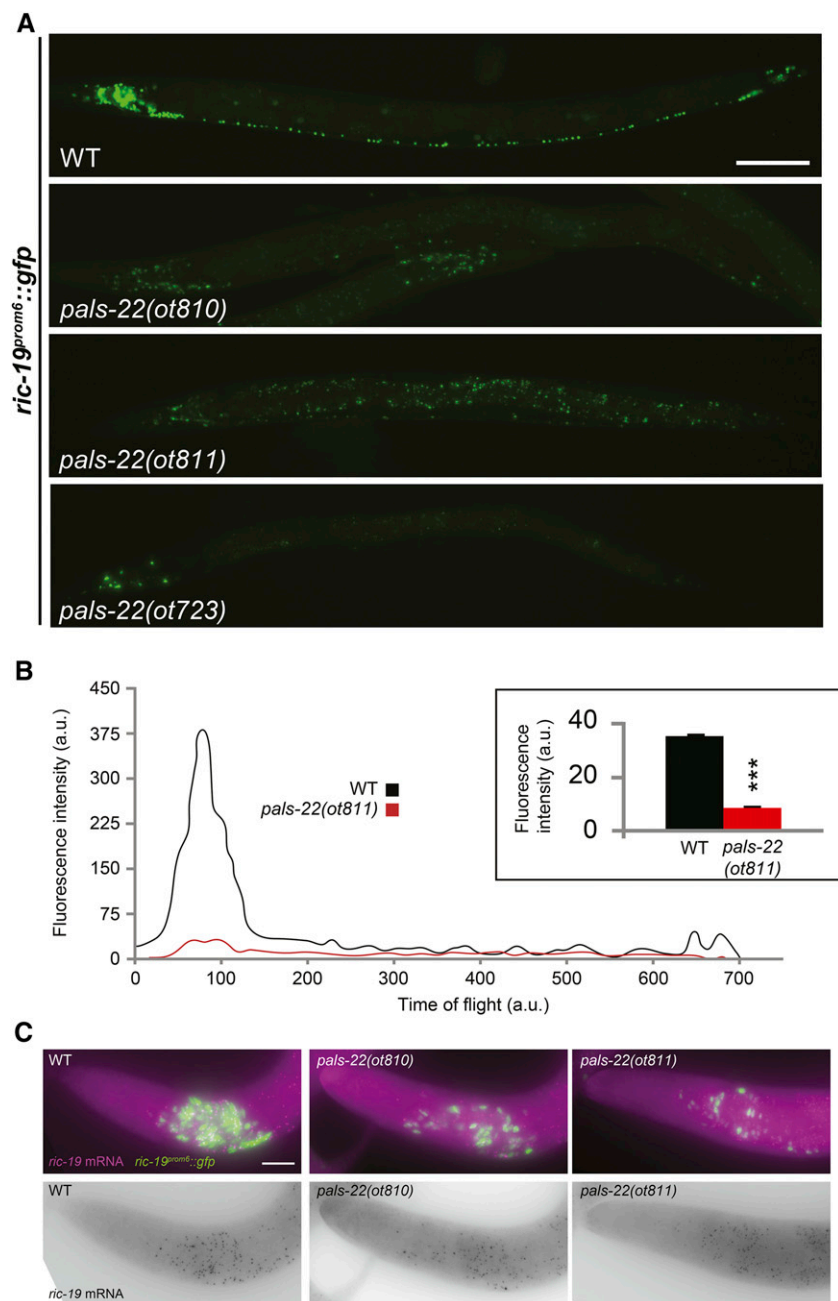


Figure 1 Loss of pan-neuronal reporter gene expression in *pals-22* mutants. (A) The *ric-19^{prom6}::NLS::gfp* transcriptional reporter is brightly expressed in all neurons in wild-type (WT) N2 worms, with a nuclear localization. In *pals-22(ot723)*, *pals-22(ot810)*, and *pals-22(ot811)* mutants, expression is reduced throughout the nervous system. All images correspond to L4 worms. Bar, 50 μ m. (B) Fluorescence profiles of a representative individual day 1 adult worm expressing *ric-19^{prom6}::NLS::gfp* in WT (black line) or *pals-22(ot811)* mutant background (red line) obtained with a COPAS FP-250 system. Time of flight indicates worm length, with lower values corresponding to the head of the worms. a.u., arbitrary units. Inset bar graph displays quantification of total fluorescence intensity averaged over 500 animals analyzed by COPAS. The data are presented as mean + SEM. Unpaired *t*-tests were performed for *pals-22(ot811)* compared to WT; *** *P* < 0.001. (C) WT (left), *pals-22(ot810)* (middle), and *pals-22(ot811)* (right) images of the anterior part of L3 worms showing equal *ric-19* mRNA levels in control and mutant worms as assessed by single-molecule fluorescence *in situ* hybridization. Individual transcripts shown as purple dots in top and as black dots in bottom panels. GFP expression of the reporter transgene is shown in green. At least 20 animals examined for each genotype displayed indistinguishable staining. Bar, 10 μ m.

body wall and pharyngeal muscle, gut, and hypodermal cells (Figure 4C). A single copy, chromosomal integrant of this translational reporter, generated by MosSCI (kindly provided by E. Troemel), shows a similar expression pattern (Figure 4D). Both the rescuing, translational multicopy reporter transgene, as well as the single-copy integrant, revealed a strong if not exclusive enrichment in the cytoplasm and appears to be excluded from the nucleus in most tissues (Figure 4, C and D). However, we cannot exclude the possibility that the nucleus contains a low amount of functional *PALS-22* protein. The *cis*-regulatory information that controls broad *pals-22* gene expression appears to be located in the 5' intergenic regions, since a transcriptional reporter that only contains 791 bp of the 5'

intergenic region of *pals-22*, fused to GFP, displays the same broad tissue distribution as the translational reporter (Figure 4E).

pals-22 is a member of an unusual *C. elegans* gene family

Analysis of the *PALS-22* protein sequence revealed that it is a member of the previously completely uncharacterized ALS2CR12 protein family. This family is defined by a sequence signature in the InterPro (<https://www.ebi.ac.uk/interpro/>; family IPR026674) (Finn *et al.* 2017) and Panther (<http://www.pantherdb.org/>; family PTHR21707) (Mi *et al.* 2013, 2017) databases, hence the name “PALS” for protein containing *ALS2CR12* signature. The *ALS2CR12*

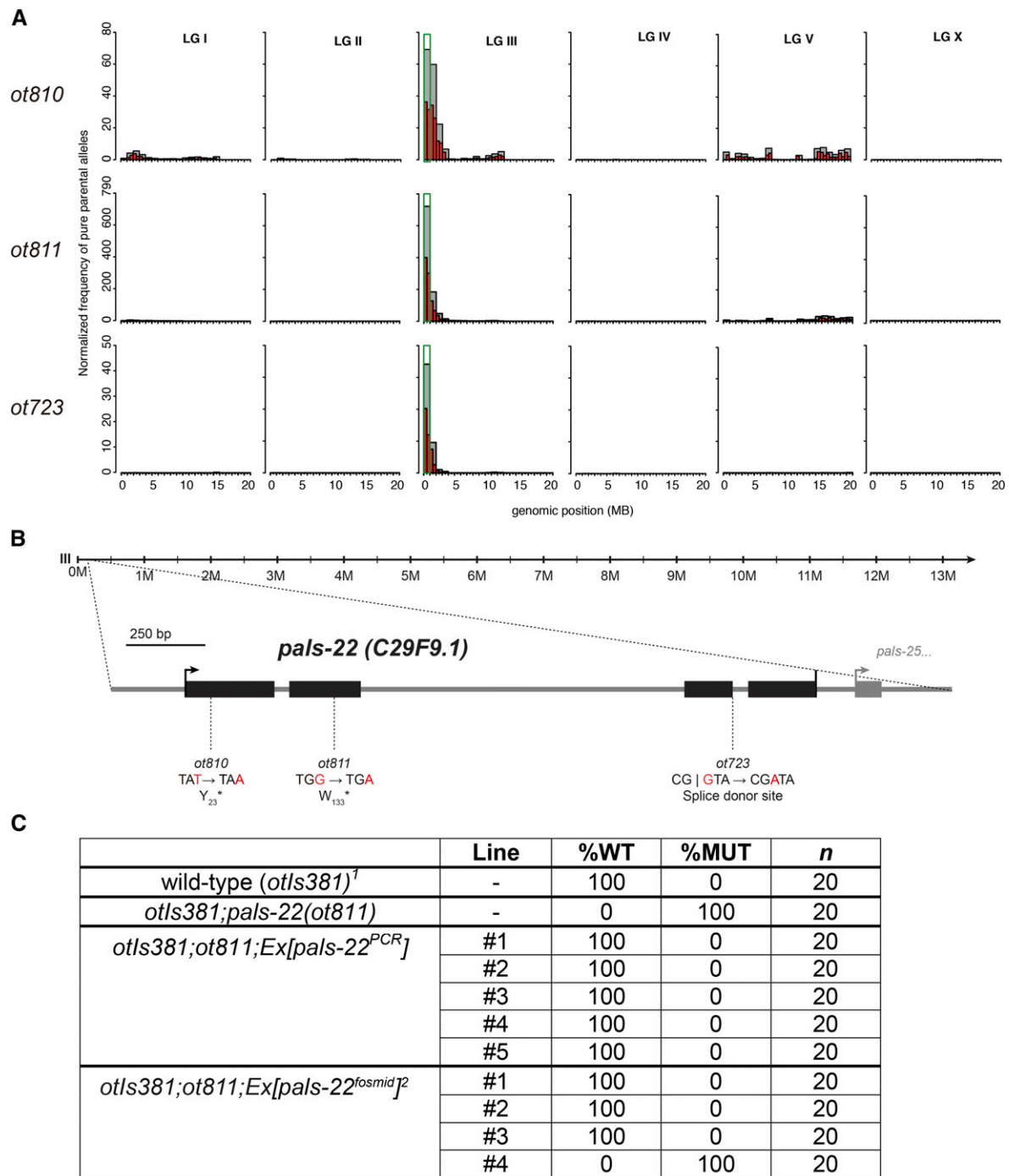


Figure 2 *pals-22* codes for a protein with an ALS2CR12 signature. (A) Single-nucleotide polymorphisms from Hawaiian wild *C. elegans* strain CB4856 mapping plots obtained from whole-genome sequencing of the following mutant alleles: *pals-22(ot810)* (top row), *pals-22(ot811)* (middle row), and *pals-22(ot723)* (bottom row). (B) Schematic of *pals-22* gene locus with mutant allele annotation. (C) *pals-22* rescue data. %WT = % percent animals that express the *ric-19* reporter strongly (as in wild-type animals). %MUT = % percent animals that express the *ric-19* reporter more weakly than wild-type.¹ *otls381* = *ric-19^{prom6}::NLS::gfp*. ² WRM0616DC09 fosmid.

sequence signature is comprised of a 404 amino acid position weight matrix modeled from an alignment of the human ALS2CR12 protein (Hadano *et al.* 2001) and 27 homologs (Figure S4A in File S1). No biochemical or cellular function has yet been assigned to the human ALS2CR12 protein or any of its homologs in other organisms.

Curiously, while there is only one member of this protein family in mouse and human, the number of ALS2CR12 protein family members is expanded to a total of 39 distinct proteins in *C. elegans* (Figure 5A and Table 2). In striking contrast, *Drosophila* seems to be completely devoid of ALS2CR12 family proteins. The 39 *C. elegans* PALS proteins

Table 1 *pals-22* effects on transgene reporter expression

Array Type	Construct	Expression Pattern	Relative Intensities of Reporters	
			Wild-Type	<i>pals-22</i> ^a
Simple array	<i>ric-19^{prom6}::gfp</i> ^b	Pan-neuronal	+++	+
	<i>myo-3^p::gfp</i>	Body-wall muscle	++++	+
	<i>cat-2^{prom}::gfp</i>	Dopaminergic neurons	++++	++
	<i>rab-3^{prom1}::rfp</i>	Pan-neuronal	+++	+
	<i>unc-3^p::mCherry</i>	Cholinergic neurons	++++	++
	<i>tdc-1^p::yfp</i>	RIC and RIM neurons	++	+
	<i>ric-4^{p26}::yfp</i> ^c	Neuronal	++++	+
	<i>unc-11^{p8}::gfp</i>	Pan-neuronal	+++	+
Complex array	<i>cho-1^{fsmid}::yfp</i>	Cholinergic neurons	++	++
	<i>ric-4^{fsmid}::yfp</i>	Pan-neuronal	++	++
	<i>lin-4^p::yfp</i>	Ubiquitous	++	++
Single copy	<i>myo-3^p::mRuby</i>	Pharyngeal muscle	++	++
None (endogenous tag)	<i>che-1::gfp</i>	ASE neurons	+	+

Transgene expression in wild-type and *pals-22* mutants. See Table S1 in File S1 for details on transgenic arrays. Number of plus signs (+) indicates the relative intensity of GFP fluorescence. At least 50 animals examined for each genotype. Unless otherwise indicated, all simple and complex arrays correspond to stable genome integrated transgenes. Single-copy reporters were generated by miniMos. GFP tagging of the *che-1* locus was achieved using clustered regularly interspersed short palindromic repeats/Cas9 technology.

^a *pals-22(ot811)* mutant background.

^b *otIs381* and *otIs380* strains (independently integrated lines).

^c Extrachromosomal array.

are very divergent from one another, as reflected in Table 2, with similarity scores of *PALS-22* compared to other PALS proteins and reflected in Figure S4B in File S1 with a sequence alignment of *PALS-22* with its closest paralogues. Besides poor paralogy, there is also poor orthology. For example, BLAST searches with *PALS-22* picks up no sequence ortholog in *C. briggsae*, and the best homolog from another species (*C. brenneri*, *CBN22612*) is less similar to *PALS-22* than some of the *C. elegans* *PALS-22* paralogues.

The expansion of *C. elegans* ALS2CR12 signature-containing proteins appears to be nematode species-specific, as *C. briggsae* only contains eight predicted proteins with the ALS2CR12 sequence signature; other nematodes also contain significantly less ALS2CR12 family proteins (Figure 5A). Such *C. elegans*-specific gene expansion is highly unusual, as shown in Figure 5B. Among 3874 Panther protein families analyzed, 2759 (71.2%) are present in the same number of genes in both *C. elegans* and *C. briggsae*, 128 (3.3%) protein families are present only in *C. elegans*, while 74 (1.9%) are present only in *C. briggsae*. Most of the remaining families are only slightly enriched in one species or the other. Just 10 protein families (0.3%) are enriched five times or more in *C. elegans* vs. *C. briggsae* (Figure 5B). Most of these families contain uncharacterized genes, even though many of them contain human and vertebrate orthologs. Among them, the ALS2CR12 family has not only a noteworthy enrichment, but it also constitutes the family with the largest absolute number of genes (Figure 5B).

As perhaps expected from a species-specific expansion, the *C. elegans* *pals* genes are genomically clustered (Figure 6). Fourteen genes are clustered in chromosome I (position I: 17.22 cM); while three clusters with seven genes (III: −21.90 cM), four genes (III: −26.98 cM), and three genes (III: −3.18 cM) are present in chromosome III; and

lastly, four genes are clustered in chromosome V (V: 4.22) (Figure 6 and Table 2). Taking into account the *C. elegans*-specific family expansion, it is not surprising to find a total lack of conservation in the regions encompassing most of the *pals* clusters (Figure 6). Only one cluster on chromosome V shows some degree of conservation among other nematode species. Genes surrounding these regions are conserved, suggesting recent gene duplications in the nonconserved areas. The low conservation region in cluster III: −21.90 cM, contains many additional, nonconserved genes that are largely expanded in a nematode-specific manner, namely the previously analyzed *fbxa* genes (Thomas 2006).

Local gene duplications seem a plausible mechanism for the origin of the expanded *C. elegans* *pals* gene family. Consequently, we reasoned that *pals* genes within the same cluster should be more similar among each other than to other *pals* genes. To explore this possibility, we built a phylogenetic tree to visualize the phylogenetic relationships between *pals* genes (Figure 5C). We included all *C. elegans* *pals* genes plus orthologs from *C. briggsae*, *C. remanei*, *C. brenneri*, mouse, and human (based on presence of InterPro IPR026674). As expected, *pals* genes within the same cluster have a closer phylogenetic relationship, suggesting a shared origin.

According to modENCODE expression data (Celniker *et al.* 2009), most of the genes within each cluster are expressed in the same stage (e.g., all genes clustered in chromosome I are only found in L4 males), suggesting related functions (Table 2). Perhaps most intriguingly, though, out of just a few hundred significantly enriched genes, the majority of *pals* genes become upregulated upon exposure to specific pathogens, specifically the exposure to intracellular fungal pathogen (microsporidia) or by viral infection (Bakowski *et al.* 2014; Chen *et al.* 2017). Induction of *pals* gene expression is also observed upon various other environmental insults

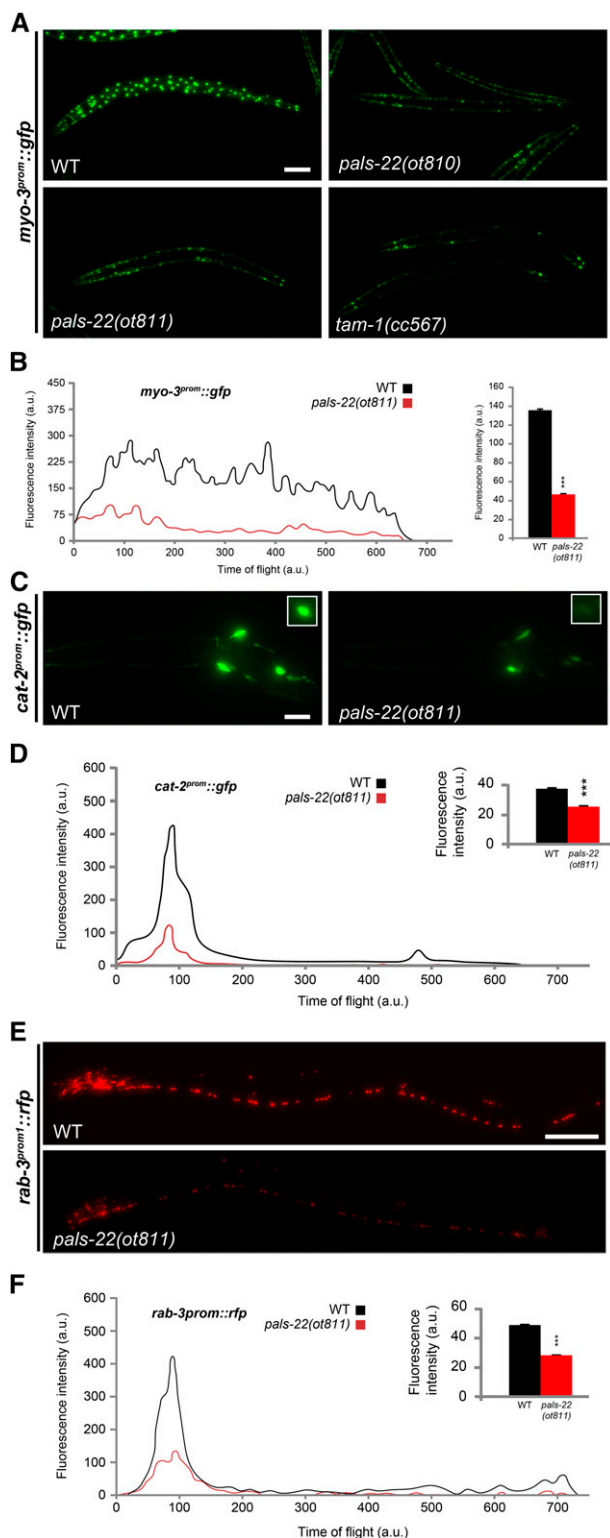


Figure 3 *pals-22* mutants show silencing of several multicopy arrays. (A) The *myo-3^{prom}::gfp* transcriptional reporter is brightly expressed in all body muscles in wild-type (WT) N2 worms, with a combination of mitochondrial and nuclear localization. In *pals-22(ot810)*, *pals-22(ot811)*, and *tam-1(cc567)* mutants, a generalized reduction in GFP fluorescence is observed. All images correspond to L4 worms. Bar, 50 μ m. (B) Fluorescence profile of a representative individual day 1 adult worm expressing *myo-3^{prom}::gfp* in WT (black line) or *pals-22(ot811)* mutant background

(exposure to toxic compounds) (Cui *et al.* 2007) (summarized in Table 2). Several *fbxa* genes present in the low conservation region in cluster III: -21.90 cM, also become upregulated upon exposure to microsporidia or by viral infection (Bakowski *et al.* 2014; Chen *et al.* 2017).

Somatic transgene silencing in *pals-22* mutants requires *rde-4*-dependent small RNAs

We examined whether two *pals-22* paralogs, *pals-19* and *pals-25*, may also be involved in transgene silencing. Both genes show significant sequence similarity to *pals-22* (Table 2), and one (*pals-25*) is directly adjacent to *pals-22* (Figure 6). We tested whether two nonsense alleles generated by the Million Mutant Project, *pals-19(gk166606)* and *pals-25(gk891046)* (Thompson *et al.* 2013), silence the *ric-19^{prom}::NLS::gfp* and *myo-3::gfp* multicopy transgenes. Neither array shows obvious changes in expression in *pals-19* or *pals-25* mutant backgrounds (data not shown).

To further pursue the role of *pals-22* in transgene silencing, we considered previous reports on transgene silencer mutations. Several genes with transgene-silencing effects are known to be involved in modifying chromatin (Cui and Han 2007). However, the cytoplasmic enrichment of *PALS-22* described above argues against a direct role in controlling chromatin architecture (although, a function in the nucleus cannot be ruled out). Nevertheless, we do find that *pals-22* affects transgene silencing at the mRNA level. smFISH against *gfp* mRNA shows that silenced *gfp* transgenic arrays display a significantly reduced number of transcripts in *pals-22* mutants (Figure 7A). Reduced transcript levels can be

(red line) obtained with a COPAS FP-250 system. Time of flight indicates worm length, with lower values corresponding to the head of the worms. a.u., arbitrary units. Inset bar graph displays quantification of total fluorescence intensity averaged over 1000 animals analyzed by COPAS. The data are presented as mean + SEM. Unpaired *t*-tests were performed for *pals-22(ot811)* compared to WT; *** $P < 0.001$. (C) GFP images showing silencing of *cat-2^{prom}::gfp* expression in the head of *pals-22(ot811)* mutants (right) compared to WT L4 worms (left). The *cat-2^{prom}::gfp* transcriptional reporter is expressed in all dopamine neurons: CEPD, CEPV, and ADE in the head, and PDE in the posterior midbody (top right insets). Bar, 10 μ m. (D) Fluorescence profile of a representative individual day 1 adult worm expressing *cat-2^{prom}::gfp* in WT (black line) or *pals-22(ot811)* mutant background (red line) obtained with a COPAS FP-250 system. Time of flight indicates worm length, with lower values corresponding to the head of the worms. a.u., arbitrary units. Inset bar graph displays quantification of total fluorescence intensity averaged over 1000 animals analyzed by COPAS. The data are presented as mean + SEM. Unpaired *t*-tests were performed for *pals-22(ot811)* compared to WT; *** $P < 0.001$. (E) Red fluorescent protein images showing silencing of *rab-3^{prom}::rfp* expression in the head of *pals-22(ot811)* mutants (right) compared to WT L4 worms (left). (F) Fluorescence profile of a representative individual day 1 adult worm expressing *rab-3^{prom}::rfp* in WT (black line) or *pals-22(ot811)* mutant background (red line) obtained with a COPAS FP-250 system. Time of flight indicates worm length, with lower values corresponding to the head of the worms. a.u., arbitrary units. Inset bar graph displays quantification of total fluorescence intensity averaged over 1000 animals analyzed by COPAS. The data are presented as mean + SEM. Unpaired *t*-tests were performed for *pals-22(ot811)* compared to WT; *** $P < 0.001$.

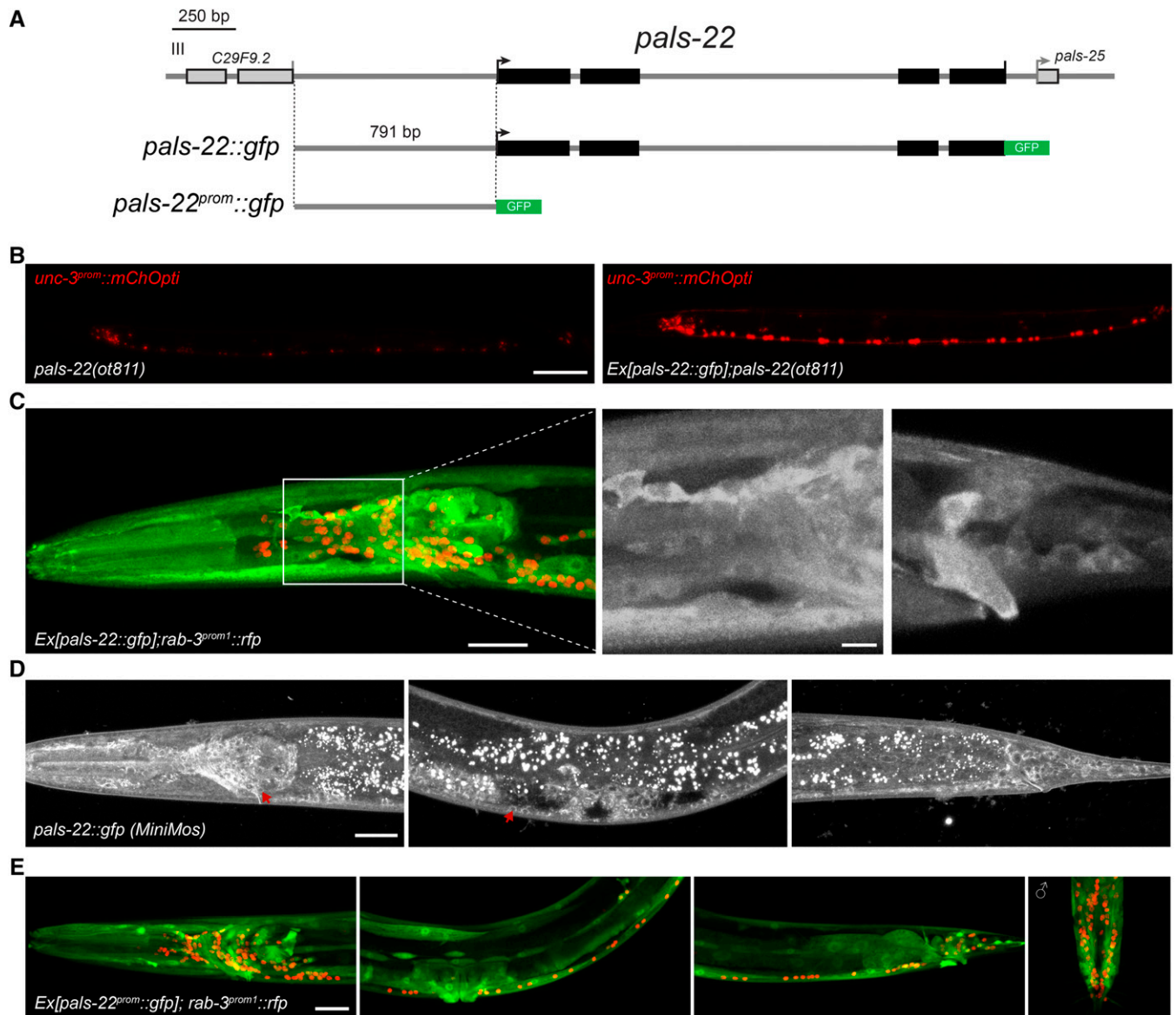


Figure 4 PALS-22 is a broadly expressed, cytoplasmically enriched protein. (A) Schematic of *pals-22* transcriptional and translational GFP reporters. The *gfp* reporter is followed by the 3'-UTR from the *unc-54* gene. (B) PALS-22::GFP (right) rescues the expression of *unc-3^{prom}::mChOpti* in *pals-22(ot811)* mutants (left). At least 50 animals were examined for each genotype. (C) Expression of the *pals-22::gfp* translational reporter (*otEx7037*) is shown in green in the head (left) in a representative hermaphrodite L4 worm. High-magnification images in black and white for the head (middle) and tail (right) show PALS-22 cytoplasmic enrichment. Nuclear pan-neuronal *rab-3^{prom1}::rfp* expression is shown in red on the background. Two distinct extrachromosomal arrays show the same pattern. (D) Expression of the *pals-22::gfp* MosSCI insertion (*jySi37*) is shown in black and white in the head (left panel), midbody (second panel), and tail (third panel) in a representative hermaphrodite L4 worm. Red arrows point to individual, isolated neuron cell bodies where cytoplasmic localization is particularly evident. (E) Expression of the *pals-22^{prom}::gfp* transcriptional reporter (*otEx7036*) is shown in green in the head (left panel), midbody (second panel), and tail (third panel) in a representative hermaphrodite L4 worm; and in the male tail (right panel). Nuclear pan-neuronal *rab-3^{prom1}::rfp* expression is shown in red on the background. Three distinct extrachromosomal arrays show the same pattern. Bar, 20 μ m (B, C, and E), 5 μ m ((B), high-magnification images), and 50 μ m (D).

explained by transcriptional and/or post-transcriptional alterations and we therefore considered the possibility that the effects of cytoplasmically enriched PALS-22 protein on transcript levels may be controlled by intermediary factors. Small interfering RNAs can affect gene expression both at the transcriptional or post-transcriptional level (Zamore *et al.* 2000; Elbashir *et al.* 2001; Le Thomas *et al.* 2013),

and through a genetic epistasis test, we asked whether *pals-22* requires small RNAs for its function. To this end, we turned to *rde-4* mutant animals. The dsRNA-binding protein RDE-4 initiates gene silencing by recruiting an endonuclease to process long dsRNA into short dsRNA and is involved in exogenous as well as endogenous RNAi pathways (Tabara *et al.* 2002; Parker *et al.* 2006; Gent *et al.* 2010; Vasale *et al.*

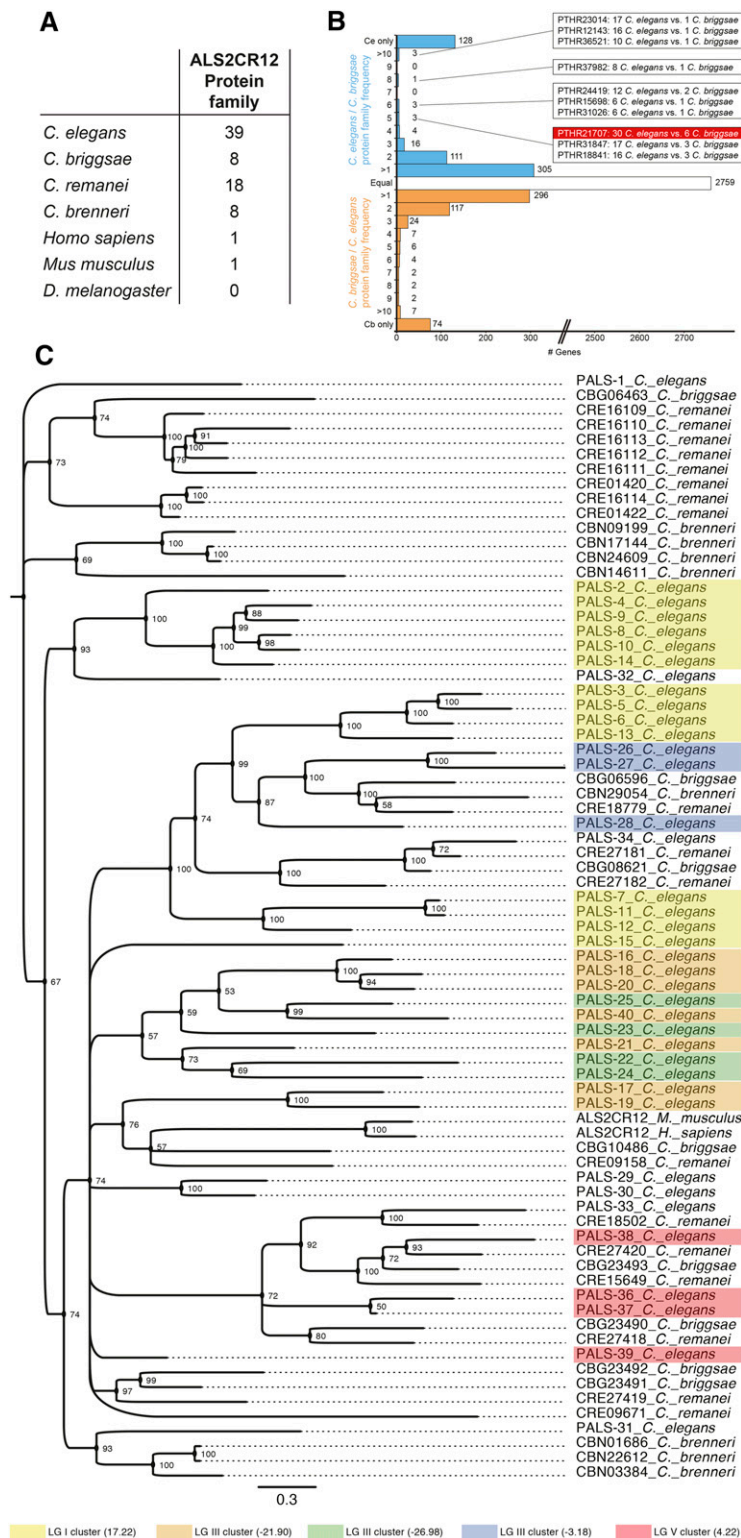


Figure 5 Sequence and genomic analysis of *pals* gene family in *C. elegans* and other species. (A) Number of genes member of the ALS2CR12 protein family as predicted by InterPro (IPR026674) and/or Panther (PTHR21707). (B) Graph representing the *C. elegans* to *C. briggsae* Panther protein family frequency for protein families enriched in *C. elegans* (blue), *C. briggsae* to *C. elegans* ratio for protein families enriched in *C. briggsae* (orange), or white for Panther protein families predicted to occur in the same number of genes in both species. Boxes indicate the gene counts for highly enriched protein families in *C. elegans*. ALS2CR12 family (Panther protein family PTHR21707) is highlighted in red. The number of Panther protein family hits for the *C. elegans* and *C. briggsae* genome were obtained from WormBase. (C) Phylogram of ALS2CR12 family-containing genes, including *C. elegans* paralogs and orthologs from *C. briggsae*, *C. remanei*, *C. brenneri*, human, and mouse. Node values indicate posterior probabilities for each split as percent. The scale bar indicates average branch length measured in expected substitutions per site. #, number; Cb, *C. briggsae*; Ce, *C. elegans*; LG, linkage group.

2010). We find that loss of *rde-4* completely suppresses the *pals-22* mutant phenotype (Figure 7, B and C). This result might be in contrast with previous reports showing that *rde-4* can also act as a transgene antisilencing factor (Fischer *et al.* 2013). Interestingly, we find that *ric-19^{prom6}::NLS::gfp* expression increases in the *rde-4(ne301)* alone background,

in agreement with previous studies suggesting a background level of transgene silencing in wild-type worms (Mello and Fire 1995; Lehner *et al.* 2006). Although *rde-4* could also play other roles in gene expression independently of siRNA production, our result suggests that the gene silencing mediated by *pals-22* deficiency requires the production of small dsRNAs.

Table 2 List of all *pals* genes

Linkage Group	Gene	Cosmid-Based Name	WB Gene ID	Location	Protein Size (kDa) ^a	Similarity to PALS-22 ^b	Upregulated After Microsporidal and/or Viral Infection ^c	Expression Enrichment (ModEncode) ^d
LGI	<i>pals-1</i>	F15D3.8	WBGene00008858	I:9.26	34.9		Yes	Broad
	<i>pals-2</i>	C17H1.3	WBGene00007656	I:17.22	45.2		Yes ^e	Male L4
	<i>pals-3</i>	C17H1.4	WBGene00007657	I:17.22	39.2		Yes ^e	Male L4
	<i>pals-4</i>	C17H1.5	WBGene00007658	I:17.22	40.5		Yes	Male L4
	<i>pals-5</i>	C17H1.6	WBGene00007659	I:17.22	35.4		Yes	Male L4
	<i>pals-6</i>	C17H1.7	WBGene00007660	I:17.22	46.8		Yes	Male L4
	<i>pals-7</i>	C17H1.8	WBGene00007661	I:17.22	39.3		Yes ^e	Male L4
	<i>pals-8</i>	C17H1.9	WBGene00007662	I:17.22	41.1		Yes ^e	Male L4
	<i>pals-9</i>	C17H1.10	WBGene00044237	I:17.22	34.9		Yes	Male L4
	<i>pals-10</i>	C17H1.11	WBGene00044711	I:17.22	8.4			Male L4
	<i>pals-11</i>	C17H1.13	WBGene00044708	I:17.22	38.5		Yes	Male L4
	<i>pals-12</i>	C17H1.14	WBGene00044709	I:17.22	46.4		Yes	Male L4
	<i>pals-13</i>	Y26D4A.8	WBGene00012505	I:17.22	43.5			Male L4
	<i>pals-14</i>	F22G12.1	WBGene00009061	I:17.22	41.3		Yes	Male L4
	<i>pals-15</i>	F22G12.7	WBGene00044707	I:17.22	41.7		Yes	Male L4
LGIII	<i>pals-16</i>	Y82E9BR.4	WBGene00022337	III:−21.90	37.7			Broad
	<i>pals-17</i>	Y82E9BR.13	WBGene00022346	III:−21.90	30	3e−05	Yes	Broad
	<i>pals-18</i>	Y82E9BR.21	WBGene00022353	III:−21.90	29.2			Broad
	<i>pals-19</i>	Y82E9BR.23	WBGene00044761	III:−21.90	30.6	3e−05		Broad
	<i>pals-20</i>	Y82E9BR.25	WBGene00194746	III:−21.90	27.5			Low
	<i>pals-21</i>	Y82E9BR.32	WBGene00255597	III:−21.90	20.7	3e−12		Low
	<i>pals-40</i>	Y82E9BR.11	WBGene00022344	III:−21.90	24	4e−06	Yes	Low
	<i>pals-22</i>	C29F9.1	WBGene00016216	III:−26.98	32.7	0.0		Broad
	<i>pals-23</i>	C29F9.3	WBGene00016218	III:−26.98	34.7	6e−05		Broad
	<i>pals-24</i>	C29F9.4	WBGene00016219	III:−26.98	35.7	2e−10		Broad
	<i>pals-25</i>	T17A3.2	WBGene00020540	III:−26.98	34.7	2e−06		Broad
	<i>pals-26</i>	B0284.1	WBGene00007131	III:−3.18	49.2		Yes	Broad
	<i>pals-27</i>	B0284.2	WBGene00007132	III:−3.18	48.2		Yes ^e	Broad
	<i>pals-28</i>	B0284.4	WBGene00007134	III:−3.18	48		Yes ^e	Male L4
	<i>pals-29</i>	T27E7.6	WBGene00012091	IV:12.20	48.4	7e−05	Yes	Male L4 + dauer
LGIV	<i>pals-30</i>	Y57G11B.1	WBGene00013294	IV:12.23	48.9	3e−04	Yes	Broad
	<i>pals-31</i>	F48G7.2	WBGene00018614	V:−19.97	25.6		Yes	Broad
LGV	<i>pals-32</i>	C31B8.4	WBGene00016281	V:−12.86	43.7	9e−05	Yes ^e	Male L4 + dauer
	<i>pals-33</i>	W08A12.4	WBGene00021081	V:−8.24	52		Yes ^e	Male L4 + dauer
	<i>pals-34</i>	F26D11.6	WBGene00017823	V:0.98	19.1			broad
	<i>pals-36^f</i>	C54D10.12	WBGene00044787	V:4.22	55.4	2e−04	Yes	Male L4 + dauer
	<i>pals-37</i>	C54D10.14	WBGene00138721	V:4.22	87.3		Yes ^e	Male L4 + dauer
	<i>pals-38</i>	C54D10.8	WBGene00008302	V:4.22	48.1		Yes ^e	Male L4 + dauer
	<i>pals-39</i>	C54D10.7	WBGene00008301	V:4.22	48.5	2e−05	Yes	Male L4 + dauer

Protein family assignments made by either Interpro (IPR026673) or Panther (PTHR21707), v9.0. A newer release of Panther does not include *pals-7*, *10*, or *11* in the ALS2CR12 family, yet the sequence similarity of these genes to neighboring *pals* genes is clear. WB Gene ID, WormBase gene identifier.

^a Protein size of largest predicted isoform.

^b PSI Blast e-value. Similarities only above e−04 threshold are shown.

^c Twenty-eight *pals* genes upregulated in response to microsporidal and/or viral infection according to Bakowski *et al.* (2014) and Chen *et al.* (2017) (out of 218 significantly upregulated genes).

^d According to Celniker *et al.* (2009).

^e Ten *pals* genes upregulated after cadmium exposure according to Cui *et al.* (2007) (out of 237 significantly upregulated genes).

^f Possibly a pseudogene.

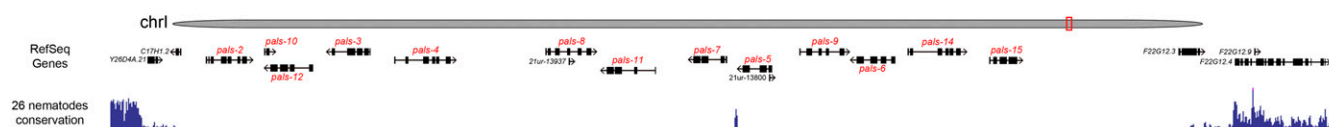
Transgene-silencing phenotypes have been observed in mutants that affect multiple distinct small RNA pathways, and exogenous RNAi responses are often enhanced in these mutants (Simmer *et al.* 2002; Lehner *et al.* 2006; Fischer *et al.* 2013). Thus, we tested whether *pals-22(ot811)* shows an enhanced exogenous RNAi response. Using *rnf-3(pk1426)* as a positive control (Simmer *et al.* 2002), we detect enhanced *dpy-13* or *cel-1* RNAi phenotypes in *pals-22(ot811)* (dsRNA delivered by feeding; Figure 7D). We conclude that *pals-22* physiological function might be related to the regulation of RNAi-dependent

silencing, via a mechanism critical to its action as an anti-silencing factor.

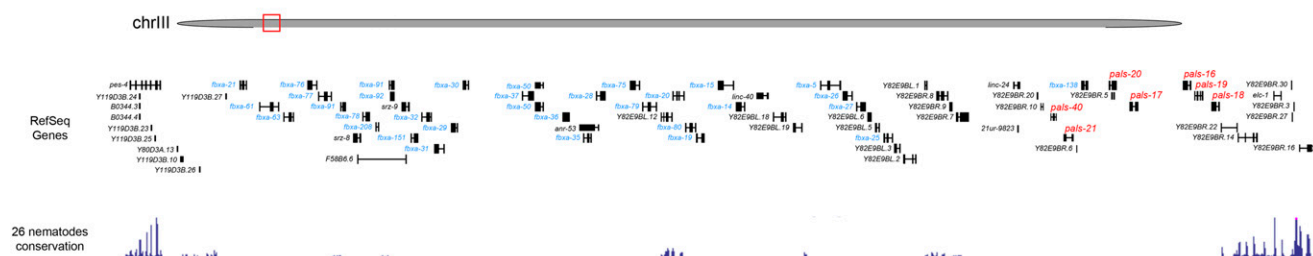
Locomotory and aging defects of *pals-22* mutant animals

The loss of *pals-22* has striking physiological consequences. Since casual observation of *pals-22* mutants indicates defects in locomotion, we quantified these defects, by measured swimming behavior (Restif *et al.* 2014) and crawling activity, comparing wild-type, *pals-22(ot810)*, *pals-22(ot811)*, and *pals-22(ot811)* carrying a wild-type copy of *pals-22* on an

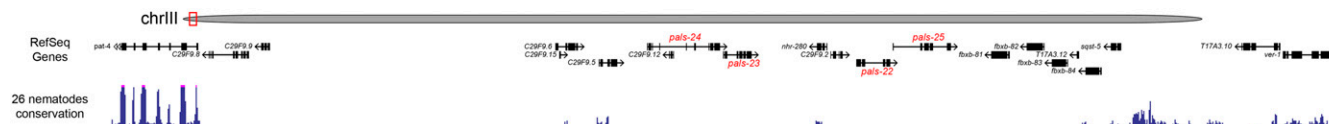
LG I cluster (17.22 cM)



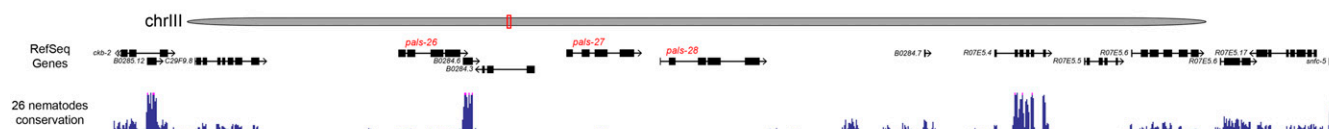
LG III cluster (-21.90 cM)



LG III cluster (-26.98 cM)



LG III cluster (-3.18 cM)



LG V cluster (4.22 cM)

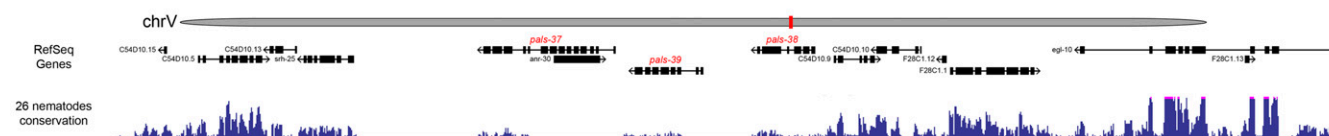


Figure 6 *C. elegans* *pals* genes are clustered and these clusters are poorly conserved. Schematics of different *C. elegans* genomic regions adapted from the University of California, Santa Cruz Genome Browser (<https://genome.ucsc.edu/>). One isoform per gene is shown. *pals* genes are indicated in red and *fbxa* genes in blue. The following regions are shown: chromosome I, cluster at position 17.22 cM, chrI: 13,099,564–13,160,497 bp; chromosome III, cluster at position –21.90 cM, chrIII: 1,215,665–1,423,550 bp; chromosome III, cluster at position –26.98 cM, chrIII: 89,907–159,962 bp; chromosome III, cluster at position –3.18 cM, chrIII: 4,368,479–4,405,090 bp; and chromosome V, cluster at position 4.22 cM, chrV: 12,427,096–12,462,015 bp. Chr; chromosome; LG, linkage group.

extrachromosomal array (Figure 8). Day 1 adult *pals-22* mutants show a poor performance in swimming assays as evaluated by multiple parameters, including low wave initiation rate (akin to thrash speed), travel speed (distance moved over time), brush stroke area (area covered in unit time), and activity index (Figure 8A). On agar plates animal movement was also clearly impaired, displaying significantly decreased traveling speed (Figure 8B). All

locomotory defects were rescued by the *pals-22*(+) extra-chromosomal array.

Apart from the locomotory defects, we also noted abnormal survival of *pals-22* mutants. Using standard life span assays, we observed premature death of *pals-22* mutants, commencing at about day 10 of adulthood (Figure 8C). We also find that 5-day-old adult animals display a signature change in aging, namely increased age pigment in the gut of adult

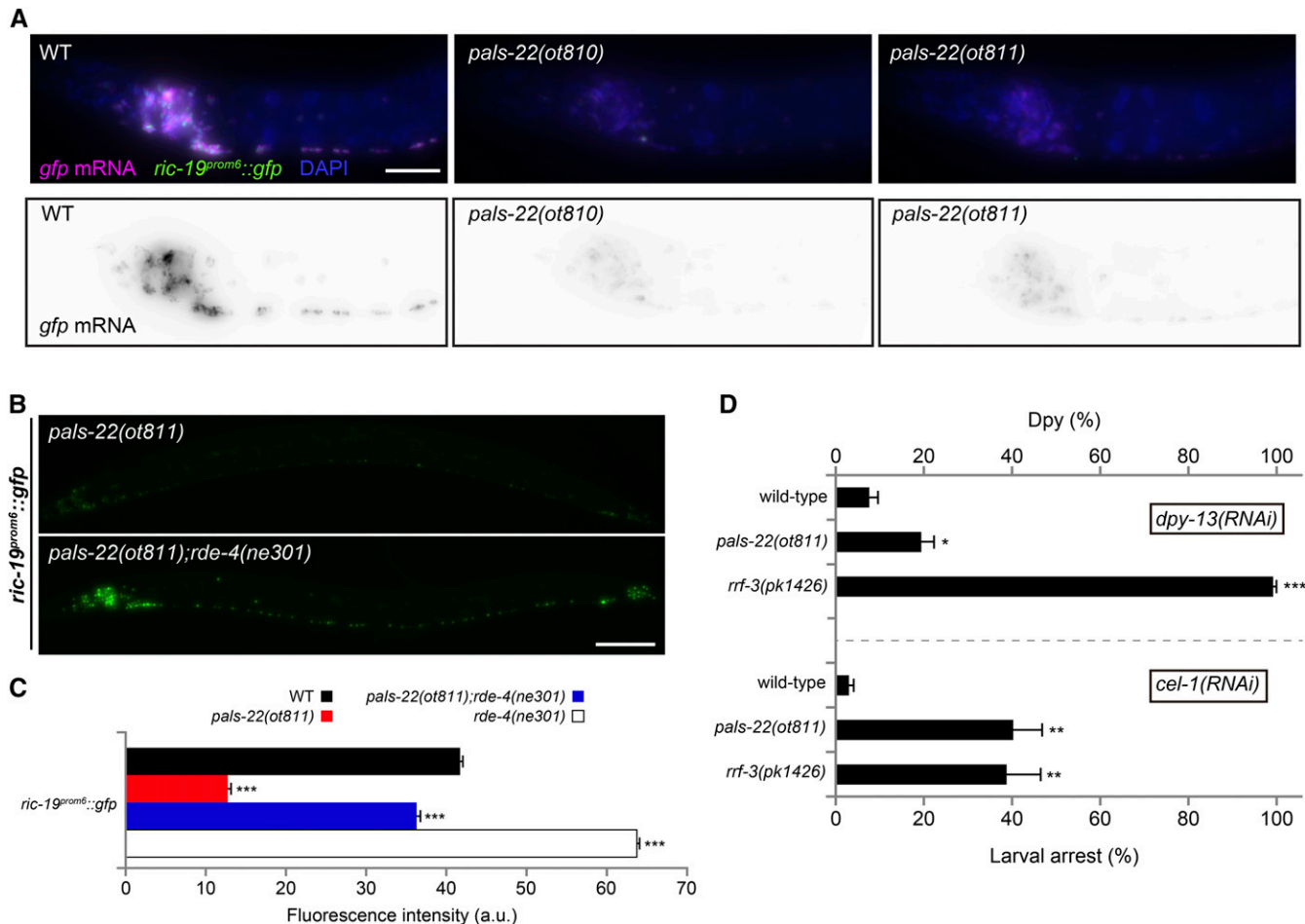


Figure 7 Somatic transgene silencing in *pals-22* depends on the RNAi pathway. (A) Wild-type (left), *pals-22(ot810)* (middle), and *pals-22(ot811)* (right) images of the anterior part of L3 worms showing reduced *gfp* mRNA levels in mutants compared to control wild-type worms as assessed by single-molecule fluorescence *in situ* hybridization. Individual transcripts shown as purple dots in top and as black dots in bottom panels. GFP expression is shown in green and DAPI staining is shown in blue. At least 20 animals were examined for each genotype. (B) Silencing of *ric-19^{prom6}::NLS::gfp* in *pals-22(ot811)* (top) is suppressed in *pals-22(ot811);rde-4(ne301)* double mutants (bottom). (C) Quantification of *otIs381[ric-19^{prom6}::NLS::gfp]* fluorescence intensity in wild-type, *pals-22(ot811)*, *pals-22(ot811);rde-4(ne301)*, and *rde-4(ne301)* mutants. The data are presented as mean + SEM. Unpaired *t*-test (comparisons to WT), *** $P < 0.001$; $n \geq 950$ for all genotypes. (D) Animals of the indicated genotype were grown on bacteria expressing *dpy-13* or *cel-1* dsRNA (feeding RNAi). For *dpy-13(RNAi)* (top), progeny were scored for the percentage of animals having a dumpy body shape (Dpy). For *cel-1(RNAi)* (bottom), the percentage of their progeny arresting at the L2 larval stage was determined. Wild-type *otIs381[ric-19^{prom6}::NLS::gfp]* and *rrf-3(pk1426)* mutants were used as negative and positive controls. The data are presented as mean + SEM among the data collected from at least four independent experiments. Unpaired *t*-tests were performed for *pals-22(ot811)* and *rrf-3(pk1426)* compared to WT; * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$. Bar, 10 μ m (A) and 50 μ m (B). a.u., arbitrary units; RNAi, RNA interference; WT, wild-type.

worms (Figure 8D). In light of these premature aging phenotypes, we surmise that the locomotory defects described above may also be an indication of premature aging.

Discussion

We have described here a large, unusual gene family in *C. elegans*. In the context of whole-animal transcriptome profiling under different conditions, expression of members of the *pals* gene family has previously been shown to be induced upon various forms of cellular insults, ranging from exposure to intracellular fungal infections, to viral infection, and to toxic compound exposure (Cui *et al.* 2007; Bakowski *et al.* 2014; Chen *et al.* 2017). We define here a function for one of

the family members, *pals-22*, in controlling the silencing of repetitive DNA sequences. Even though the biochemical function of PALS proteins is presently unclear, the upregulation of many *pals* genes under conditions of cellular stress suggest that this gene family may be part of a host defense mechanism that protects animals from specific insults. The *C. elegans*-specific expansion of *pals* genes may relate to their potential function in fending off species-specific stressors and/or encounters with species-specific pathogens.

The mutant phenotype of *pals-22* as a transgene silencer, as well as the dependence of this phenotype on small RNA production, indicates that PALS-22 may control gene expression via small RNA molecules. A role of PALS-22 in controlling gene expression is also illustrated in a parallel study in

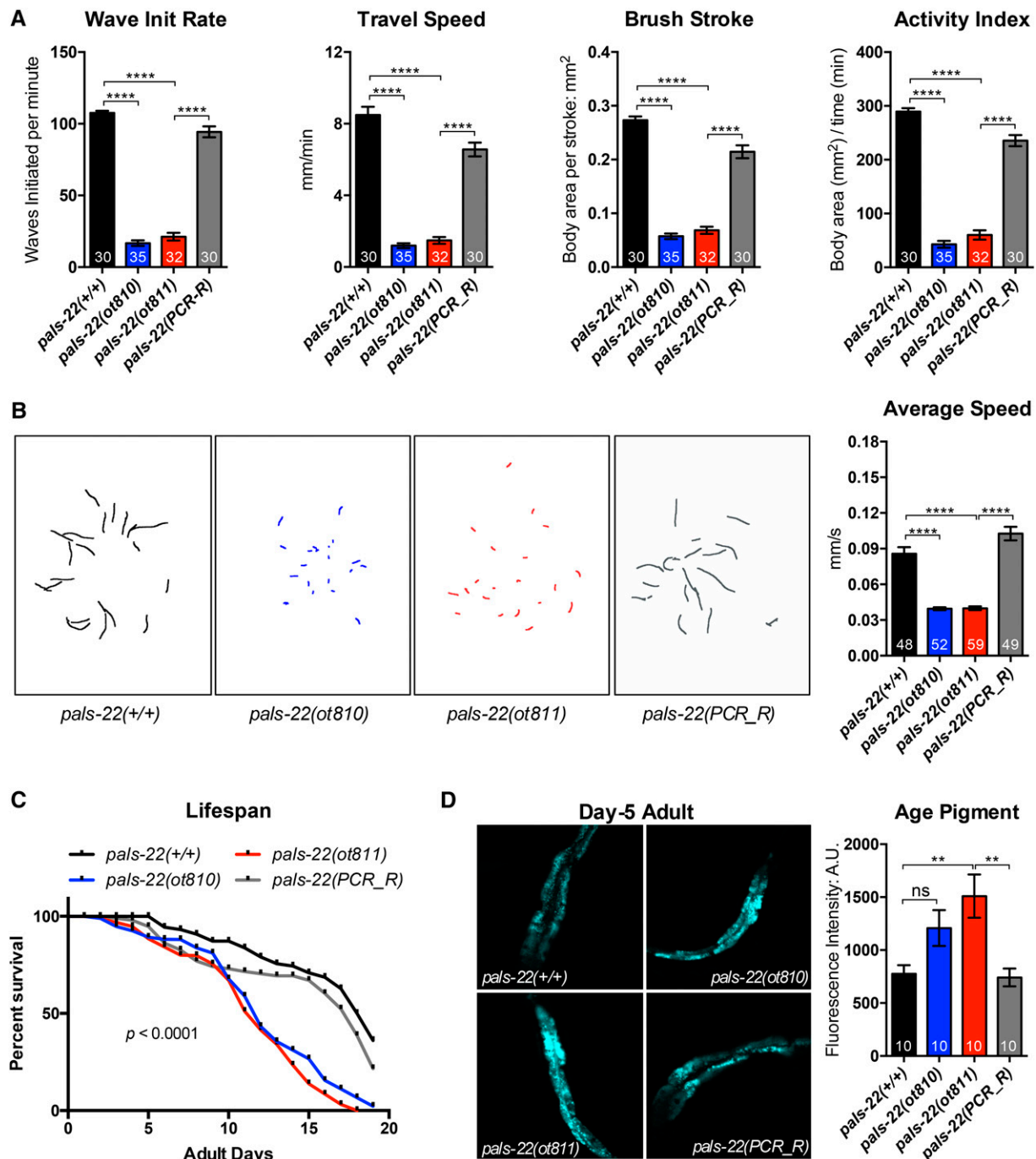


Figure 8 *pals-22* mutants show defective locomotion and early onset of aging traits. (A) Day 1 adult *pals-22* mutants exhibit defective swimming features, including decreased wave initiation (init) rate, travel speed, brush stroke, and activity index. *pals-22(PCR_R)* is the *pals-22(ot811)* mutant carrying a wild-type copy of *pals-22* on an extrachromosomal array amplified by PCR. Data shown are mean $\pm$ SEM of each parameters, $n = 30$ – 35 (number indicated in each bar) from two independent experiments, **** $P < 0.0001$ (one-way ANOVA) compared to related control. (B) *pals-22* mutants crawl slowly on agar plates. The left panel shows the crawling path of each animal in 30 sec. The right panel shows the mean $\pm$ SEM of average crawling speed (millimeters per second) for day 1 adults from two independent experiments, $n = \sim 55$ worms for each genotype, **** $P < 0.0001$ (one-way ANOVA) compared to related control. (C) *pals-22* mutants have shorter life spans compared to wild-type or *pals-22(PCR_R)*. Survival study was initiated with 96 worms (for each genotype) at L4 stage (day 0). Data shown is one represented trial of life span; two additional independent trials also show similar changes of life span in *pals-22* mutants, $P < 0.0001$ (Log-rank test), comparing the wild-type to mutants, or comparing *pals-22(PCR_R)* against *pals-22(ot811)*. (D) *pals-22* mutants exhibit early accumulation of age pigment. The left panel shows representative pictures of age pigment (excitation: 364 nm; emission: 380–420 nm). The right panel shows the mean $\pm$ SEM of age pigment auto-fluorescence intensity of day 5 adults, $n = 10$ worms from each genotype, ns (not significant) or ** $P < 0.01$ (one-way ANOVA) compared to related control. A.U., arbitrary units.

which *pals-22* has been found to be required for the proper regulation of a battery of stress- and microsporidial infection-induced genes, including many of the *pals* genes themselves (Reddy *et al.* 2017). While the function of *pals-22* in the RNAi process is not clear, there are numerous examples of mutants in which somatic transgene silencing is induced as a result of an increase in RNAi sensitivity, including *rrf-3*, *eri-1*, *lin-35*, and others (Simmer *et al.* 2002; Kennedy *et al.* 2004; Kim *et al.* 2005; Wang *et al.* 2005; Lehner *et al.* 2006; Fischer *et al.* 2008, 2011, 2013). A set of genes classically linked to somatic transgene silencing are the class B synMuv genes. synMuv B genes are so-called due to a synthetic phenotype found when mutations are combined with those in synMuv A genes (Ferguson and Horvitz 1989). As *pals-22*-defective worms, mutants for synMuv B genes show reduced expression from repetitive DNA transgenes. While mutations in some synMuv B genes do not result in transgene silencing (e.g., *lin-36*) (Hsieh *et al.* 1999) it has been shown that the silencing of repetitive transgenes in synMuv B mutants results from an enhanced somatic cell RNAi (Wang *et al.* 2005; Lehner *et al.* 2006). Interestingly, *tam-1* and *lex-1* mutants are genetically class B synMuvs; however, some synMuv B mutant characteristics such as enhanced RNAi or defects in P granule localization are not observed in *tam-1* mutants, while they still show a transgene-silencing phenotype (Wang *et al.* 2005; Lehner *et al.* 2006; Tseng *et al.* 2007). This suggests that they act through a different mechanism than other synMuv B. While *pals-22* animals present some overlapping phenotypes with synMuv B genes, further analysis will reveal in detail the mechanism for somatic transgene silencing in *pals-22* mutants. Intriguingly, since the expression levels per copy of repetitive tandem arrays are much lower than for endogenous genes, transgene-targeting siRNAs may be already abundant in wild-type transgenic strains, indicating a background level of transgene silencing in wild-type worms (Mello and Fire 1995). An interesting hypothesis to be tested is that in a mutant, such as *pals-22*, these transgene-targeted siRNAs may be reduced, perhaps because of a shift in the balance between the loading of transgene siRNAs into a silencing Argonaute (e.g., *NRDE-3*) vs. an antisilencing Argonaute (e.g., *CSR-1*) (Shirayama *et al.* 2012; Fischer *et al.* 2013).

The precise biochemical function of any PALS protein remains obscure. The only notable sequence relationship that we could find points to potential biochemical function of the PALS proteins in proteostasis. One of the LGIII clusters of *pals* genes (III: −21.90 cM) also contains a large number of *fbxa* genes, another large gene family specifically expanded in *C. elegans* (Thomas 2006) (Figure 6). *fbxa* genes code for F-box proteins that are involved in protein degradation (Thomas 2006). One of the FBXA proteins in the LGIII cluster, *FBXA-138*, displays sequence similarities to a number of distinct PALS proteins within and outside the LGIII *pals* cluster, including *PALS-22*, *PALS-23*, *PALS-32*, *PALS-25*, and *PALS-1*. Even though PALS proteins are not predicted to contain a canonical F-box, it is conceivable that their distant sequence relationship to F-box proteins may suggest a role of PALS

proteins in protein degradation. How, in the case of *PALS-22*, such a function may relate to the control of gene expression via small RNA molecules is not clear.

Although it is difficult to unambiguously distinguish accelerated aging from general sickness, young adult *pals-22* mutants clearly exhibit multiple features of aged animals: impaired mobility, elevated age pigments/lipofuscin, and shortened life span. Increased expression of repetitive sequences has been documented in aging *C. elegans*, *Drosophila*, and humans, and has been suggested to contribute to genomic instability and cell dysfunction (Sedivy *et al.* 2013); the physiological effect of decreases in the expression of repetitive sequences has not been explored. However, we note that other transgene-silencing mutants (which are also RNAi hypersensitive), like *rrf-3* and *eri-1*, do not display an aging defect (Zhang *et al.* 2009; Ren *et al.* 2012). *pals-22* may therefore be involved in novel aspects of small RNA-dependent gene silencing that may control the expression of genes involved in animal physiology.

Acknowledgments

We thank Chi Chen for the generation of transgenic strains, the *Caenorhabditis* Genetics Center for providing strains (funded by the National Institutes of Health Office of Research Infrastructure Programs, P40 OD010440), Kevin Howe and Gary Williams at WormBase for their help obtaining and analyzing protein family data, Sangeena Salam for initial life span assays, the Million Mutation Project for providing the *pals-19(gk166606)* and *pals-25(gk891046)* alleles, Emily Troemel for communicating unpublished results and providing the *jySi37* strain, Dylan Rahe for providing the GFP-tagged *che-1(ot856)* allele, HaoSheng Sun for providing the *lin-4^{prom}::yfp* MiniMos strain, and members of the Hobert laboratory for comments on this manuscript. This work was supported by grant R01-AG046358 to M.D., and a New Jersey Commission on Cancer Research Postdoctoral Fellowship (DHFS16PPC070) to G.W. and the Howard Hughes Medical Institute (O.H.). E.L.-D. was funded by an European Molecular Biology Organization long-term fellowship.

Literature Cited

- Bakowski, M. A., C. A. Desjardins, M. G. Smelkinson, T. L. Dunbar, I. F. Lopez-Moyado *et al.*, 2014 Ubiquitin-mediated response to microsporidia and virus infection in *C. elegans*. *PLoS Pathog.* 10: e1004200.
- Brenner, S., 1974 The genetics of *Caenorhabditis elegans*. *Genetics* 77: 71–94.
- Celniker, S. E., D. A. Wheeler, B. Kronmiller, J. W. Carlson, A. Halpern *et al.*, 2002 Finishing a whole-genome shotgun: release 3 of the *Drosophila melanogaster* euchromatic genome sequence. *Genome Biol.* 3: RESEARCH0079.
- Celniker, S. E., L. A. Dillon, M. B. Gerstein, K. C. Gunsalus, S. Henikoff *et al.*, 2009 Unlocking the secrets of the genome. *Nature* 459: 927–930.

- Chen, K., C. J. Franz, H. Jiang, Y. Jiang, and D. Wang, 2017 An evolutionarily conserved transcriptional response to viral infection in *Caenorhabditis nematodes*. *BMC Genomics* 18: 303.
- Chuong, E. B., N. C. Elde, and C. Feschotte, 2017 Regulatory activities of transposable elements: from conflicts to benefits. *Nat. Rev. Genet.* 18: 71–86.
- Cordaux, R., and M. A. Batzer, 2009 The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* 10: 691–703.
- Cui, M., and M. Han, 2007 Roles of chromatin factors in *C. elegans* development (May 3, 2007), *WormBook*, ed. The *C. elegans* Research Community, *WormBook*, doi/10.1895/wormbook.1.139.1, <http://www.wormbook.org>.
- Cui, Y., S. J. McBride, W. A. Boyd, S. Alper, and J. H. Freedman, 2007 Toxicogenomic analysis of *Caenorhabditis elegans* reveals novel genes and pathways involved in the resistance to cadmium toxicity. *Genome Biol.* 8: R122.
- de Koning, A. P., W. Gu, T. A. Castoe, M. A. Batzer, and D. D. Pollock, 2011 Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.* 7: e1002384.
- Doitsidou, M., N. Flames, A. C. Lee, A. Boyanov, and O. Hobert, 2008 Automated screening for mutants affecting dopaminergic-neuron specification in *C. elegans*. *Nat. Methods* 5: 869–872.
- Doitsidou, M., R. J. Poole, S. Sarin, H. Bigelow, and O. Hobert, 2010 *C. elegans* mutant identification with a one-step whole-genome-sequencing and SNP mapping strategy. *PLoS One* 5: e15435.
- Doolittle, W. F., and C. Sapienza, 1980 Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284: 601–603.
- Edelstein, A., N. Amodaj, K. Hoover, R. Vale, and N. Stuurman, 2010 Computer control of microscopes using μ Manager. *Curr. Protoc. Mol. Biol.* CHAPTER: Unit14.20.
- Elbashir, S. M., W. Lendeckel, and T. Tuschl, 2001 RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* 15: 188–200.
- Ferguson, E. L., and H. R. Horvitz, 1989 The multivulva phenotype of certain *Caenorhabditis elegans* mutants results from defects in two functionally redundant pathways. *Genetics* 123: 109–121.
- Finn, R. D., T. K. Attwood, P. C. Babbitt, A. Bateman, P. Bork *et al.*, 2017 InterPro in 2017-beyond protein family and domain annotations. *Nucleic Acids Res.* 45: D190–D199.
- Fischer, S. E., M. D. Butler, Q. Pan, and G. Ruvkun, 2008 Trans-splicing in *C. elegans* generates the negative RNAi regulator ERI-6/7. *Nature* 455: 491–496.
- Fischer, S. E., T. A. Montgomery, C. Zhang, N. Fahlgren, P. C. Breen *et al.*, 2011 The ERI-6/7 helicase acts at the first stage of an siRNA amplification pathway that targets recent gene duplications. *PLoS Genet.* 7: e1002369.
- Fischer, S. E., Q. Pan, P. C. Breen, Y. Qi, Z. Shi *et al.*, 2013 Multiple small RNA pathways regulate the silencing of repeated and foreign genes in *C. elegans*. *Genes Dev.* 27: 2678–2695.
- Friedli, M., and D. Trono, 2015 The developmental control of transposable elements and the evolution of higher species. *Annu. Rev. Cell Dev. Biol.* 31: 429–451.
- Gent, J. I., A. T. Lamm, D. M. Pavelec, J. M. Maniar, P. Parameswaran *et al.*, 2010 Distinct phases of siRNA synthesis in an endogenous RNAi pathway in *C. elegans* soma. *Mol. Cell* 37: 679–689.
- Gerstbrein, B., G. Stamatas, N. Kollias, and M. Driscoll, 2005 In vivo spectrofluorimetry reveals endogenous biomarkers that report healthspan and dietary restriction in *Caenorhabditis elegans*. *Aging Cell* 4: 127–137.
- Goke, J., and H. H. Ng, 2016 CTRL+INSERT: retrotransposons and their contribution to regulation and innovation of the transcriptome. *EMBO Rep.* 17: 1131–1144.
- Goodier, J. L., and H. H. Kazazian, 2008 Retrotransposons revisited: the restraint and rehabilitation of parasites. *Cell* 135: 23–35.
- Grishok, A., J. L. Sinskey, and P. A. Sharp, 2005 Transcriptional silencing of a transgene by RNAi in the soma of *C. elegans*. *Genes Dev.* 19: 683–696.
- Hacein-Bey-Abina, S., C. Von Kalle, M. Schmidt, M. P. McCormack, N. Wulffraat *et al.*, 2003 LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 302: 415–419.
- Hadano, S., C. K. Hand, H. Osuga, Y. Yanagisawa, A. Otomo *et al.*, 2001 A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. *Nat. Genet.* 29: 166–173.
- Hobert, O., 2002 PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *Biotechniques* 32: 728–730.
- Hsieh, J., J. Liu, S. A. Kostas, C. Chang, P. W. Sternberg *et al.*, 1999 The RING finger/B-box factor TAM-1 and a retinoblastoma-like protein LIN-35 modulate context-dependent gene silencing in *Caenorhabditis elegans*. *Genes Dev.* 13: 2958–2970.
- Huelsenbeck, J. P., and F. Ronquist, 2001 MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.
- Jacobs, F. M. J., D. Greenberg, N. Nguyen, M. Haeussler, A. D. Ewing *et al.*, 2014 An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. *Nature* 516: 242–245.
- Ji, N., and A. van Oudenaarden, 2012 Single molecule fluorescent in situ hybridization (smFISH) of *C. elegans* worms and embryos (December 13, 2012), *WormBook*, ed. The *C. elegans* Research Community, *WormBook*, doi/10.1895/wormbook.1.153.1, <http://www.wormbook.org>.
- Kamath, R. S., and J. Ahringer, 2003 Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods* 30: 313–321.
- Kazazian, Jr., H. H., 2004 Mobile elements: drivers of genome evolution. *Science* 303: 1626–1632.
- Kelly, W. G., S. Xu, M. K. Montgomery, and A. Fire, 1997 Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* 146: 227–238.
- Kennedy, S., D. Wang, and G. Ruvkun, 2004 A conserved siRNA-degrading RNase negatively regulates RNA interference in *C. elegans*. *Nature* 427: 645–649.
- Kim, J. K., H. W. Gabel, R. S. Kamath, M. Tewari, A. Pasquinelli *et al.*, 2005 Functional genomic analysis of RNA interference in *C. elegans*. *Science* 308: 1164–1167.
- Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody *et al.*, 2001 Initial sequencing and analysis of the human genome. *Nature* 409: 860–921.
- Law, J. A., and S. E. Jacobsen, 2010 Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11: 204–220.
- Le Thomas, A., A. K. Rogers, A. Webster, G. K. Marinov, S. E. Liao *et al.*, 2013 Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev.* 27: 390–399.
- Lehner, B., A. Calixto, C. Crombie, J. Tischler, A. Fortunato *et al.*, 2006 Loss of LIN-35, the *Caenorhabditis elegans* ortholog of the tumor suppressor p105Rb, results in enhanced RNA interference. *Genome Biol.* 7: R4.
- Lynch, M., and J. S. Conery, 2003 The origins of genome complexity. *Science* 302: 1401–1404.
- Mello, C., and A. Fire, 1995 DNA transformation. *Methods Cell Biol.* 48: 451–482.
- Mi, H., A. Muruganujan, J. T. Casagrande, and P. D. Thomas, 2013 Large-scale gene function analysis with the PANTHER classification system. *Nat. Protoc.* 8: 1551–1566.
- Mi, H., X. Huang, A. Muruganujan, H. Tang, C. Mills *et al.*, 2017 PANTHER version 11: expanded annotation data from gene ontology and reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.* 45: D183–D189.

- Minevich, G., D. S. Park, D. Blankenberg, R. J. Poole, and O. Hobert, 2012 CloudMap: a cloud-based pipeline for analysis of mutant genome sequences. *Genetics* 192: 1249–1269.
- Nowick, K., A. T. Hamilton, H. Zhang, and L. Stubbs, 2010 Rapid sequence and expression divergence suggest selection for novel function in primate-specific KRAB-ZNF genes. *Mol. Biol. Evol.* 27: 2606–2617.
- Nussbaum-Krammer, C. I., M. F. Neto, R. M. Brielmann, J. S. Pedersen, and R. I. Morimoto, 2015 Investigating the spreading and toxicity of prion-like proteins using the metazoan model organism *C. elegans*. *J. Vis. Exp.* 95: 52321.
- Orgel, L. E., and F. H. Crick, 1980 Selfish DNA: the ultimate parasite. *Nature* 284: 604–607.
- Parker, G. S., D. M. Eckert, and B. L. Bass, 2006 RDE-4 preferentially binds long dsRNA and its dimerization is necessary for cleavage of dsRNA to siRNA. *RNA* 12: 807–818.
- Quenneville, S., P. Turelli, K. Bojkowska, C. Raclot, S. Offner *et al.*, 2012 The KRAB-ZFP/KAP1 system contributes to the early embryonic establishment of site-specific DNA methylation patterns maintained during development. *Cell Rep.* 2: 766–773.
- Reddy, K. C., T. Dror, J. Dowa, J. Panek, K. Chen *et al.*, 2017 An intracellular pathogen response pathway promotes proteostasis in *C. elegans*. *bioRxiv*. DOI: <https://doi.org/10.1101/145235>.
- Ren, Y., S. Yang, G. Tan, W. Ye, D. Liu *et al.*, 2012 Reduction of mitoferrin results in abnormal development and extended lifespan in *Caenorhabditis elegans*. *PLoS One* 7: e29666.
- Restif, C., C. Ibáñez-Ventoso, M. M. Vora, S. Guo, D. Metaxas *et al.*, 2014 CeleST: computer vision software for quantitative analysis of *C. elegans* swim behavior reveals novel features of locomotion. *PLoS Comput. Biol.* 10: e1003702.
- Ronquist, F., and J. P. Huelsenbeck, 2003 MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
- Rowe, H. M., J. Jakobsson, D. Mesnard, J. Rougemont, S. Reynard *et al.*, 2010 KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* 463: 237–240.
- Schlesinger, S., and S. P. Goff, 2015 Retroviral transcriptional regulation and embryonic stem cells: war and peace. *Mol. Cell. Biol.* 35: 770–777.
- Schneider, C. A., W. S. Rasband, and K. W. Eliceiri, 2012 NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9: 671–675.
- Sedivy, J. M., J. A. Kreiling, N. Neretti, M. De Cecco, S. W. Criscione *et al.*, 2013 Death by transposition - the enemy within? *Bioessays* 35: 1035–1043.
- Shapiro, J. A., and R. von Sternberg, 2005 Why repetitive DNA is essential to genome function. *Biol. Rev. Camb. Philos. Soc.* 80: 227–250.
- Shirayama, M., M. Seth, H.-C. Lee, W. Gu, T. Ishidate *et al.*, 2012 piRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline. *Cell* 150: 65–77.
- Simmer, F., M. Tijsterman, S. Parrish, S. Koushika, M. Nonet *et al.*, 2002 Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi. *Curr. Biol.* 12: 1317.
- Stefanakis, N., I. Carrera, and O. Hobert, 2015 Regulatory logic of pan-neuronal gene expression in *C. elegans*. *Neuron* 87: 733–750.
- Stein, L. D., Z. Bao, D. Blasiar, T. Blumenthal, M. R. Brent *et al.*, 2003 The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLoS Biol.* 1: E45.
- Stinchcomb, D. T., J. E. Shaw, S. H. Carr, and D. Hirsh, 1985 Extrachromosomal DNA transformation of *Caenorhabditis elegans*. *Mol. Cell. Biol.* 5: 3484–3496.
- Tabara, H., M. Sarkissian, W. G. Kelly, J. Fleenor, A. Grishok *et al.*, 1999 The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* 99: 123–132.
- Tabara, H., E. Yigit, H. Siomi, and C. C. Mello, 2002 The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DEXH-box helicase to direct RNAi in *C. elegans*. *Cell* 109: 861–871.
- Thomas, J. H., 2006 Adaptive evolution in two large families of ubiquitin-ligase adapters in nematodes and plants. *Genome Res.* 16: 1017–1030.
- Thompson, O., M. Edgley, P. Strasbourger, S. Flibotte, B. Ewing *et al.*, 2013 The million mutation project: a new approach to genetics in *Caenorhabditis elegans*. *Genome Res.* 23: 1749–1762.
- Toth, K. F., D. Pezic, E. Stuwe, and A. Webster, 2016 The piRNA pathway guards the germline genome against transposable elements. *Adv. Exp. Med. Biol.* 886: 51–77.
- Tseng, R. J., K. R. Armstrong, X. Wang, and H. M. Chamberlin, 2007 The bromodomain protein LEX-1 acts with TAM-1 to modulate gene expression in *C. elegans*. *Mol. Genet. Genomics* 278: 507–518.
- Vasale, J. J., W. Gu, C. Thivierge, P. J. Batista, J. M. Claycomb *et al.*, 2010 Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. *Proc. Natl. Acad. Sci. USA* 107: 3582–3587.
- Wallace, I. M., O. O'Sullivan, D. G. Higgins, and C. Notredame, 2006 M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res.* 34: 1692–1699.
- Wang, D., S. Kennedy, D. Conte, Jr., J. K. Kim, H. W. Gabel *et al.*, 2005 Somatic misexpression of germline P granules and enhanced RNA interference in retinoblastoma pathway mutants. *Nature* 436: 593–597.
- Wheeler, T. J., J. Clements, and R. D. Finn, 2014 Skylign: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinformatics* 15: 7.
- Wolf, D., and S. P. Goff, 2009 Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature* 458: 1201–1204.
- Zamore, P. D., T. Tuschl, P. A. Sharp, and D. P. Bartel, 2000 RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* 101: 25–33.
- Zhang, M., M. Poplawski, K. Yen, H. Cheng, E. Bloss *et al.*, 2009 Role of CBP and SATB-1 in aging, dietary restriction, and insulin-like signaling. *PLoS Biol.* 7: e1000245.

Communicating editor: D. Greenstein