Supplemental information to:

PETISCO is a novel protein complex required for 21U RNA biogenesis and embryonic viability

Ricardo J. Cordeiro Rodrigues[1], António Miguel de Jesus Domingues[1], Svenja Hellmann[1], Sabrina Dietz[2], Bruno F. M. de Albuquerque[1,3], Christian Renz[4], Helle D. Ulrich[4], Falk Butter[2], René F. Ketting[1]

Institute of Molecular Biology (IMB), Ackermannweg 4, 55128, Mainz, Germany

[1] IMB, Biology of Non-coding RNA Group

[2] IMB, Quantitative Proteomics Group

[3]Graduate Program in Areas of Basic and Applied Biology, University of Porto, 4099-003 Porto, Portugal.

[4]IMB, Maintenance of Genome Stability Group

correspondence: r.ketting@imb.de

Materials and Methods

Caenorhabditis elegans genetics and culture

Caenorhabditis elegans strains were cultured at 20°C on OP50 bacteria according to standard laboratory conditions [1] unless otherwise stated and with the exception of IP-LFQP experiments. Material for IP-LFQP were grown at 20°C in OP50 high density plates [2] for two generations and synchronized and plated in standard plates for the generation before harvest, lest indicated otherwise. Bristol N2 strain was used as reference wild type strain. Strain list of this study can be consulted below.

Strain reference	Genotype	Usage
N2	Wild Type	NGS; IP-LFQP and RIP
EG7833	*oxTi559[Peft-3::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119] I; unc-119(ed3)* III	CRISPR/Cas9
EG7893	*oxTi615[Peft-3::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119]; unc-119(ed3)* III	CRISPR/Cas9
EG8897	*unc-119(ed3)* III*; oxTi947[Peft-3::GFP::2xNLS::tbb-2 3'UTR + Cbr-unc-119]* V	CRISPR/Cas9
HT1593	*unc-119(ed3)* III	miniMos Transgenes
KK359	*tofu-6(it20); unc-4(e120)/mnC1 dpy-10(e128) unc-52(e444)* II.	
QA137	*tofu-6(yt2)* II*; ytEx100*	
RFK180	*mjIs144[Pmex-5::egfp::his-58::21UR-1_as::tbb-2(3'UTR)]; pid-1(xf14)* II	
RFK182	*pid-1(xf35)* II	NGS
RFK183	*pid-1(xf36)* II	
RFK184	*mjSi22[Pmex-5::mCherry::his-58::21UR-1_as::tbb-2(3'UTR)] I; pid-1(xf35)* II	Microscopy
RFK514	*unc-119(ed3)* III*; ife-3(xf101); oxTi947[Peft-3::GFP::2xNLS::tbb-2 3'UTR + Cbr-unc-119]/ oxTi664[Peft-3::TdTomato::H2B::unc-54 3'UTR + Cbr-unc-119]* V	
RFK515	*unc-119(ed3)* III*; ife-3(xf102); oxTi947[Peft-3::GFP::2xNLS::tbb-2 3'UTR + Cbr-unc-119]/ oxTi664[Peft-3::TdTomato::H2B::unc-54 3'UTR + Cbr-unc-119]* V	NGS
RFK523	*pid-3(tm2417) I/hT2[bli-4(e937) let-?(q782) qIs48]*(I;III*).*	NGS
RFK625	*unc-119(ed3)* III*; xfIs137[Ppid-3::pid-3::mCherry::Myc::pid-3(3'UTR); Cbr-unc-119]* II.	Microscopy
RFK647	*pid-1(xf14); mjIs144[Pmex-5::egfp::his-58::21UR-1_as::tbb-2(3'UTR)]* II; *xfIs117[Ppid-1::pid-1::mCherry::V5::pid-1(3'UTR); Cbr-unc-119]* V	RNAi essay
RFK679	*pid-3(tm2417); xfIs136[Ppid-3::pid-3::mCherry::Myc::pid-3*	IP-LFQP and

	(3'UTR); Cbr-unc-119] I	RIP
RFK684	*xfIs123[Ptofu-6::tofu-6::GFP::HA::tofu-6(3'UTR); Cbr-unc-119]* V	Microscopy
RFK696	*xfIs121[Pife-3::3xFLAG::mCherry::ife-3::ife-3(3'UTR) + Cbr-unc-119]* II*; ife-3(xf101); oxTi947[Peft-3::GFP::2xNLS::tbb-2 3'UTR + Cbr-unc-119]* V	IP-LFQP and RIP
RFK697	*xfIs121[Pife-3::3xFLAG::mCherry::ife-3::ife-3(3'UTR) + Cbr-unc-119]* II*; xfIs123 [Ptofu-6::tofu-6::GFP::HA::tofu-6(3'UTR) + Cbr-unc-119]* V	Microscopy
RFK700	*xfIs136[Ppid-3::pid-3::mCherry::Myc::pid-3(3'UTR); + Cbr-unc-119*] I;*xfIs123 [Ptofu-6::tofu-6::GFP::HA::tofu-6(3'UTR) + Cbr-unc-119]* V	Microscopy
RFK701	*xfIs136[Ppid-3::pid-3::mCherry::Myc::pid-3(3'UTR); + Cbr-unc-119]* I*; pid-1(xf35)* II	IP-LFQP
RFK703	*bnIs1[Ppie-1::GFP::pgl-1 + unc-119(+)], xfIs136[Ppid-3::pid-3::mCherry::Myc::pid-3(3'UTR)+ Cbr-unc-119]* I	Microscopy
RFK721	*tofu-6(it20), unc-4(e120)* II*; xfIs123[Ptofu-6::tofu-6::GFP::HA::tofu-6(3'UTR)+ Cbr-unc-119]* V	IP-LFQP
RFK742	*xfIs167[Perh-2::erh-2::EGFP::OLLAS::erh-2 (3'UTR) + Cbr-unc-119]* I*; erh-2(xf168), oxTi615[eft-3p::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119]; unc-119(ed3)* III	IP-LFQP and Microscopy
RFK810	*erh-2(xf168); oxTi615[Peft-3::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119]/ qC1[dpy-19(e1259) glp-1(q339) qIs26]* III	NGS
RFK861	*tost-1(xf191); oxTi615[Peft-3::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119] III*	
RFK874	*pid-3(xf149), oxTi559[Peft-3::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119] III/ hT2[bli-4(e937) let-?(q782) qIs48](I;III)*	
RFK875	*pid-3(xf153), oxTi559[Peft-3::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119] III/ hT2[bli-4(e937) let-?(q782) qIs48](I;III)*	
RFK876	*pid-3(xf151), oxTi559[Peft-3::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119] III/ hT2[bli-4(e937) let-?(q782) qIs48](I;III)*	
RFK905	*tost-1(xf194), oxTi615[Peft-3::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119] III/ qC1[dpy-19(e1259) glp-1(q339) qIs26] III*	NGS
RFK912	*tost-1(xf196), oxTi615[Peft-3::tdTomato::H2B::unc-54 3'UTR + Cbr-unc-119] III*	Temperature Shift

Mutant generation with CRISPR/Cas9 system

Mutant alleles were generated as described in [3]. gRNAs were selected under the criteria: NGG PAM site, highest GC content and specificity according to CRISPRdirect [4] and Zhang Lab's http://crispr.mit.edu. Two to three gRNA, singularly cloned into Addgene plasmid #46169, were injected (35ng/µl) together with Addgene plasmid #46168 (50ng/µl) and co-

injection marker pRR83 (5ng/µl) into adult worms (specific strains below). F1 worms positive for pharynx GFP expression were isolated, allowed reproduction, lysed in single worm lysis buffer (5 mM KCl, 2,5 mM MgCl2, 10 mM Tris pH=8,3, 0,45% IGEPAL, 0,45% Tween-20, 0,01% gelatin) and genotyped for mutations using NEB Taq DNA Polymerase (M0273X) according to manufacturer's instructions. Isolated mutants were outcrossed at least two times before balancing.

Target	**CRISPR/Cas9 Guide RNA**
ife-3	GCCTCCGTGCCGGGATTCGA
ife-3	GACACCCCCTCCAGAATCGC
ife-3	GAGCCCAGCGATTCTGGAGG
pid-3	gaaaATGGTTGCCCATCAGA
pid-3	GTGGAAGAATGTGCACGACG
pid-3	GGCGGATTTCAAGTCGAAAT
erh-2	gtgagaattattatgtttaa
erh-2	GAGCAGCTGATTTCTTGGAA
erh-2	GAAGATCATCATAGAAACAT
tost-1	GATAGTTctgaaacataacc
tost-1	GAGCTTCTTCTCATCAGTAG
tost-1	GATGGCAGTAGTCATtctga

miniMos transgene insertion and mapping

Random miniMos insertions were made through injection of unc-119(ed3) carrying worms. C. briggsae unc-119 was used as a selection marker. Injections and mapping were made in accordance to [5].

Immunostaining

Adult worms were dissected in Egg Buffer (25mM HEPES pH 7,4, 118mM NaCl, 48mM KCl, 2mM EDTA, 0,5mM EGTA) with 1%(v/v) Tween20 and fixed 5 minutes by adding 1:1 Egg Buffer+2% formaldehyde followed by a wash step in Egg Buffer. Cuticle was then removed by Freeze cracking [6]. An extra fixation step of 1 minute in -20°C Methanol preceded three washes in PBS (137mM NaCl, 2,7mM KCl, 10mM Na_2HPO_4, 2mM KH_2PO_4, pH=7,5) with 0,5%(v/v) Tween20 (PBST). After 1 hour in blocking buffer (PBST+10% Bovine Serum) samples were co-stained overnight at 4°C with 1:200 dilutions of RFP-Booster_Atto647N (Chromotek, rba647n-10) and GFP-Booster_Atto488 (Chromotek, gba488-10). Staining was followed with multiple PBST washes and samples were mounted in ProLong Gold Antifade Mountant (ThermoFisher Scientific, P36930).

Microscopy

Wide-field fluorescence microscopy images were obtained using a Leica DM6000B and Confocal microscopy images were acquired with a Leica TCS SP5. Images were processed using Leica LAS software, ImageJ and Adobe Photoshop.

RNAi experiments

HT115(DE3) bacteria carrying Timmons and Fire L4440 RNAi feeding vector [7] were grown over 10 hours and seeded directly onto RNAi plates (standard NGM; 1mM IPTG+and; 50µg/mL ampicillin). HT115(DE3) with empty L4440 or carrying *pid-3* or *tost-1* targeting RNA were taken from the Ahringer RNAi library [8]. Remaining vectors were made by inserting cDNA of its corresponding gene into L4440 and then retransformed into HT115(DE3).

RFK647 worms were synchronized at L1 larvae and seeded into RNAi plates containing induced bacteria. Worms were imaged in adulthood and harvested for RNA collection. Experiment was repeated three independent times.

RNA isolation

C. elegans were collected off plate and washed with M9 buffer(22mM KH_2PO_4, 42mM Na_2HPO_4, 85mM NaCl, 1mM $MgSO_4$) followed by a wash with ultrapure water and lysis in Worm Lysis Buffer (0,2M NaCl, 0,1M Tris pH=8,5, 50mM EDTA, 0,5% SDS) with 1mg/mL Proteinase K (Sigma-Aldrich, P2308) for 30 minutes at 65°C. After pelleting and removing debris, three volumes of TRIzol LS (ThermoFisher Scientific, 10296-028) were added to sample and RNA precipitation was carried out according to producer's instructions with the aid of Phase lock Gel – Heavy tubes (QuantaBio, 2302830). Eluted RNA samples were depleted of DNA using TURBO DNA-free Kit (Ambion, AM1907).

RNA Immunoprecipitation samples (see below) were obtained by adding TRIzol LS directly to IP beads after washes. The remaining isolation follows the previously described process.

RT-qPCR

Cultured worms and RNA samples were isolated as described above. Reverse transcription for each sample was performed with 500ng of total RNA using ProtoScript First Strand cDNA Synthesis Kit (NEB, E6300) and Oligo $d(T)_{23}VN$. qPCR 10µl reactions were set up with iTaq Universal SYBR Green Supermix (Bio-Rad, 1725121), 500mM primer concentration and a volume ratio of 1/5 cDNA. PCR cycles and measurements were made in an Applied Biosystems ViiA7 Real Time PCR System (ThermoFisher Scientific). Cycling conditions were made according to iTaq manufacture's recommendations: Standard run, temperature increments of 1,6°C/s; 95°C for 30 seconds, 40 cycles of 95°C for 15 seconds and 60°C for 1 minute; melt curve calculation: 15 seconds at 95°C, 1 minute at 60°C, temperature increments of 0,05°C/s to 95°C and hold for 15 seconds. Technical duplicates and biological triplicates were used. ΔΔCT method was used as an analysis method [9]. *pmp-3* was used as a normalization factor [10]. Error bars represent the standard deviation of three biological replicates. Used primers are listed below.

Target	**Sequence**
*pmp-3*_Fw	GTTCCCGTGTTCATCACTCAT
*pmp-3*_Rev	ACACCGTCGAGAAGCTGTAGA
*GFP*_Fw	ATGGTGTTCAATGCTTCTCG
*GFP*_Rev	TGACTTCAGcacgtgTCTTGT

Yeast two hybrid

Two-hybrid assays were performed in the haploid strain PJ69-4a and the pGAD and pGBD plasmid series as described previously [11]. Cell pinning was performed with Rotor HAD (Singer Instruments, ROT-001).

Temperature shift assay

RFK912 worms were cultured at 15°C in standard plates. At the start of experiment they were selected and singled into standard plates at L4 larvae stage. After overnight (O/N) culture at 15°C or 25°C individual worms were transferred into a new plate and shifted to corresponding temperature together with the plate of O/N egg lay. Every 2 hours individuals were transferred into a new plate. Eggs were counted in each of these plates on the day of egg lay and two days after larvae were counted for survival assay. As control we include RFK912 worms which underwent the same treatment except with no temperature shift and transferred into new plates every 4 hours.

Small RNA Library preparation and sequencing

NGS library prep was performed with NEXTflex Small RNA-Seq Kit V3 following Step A to Step G of Bioo Scientific`s standard protocol (V16.06). Libraries starting amount and PCR cycles can be consulted in the table below. Amplified libraries were purified by running an 8% TBE gel and size-selected for 18 – 40nt. Libraries were profiled in a High Sensitivity DNA on a 2100 Bioanalyzer (Agilent technologies) and quantified using the Qubit dsDNA HS Assay Kit, in a Qubit 2.0 Fluorometer (Life technologies). Samples of each individual experiment were pooled in equimolar ratio.

Experiment	Sample (each 3x)	Starting Material	PCR cycles	Equipment	Run type
pid-3	*pid-3(tm2417)*	1000ng	15	NextSeq 500 Flowcell	Highoutput 75-cycle-kit, SR for 1x 83 cycles plus 7 cycles for the index read
	pid-3(tm2417)/+	1000ng	15	NextSeq 500 Flowcell	Highoutput 75-cycle-kit, SR for 1x 83 cycles plus 7 cycles for the index read
ife-3	*ife-3(xf102)*	500ng	15	NextSeq 500/550 Flowcell	SR for 1x 75 cycles plus 7 cycles for the index read
	N2	500ng	15	NextSeq 500/550 Flowcell	SR for 1x 75 cycles plus 7 cycles for the index read
erh-2	*erh-2(xf168)*	500ng	15	NextSeq 500/550 Flowcell	SR for 1x 75 cycles plus 7 cycles for the index read
	erh-2(xf168)/+	500ng	15	NextSeq 500/550 Flowcell	SR for 1x 75 cycles plus 7 cycles for the index read

tost-1	*tost-1(xf194)*	2000ng	12	NextSeq 500 Flowcell	Midoutput 150-cycle-kit, PE for 2x 75 cycles plus 16 cycles for the index read
	pid-1(xf35)	2000ng	12	NextSeq 500 Flowcell	Midoutput 150-cycle-kit, PE for 2x 75 cycles plus 16 cycles for the index read
	N2	2000ng	12	NextSeq 500 Flowcell	Midoutput 150-cycle-kit, PE for 2x 75 cycles plus 16 cycles for the index read

Biochemistry

Worm preparation Synchronized non-gravid adult worms were collected off plate and washed with M9 buffer followed by a wash with ultrapure water. Pellets were frozen with liquid nitrogen and kept at -80^{o}C until usage.

Lysate preparation Worm Pellets were thawed on ice and mixed 1:1 with 2x Lysis Buffer (20 mM Tris.Cl, 300 mM NaCl, 1 mM EDTA, 1%(v/v) IGEPAL CO-630, pH 7,5) with 2x protease inhibitors (cOmplete Mini, EDTA-free, Roche, 11836170001). Bioruptor Plus (Diagenode) sonicator was used to lyse worms (10 cycles 30/30 seconds, 4^{o}C, high energy) and debris removed by spinning. Lysate protein concentration was determined with Pierce BCA Protein Assay Kit (ThermoFisher Scientific, 23225).

Immunoprecipitations Lysates were diluted in 1x Lysis Buffer+ 1x Protease inhibitors to a final concentration of 1,5 mg of protein/mL and a total of 0,75 mg of protein was used per IP. At this step input samples were collected into 2x NuPAGE LDS Sample buffer (Life Technologies, NP0007)+ 200 mM DTT and boiled for 10 minutes. Anti-mCherry IPs were performed with RFP-Trap_MA beads (Chromotek, rtma-20) and anti-GFP IPs with GFP-Trap_MA beads (Chromotek, gtma-20), in both cases 25 µl of bead slurry was used and samples were rotated at 4^{o}C for 2 hours. Subsequent washes were made with Wash Buffer (10 mM Tris.Cl, 150 mM NaCl, 0,5 mM EDTA, pH 7,5)+ Protease inhibitors in accordance with Chromotek protocols. Washed beads were resuspended in 2x NuPAGE LDS Sample Buffer + 200 mM DTT and boiled for 10 minutes, making the samples ready for loading.

RNAse treated immunoprecipitations followed the above described protocol with an additional RNAse A/T1 Mix (ThermoFisher Scientific, EN0551) treatment step. After lysate dilution samples were divided in two (Control and +RNAse) and 20 µL of RNAse A/T1 mix was added per 1 mL of +RNase sample. Control and +RNase samples were rotated for 20 minutes at 4^{o}C and followed by the described IP protocol.

Western blot Inputs an IP samples were loaded into 4-12% gradient gels (ThermoFisher, NP0321BOX) and run with 1x NuPAGE MES SDS Running Buffer (ThermoFisher, NP0002). Transfer to an Immobilon PVDF, 0,45 µm membrane (Merck Millipore, IPVH00010) was executed with 1x NuPAGE Transfer Buffer (ThermoFisher Scientific, NP0006) 20%(v/v) Methanol. Membrane was probed with rabbit anti-PID-1 Q5941 [12] and detected with Amersham ECL Select Western Blotting Detection Reagent (GE Healthcare, RPN2235). Background recognition by anti-PID-1 ab is used as loading control.

Endogenous PID-1 Immunoprecipitations 200 µL of synchronized adult worms were resuspended in 500 µL of IP lysis buffer (25 mM Tris pH 7,5, 150 mM NaCl, 1,5 mM $MgCl_2$, 1 mM DTT, 0,1% Triton X-100, complemented with 2x protease inhibitor) and sonicated at 4°C for 10 cycles of 30/30 seconds, high intensity using a Bioruptor Plus (Diagenode). Cell debris was removed via spinning and 30 µL of washed Dynabeads Protein G (Life Technologies, 1004D) and 10 µL of anti-PID-1 antibody (Q5941) was added to the lysates and incubated under rotation for 3 hours at 4°C. The beads were then washed 3x 5 minutes in wash buffer (25 mM Tris pH 7,5, 150 mM NaCl, 1,5 mM $MgCl_2$, 1 mM DTT, complemented with 2x protease inhibitor) and resuspended in 30 µL of NuPAGE LDS buffer.

Mass Spectrometry Samples were separated on a 4–12% gradient Bis-Tris gel (ThermoFisher, NP0321) in MOPS SDS Running Buffer (ThermoFisher, NP0001) at 180 V for 10 minutes, afterward separately processed by in-gel digest [13] [14] and desalted using a C18 StageTip [15]. The digested peptides were separated on a 25cm reverse-phase capillary (75µm inner diameter) packed with Reprosil C18 material (Dr. Maisch). Separation of the peptides was done with the EASYnLC 1000 system (Thermo) along a 2 hour gradient increasing from 2 to 40% Buffer B. For PID-1 IPs the gradient was shortened to 90 minutes. Measurement was done on a Q Exactive Plus mass spectrometer (Thermo) operated with a Top10 data-dependent MS/MS acquisition method per full scan [16]. Measurements were processed with MaxQuant version 1.5.2.8 [17] using the wormbase protein fasta database (version WS265) and standard settings except LFQ quantitation and match between runs were activated. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD011500.

RIPseq

Lysate preparation Worm Pellets were thawed on ice and mixed 1:1 with 2x Lysis Buffer (20 mM Tris.Cl, 300 mM NaCl, 1 mM EDTA, 1%(v/v) IGEPAL CO-630, pH 7,5) with 2x protease inhibitors (cOmplete Mini, EDTA-free, Roche, 11836170001) and 2x SUPERase.In RNase Inhibitor (Ambion, AM2696). Bioruptor Plus (Diagenode) sonicator was used to lyse worms (10 cycles 30/30 seconds, high energy). Lysate protein concentration was determined with Pierce BCA Protein Assay Kit (ThermoFisher Scientific, 23225). Lysates were diluted in 1x Lysis Buffer+ 1x Protease inhibitors+ 1x SUPERase.In RNase Inhibitor to a final concentration of 1,5 mg of protein/mL and a total of 2,1 mg of protein was used per IP. Each lysate was cleared with 225 µL of Binding Control magnetic agarose beads (Chromotek, bmab-20) for 1 hour at 4°C.

Immunoprecipitation Quadruplicate anti-mCherry RIPs were performed with RFP-Trap_MA beads (Chromotek, rtma-20). 75 µl of bead slurry per sample blocked for 1 hour with Blocking Buffer [2% (w/v) BSA, 2,5 mg/mL tRNA from *E.coli* MRE 600 (SigmaAldrich, 10109541001), 10 mM Tris.Cl, 150 mM NaCl, 0,5 mM EDTA, pH 7,5] and washed with Wash Buffer (10 mM Tris.Cl, 150 mM NaCl, 0,5 mM EDTA, pH 7,5). Inputs from cleared lysates were taken and mixed 3:1 with TRIzol. 75 µl of blocked bead slurry was added to the

remaining cleared lysate and samples were rotated at 4°C for 2 hours. Subsequent washes were made with Wash Buffer+ Protease inhibitors. Washed beads were resuspended in 100 µL of Nuclease free water and immediately mixed with 400 µL of TRIzol.

RppH treatment A sample of each RNA sample was collected and treated with RNA 5' Pyrophosphohydrolase (RppH) (NEB, M0356) for the purpose of removing 5'Cap structures [18]. Each was treated in ThermoPol Buffer (20 mM Tris.Cl, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0,1% Triton X-100, NEB, B9005) for 1 hour at 37°C with 10 units of RppH. Reaction was stopped by adding EDTA to 10 mM and heating to 65°C for 5 minutes. RNA was purified by ethanol precipitation.

Library preparation and Sequencing NGS library prep was performed with NEXTflex Small RNA-Seq Kit V3 following Step A to Step G of Bioo Scientific`s standard protocol (V16.06). Libraries were prepared with a starting amount of 100 ng and amplified in 18 PCR cycles. Amplified libraries were purified by running an 8% TBE gel and size-selected for 18 – 40nt. Libraries were profiled in a High Sensitivity DNA on a 2100 Bioanalyzer (Agilent technologies) and quantified using the Qubit dsDNA HS Assay Kit, in a Qubit 2.0 Fluorometer (Life technologies). Total amount of samples were divided in two pools. Each pool was mixed in equimolar ratio and sequenced on a NextSeq 500/550 Flowcell, SR for 1x 75 cycles plus 7 cycles for the index read.

Bioinformatic analysis

sRNA sequencing analysis Raw reads were first processed to remove adapters with v1.9 cutadapt (https://cutadapt.readthedocs.io/en/stable/) (seqtk trimfq -L 50 | cutadapt -a TGGAATTCTCGGGTGCCAAGG -O 5 -m 26 -M 48), followed by removal of reads containing low quality calls with the FASTX-Toolkit v0.0.14 (fastq_quality_filter -q 20 -p 100 -Q 33). The information of the read sequence and the 5' and 3' random UMIs (NNNN-RNA sequence-NNNN) was then used to collapse reads with identical sequences, including that of the UMIs using a command-line script. UMIs were then removed (seqtk trimfq -b 4 -e 4), reads shorter than 15 nucleotides were removed (seqtk seq -L 15) before mapped against the *C. elegans* genome (WBcel235, ensembl) with bowtie v0.12.8 [19] (-q –sam –phred33-quals –tryhard –best –strata -v 0 -M 1). Coverage tracks were generated with Bedtools 2.25.0 [20] (genomeCoverageBed -bg -split -scale) to summarize genomic read coverage, and bigwigs created with bedGraphToBigWig. Normalization was done to total mapped reads. For visualization, the alignments were merged with bamtools-2.3.0 merge [21]. For the RIP-seq experiments, merged alignments were further processed to create $\log_2$(IP/input) normalized tracks using DeepTools [22] (bigwigCompare –binSize 1 –ratio log2).

To identify RNA-bound to the complex in RIPseq, we used DESeq2 [23] with the formula '~replicate+condition' in which each IP is being compared to the corresponding input control. Replicates are paired as they are generated from the same biological sample. All the samples in the dataset were included in construction of the DESeq2 object, in order to estimate the dispersion more robustly.

The number of reads mapping to different RNA classes was estimated with a combination of a custom Python script to select reads by size and nucleotide bias, available at https://github.com/adomingues/filterReads/blob/master/filterReads/filterSmallRNAclasses.py, and bedtools intersect to match reads with annotated features. 21U RNAs were defined as reads with 18-40 bases mapping sense to an annotated 21U RNA *locus* (intersectBed -s -f 0.85). 22G RNAs are all reads 20-23 bases long mapping antisense to protein coding genes, pseudogenes, lincRNA and transposons. 26G RNAs are 26 nucleotide long reads mapping antisense to protein coding genes, pseudogenes and lincRNA. For 22G and 26G RNAs a minimum overlap of 1 base was required (intersectBed default). miRNAs were defined as reads mapping sense to annotated miRNAs (intersectBed -s -f 1.0). The definition of 21U RNAs in particular was kept loose to allow for the identification of 21Us which are not fully mature. Gene locations were extracted from a custom gtf (genes + transposons) using the biotype information.

Metagene profiles were created with DeepTools. Read coverage was summarized with computeMatrix scale-regions –metagene –missingDataAsZero -b 50 -a 50 –regionBodyLength 98 –binSize 1 –averageTypeBins mean. As SL genes are multicopy genes, the setting –averageTypeBins was set to “sum”. Using the “mean” of SL sequences did not alter the profile obtained (data no shown). The final metagene figure was created with plotProfile –plotType lines –perGroup.

Supplemental figure 1 - PID-1 interacts with a restricted set of proteins

A) Biological replicate of interaction data as described in Figure 1A.

B) mCherry pull-down of wild type (WT), *pid-1(xf14)* mutant and PID-3::mCherry::Myc;*pid-3(tm2417)* (PID-3::mCh) carrying worms. IPs were performed in non-gravid adult extracts. Membrane was probed for endogenous PID-1. Background recognition by the ✔PID-1 antibody is used as loading control.

C) Schematic representation of domain composition of PETISCO components used in the Y2H grid. In all cases proteins were fused to the C-terminal part of budding yeasts' GAL4 activation or binding domains.

Supplemental Figure 2 – PID-1 interactor miniMos transgenes

A) Schematic representation of miniMos transgene insertions for each of the PID-1 interactors. *C. briggsae unc-119* is used as selection marker.

B) Expression pattern of 3xFLAG::mCherry::IFE-3 under endogenous promotor and 3'UTR in germline and embryos. PGL-1::GFP is used as a P-granule marker. Images depict live worms under the wide field fluorescent microscope. Circle and arrow highlight PGL-1:IFE-3 co-localization. Scale Bar represents 10 µm. Contrast of images has been enhanced.

C) Expression pattern of miniMos transgenes of PID-1 and its interactors under respective endogenous promotors and 3'UTR. Images depict live worms under the wide field fluorescent microscope. Scale Bar represents 20 µm. Contrast of images has been enhanced.

Supplemental Figure 3 – PETISCO is a stably interacting complex

a-e) Volcano plots representing label-free proteomic quantification of quadruplicate anti-mCherry IPs from non-gravid adult extracts. In **a)** and **b),** respectively, PID-3::mCherry::Myc;*pid-3(tm2417)*; or 3xFLAG::mCherry::IFE-3;*ife-3(xf101)*; were precipitated with or without RNAse A/T1 treatment. Worms were grown in high density plates for these particular experiments. **c-e)** Independent IP-LFQP experiments of **c)** PID-3::mCherry::Myc;*pid-3(tm2417)*; **d)** PID-3::mCherry::Myc;*pid-1(xf14)*; and **d)** 3xFLAG::mCherry::IFE-3;*ife-3(xf101).* In all cases the x-axis represents the median fold enrichment of individual proteins in control (WT) versus transgenic strain. y-axis indicates $-\text{Log}_{10}$(p-value) of observed enrichments. Dashed lines represent thresholds at p=0.05 and 2-fold enrichment. Blue data points represent values out of scale. Red and Green data points represent above and below threshold respectively.

Supplemental Figure 4 – Yeast Two Hybrid interaction assays of PETISCO

b-e) Yeast two-hybrid interaction assays of PETISCO subunits in low stringency (TRP^- LEU^-HIS^-), high stringency ($TRP^-LEU^-HIS^-ADE^-$) or control (TRP^-LEU^-) plates as indicated. **b)** Full length proteins **c)** TOFU-6 and individual domains tested for interaction with

full length IFE-3 **d)** Interactions between PID-3 and TOFU-6 **e)** Interactions with ERH-1.

Supplemental Figure 5 – PETISCO mutant small RNA sequencing and gene structure.

A) Schematic representation of individual gene structures of PETISCO components. Alleles are indicated within these schemes.

b-d) Global levels of **b)** type II 21U RNAs **c)** 26G RNAs and **d)** 22G RNAs in wild type (N2), *ife-3(xf102)*, *pid-3(tm2417)* and *erh-2(xf168)* worms. For the latter two heterozygous populations were used as a control. Values are in reads per million (RPM). Individual data points of three independent replicates are shown and horizontal bar represents the total mean. Significance was tested with Student's t-test and p-values are indicated in the graph.

Supplemental Figure 6 – Masculinization of the germline in *pid-1(xf35)* mutant animals

Fluorescent microscopy of the germline of 21U sensor GFP of *pid-1(xf35)*; 21U sensor worms. *pid-1(xf35)* hermaphrodites display a low frequency (<1%) Mog phenotype (second row). In the bottom row a male germline from the same strain is presented and in the top row a hermaphrodite germline of a *pid-1* knock-down. Scale Bars represent 20 µm. Contrast of images has been enhanced.

Supplemental Figure 7 – TOST-1 and PID-1 alignment and effects on small RNA populations

A) Protein alignment of nematode PID-1 and TOST-1 orthologues. *C. elegans* orthologues are underlined in blue. Conserved motif is shaded in pink.

B) Temperature shift assays of the temperature sensitive allele *tost-1(xf196)*. L4 larvae were grown overnight (O/N) at 15°C or 25°C and shifted to 25°C or 15°C the next morning already as adults. O/N egg lay was shifted to corresponding temperature in parallel. Each individual was changed into a new plate every 2h and progeny counted at egg stage and L2 larvae. Each point represents the progeny of an individual worm.

C) Global levels of type II 21U RNAs, 26G RNAs and 22G RNAs in wild type (N2), *pid-1(xf35)* and *tost-1(xf194)* gravid adult worms. Values are in reads per million (RPM). Each dot represents a replicate of three and horizontal bar represents the total mean. Significance tested with Student's t-test and p-values are indicated in the graph.

Supplemental Figure 8 – PETISCO and snRNAs

a) Read coverage of SL1 RNA and 5S rRNA genes in wild type (N2), *pid-1(xf35)* and *tost-1(xf194)* gravid adult worms and *pid-3(tm2417), erh-2(xf168)* and *ife-3(xf102)* non-gravid adults. Line represents the average of three replicates. Colors under SL1 RNA correspond to scaled colors represented in Figure 6A.

b) Global levels of SL2 RNAs in wild type (N2), *pid-1(xf35)* and *tost-1(xf194)* gravid adult worms and *pid-3(tm2417), erh-2(xf168)* and *ife-3(xf102)* non-gravid adults. Values

are in reads per million (RPM). Individual data points of three independent replicates are shown and horizontal bar represents the total mean. Significance was tested with Student's t-test and p-values are indicated in the graph.

c) Fold enrichment of SL2 RNAs in Mock (N2), 3xFLAG::mCherry::IFE-3;*ife-3(xf101)*; and PID-3::mCherry::Myc;*pid-3(tm2417)* RIPs over paired input in non-gravid adult worms. Individual data points of four independent replicates are shown and horizontal bar represents the total mean. Significance was tested with Student's t-test and p-values are indicated in the graph.

d) Coverage profile, normalized to paired input, of SL2 in the indicated strains.

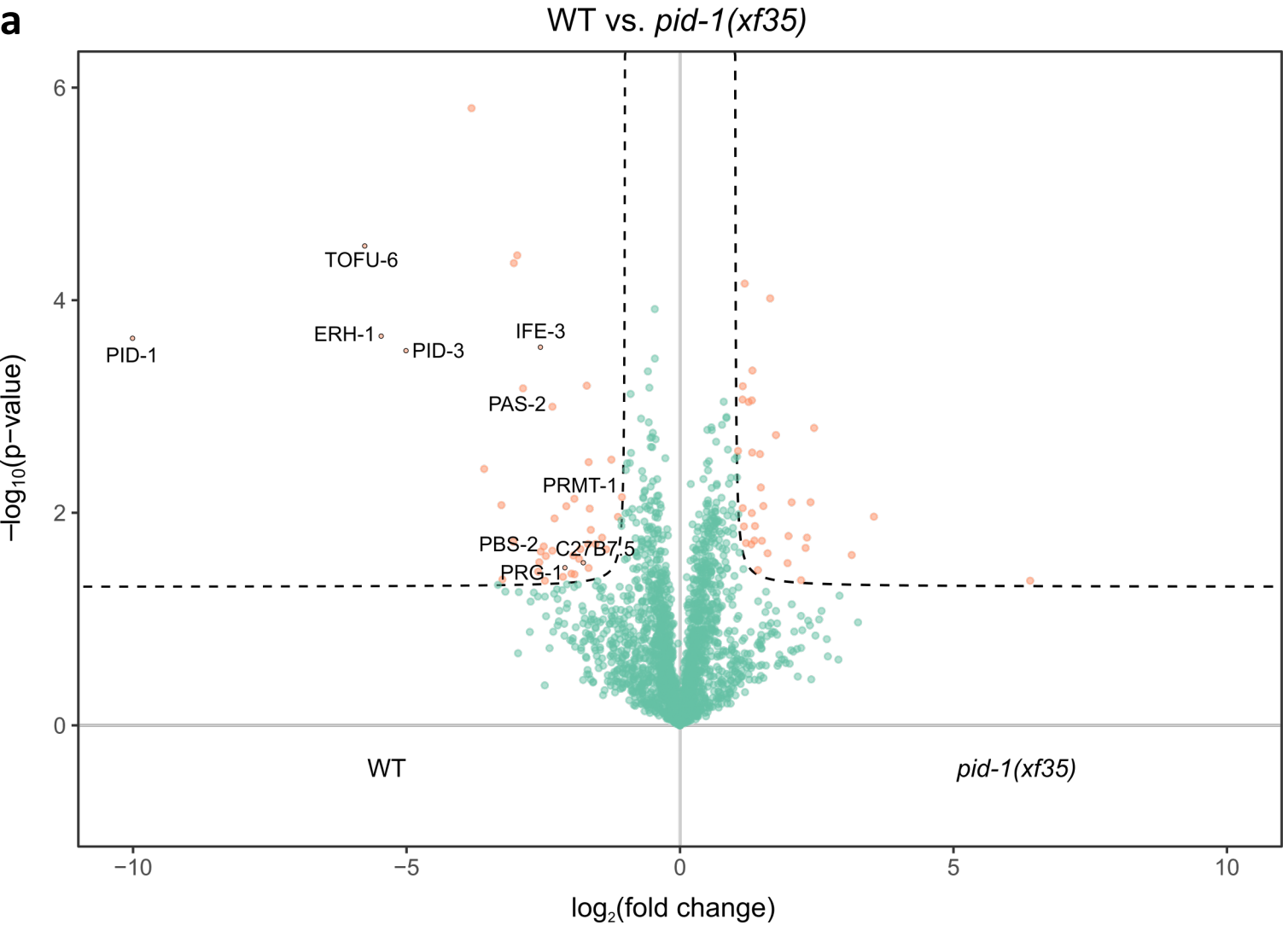
e) Violin plots showing the enrichments of 21U RNA genes and U snRNAs genes in Mock (N2), 3xFLAG::mCherry::IFE-3;*ife-3(xf101)*; and PID-3::mCherry::Myc;*pid-3(tm2417)* RIPs over paired input in non-gravid adult worms. Left panel: non-treated RNA samples; right panel: RppH treated samples. Significance was tested with a two-sided unpaired Mann-Whitney/Wilcoxon rank-sum test.

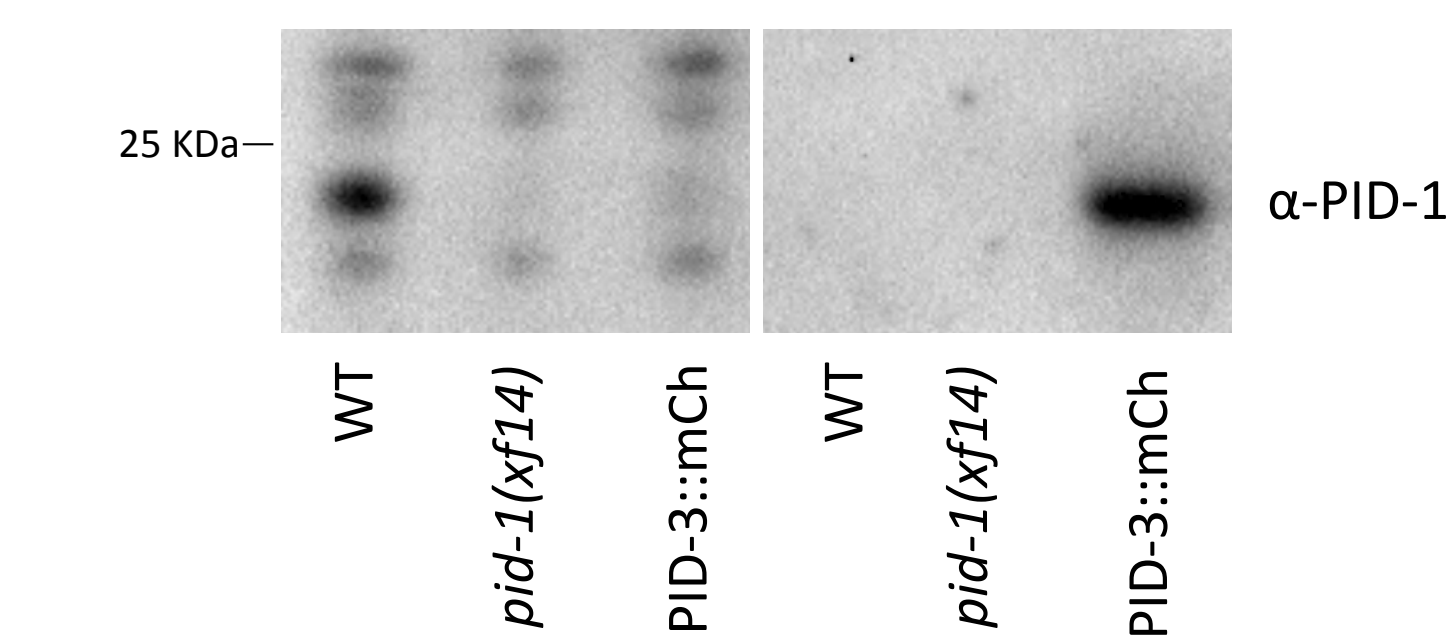
References

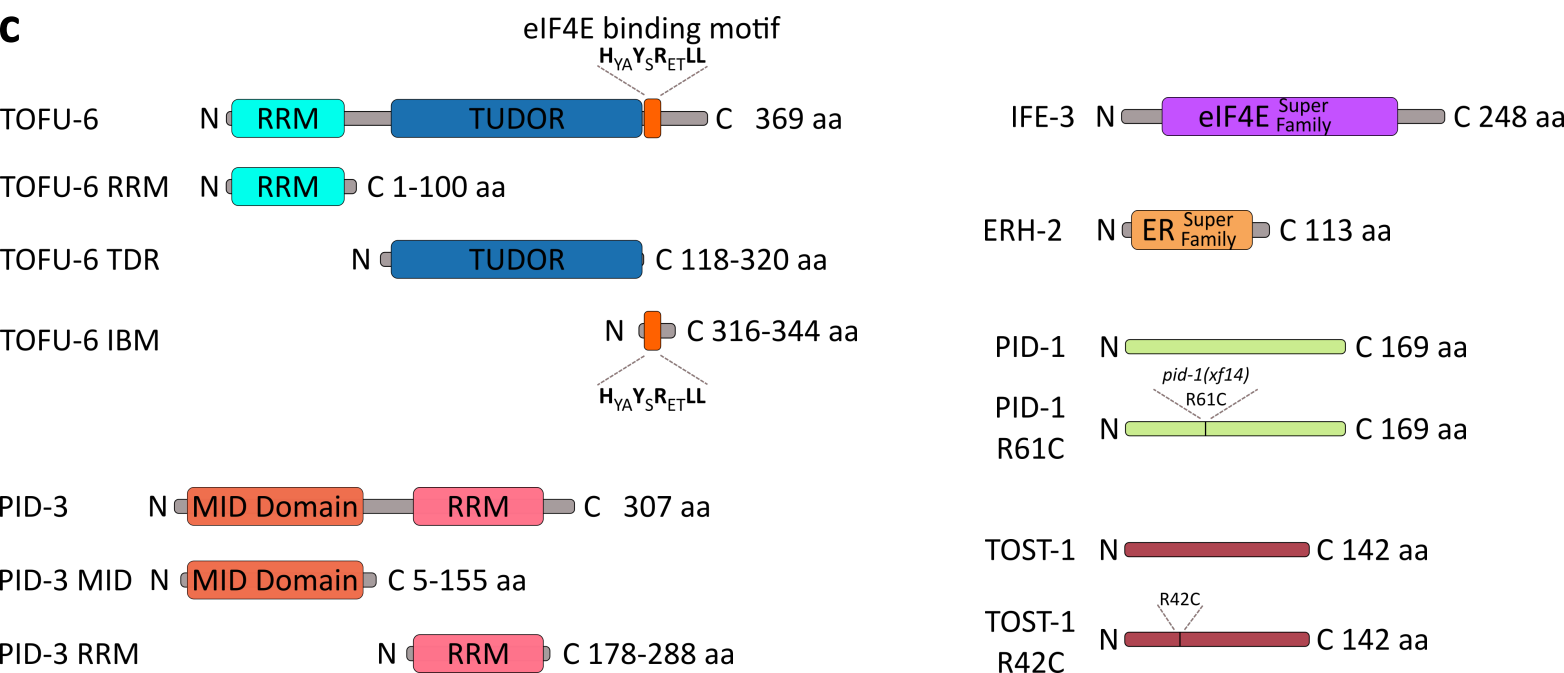
1. Brenner, S. The genetics of Caenorhabditis elegans. *Genetics* **77**, 71-94 (1974).
2. Schweinsberg, P.J. & Grant, B.D. C. elegans gene transformation by microparticle bombardment. *WormBook*, 1-10 (2013).
3. Friedland, A.E. et al. Heritable genome editing in C. elegans via a CRISPR-Cas9 system. *Nat Methods* **10**, 741-3 (2013).
4. Naito, Y., Hino, K., Bono, H. & Ui-Tei, K. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics* **31**, 1120-3 (2015).
5. Frokjaer-Jensen, C. et al. Random and targeted transgene insertion in Caenorhabditis elegans using a modified Mos1 transposon. *Nat Methods* **11**, 529-34 (2014).
6. Duerr, J.S. Antibody staining in C. elegans using "freeze-cracking". *J Vis Exp* (2013).
7. Timmons, L. & Fire, A. Specific interference by ingested dsRNA. *Nature* **395**, 854 (1998).
8. Kamath, R.S. et al. Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. *Nature* **421**, 231-7 (2003).
9. Schmittgen, T.D. & Livak, K.J. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* **3**, 1101-8 (2008).
10. Hoogewijs, D., Houthoofd, K., Matthijssens, F., Vandesompele, J. & Vanfleteren, J.R. Selection and validation of a set of reliable reference genes for quantitative sod gene expression analysis in C. elegans. *BMC Mol Biol* **9**, 9 (2008).
11. James, P., Halladay, J. & Craig, E.A. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**, 1425-36 (1996).
12. de Albuquerque, B.F. et al. PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. *Genes Dev* **28**, 683-8 (2014).
13. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* **1**, 2856-60 (2006).
14. Kappei, D. et al. HOT1 is a mammalian direct telomere repeat-binding protein contributing to telomerase recruitment. *EMBO J* **32**, 1681-701 (2013).
15. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**, 1896-906 (2007).
16. Bluhm, A., Casas-Vila, N., Scheibe, M. & Butter, F. Reader interactome of epigenetic histone marks in birds. *Proteomics* **16**, 427-36 (2016).
17. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-72 (2008).
18. Neri, F. et al. Intragenic DNA methylation prevents spurious transcription initiation. *Nature* **543**, 72-77 (2017).
19. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).
20. Quinlan, A.R. & Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-2 (2010).
21. Barnett, D.W., Garrison, E.K., Quinlan, A.R., Stromberg, M.P. & Marth, G.T. BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* **27**, 1691-2 (2011).

22. Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-5 (2016).
23. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

Supplemental Figure 1

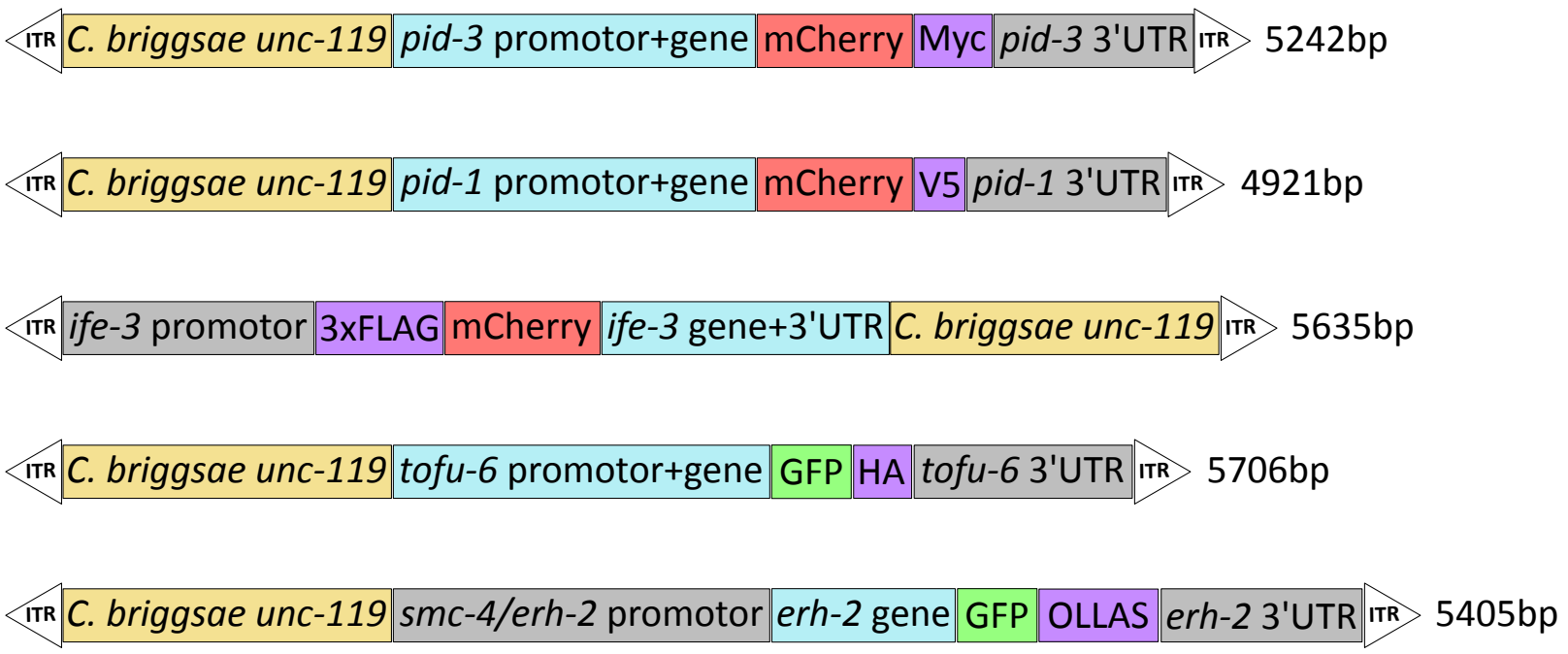




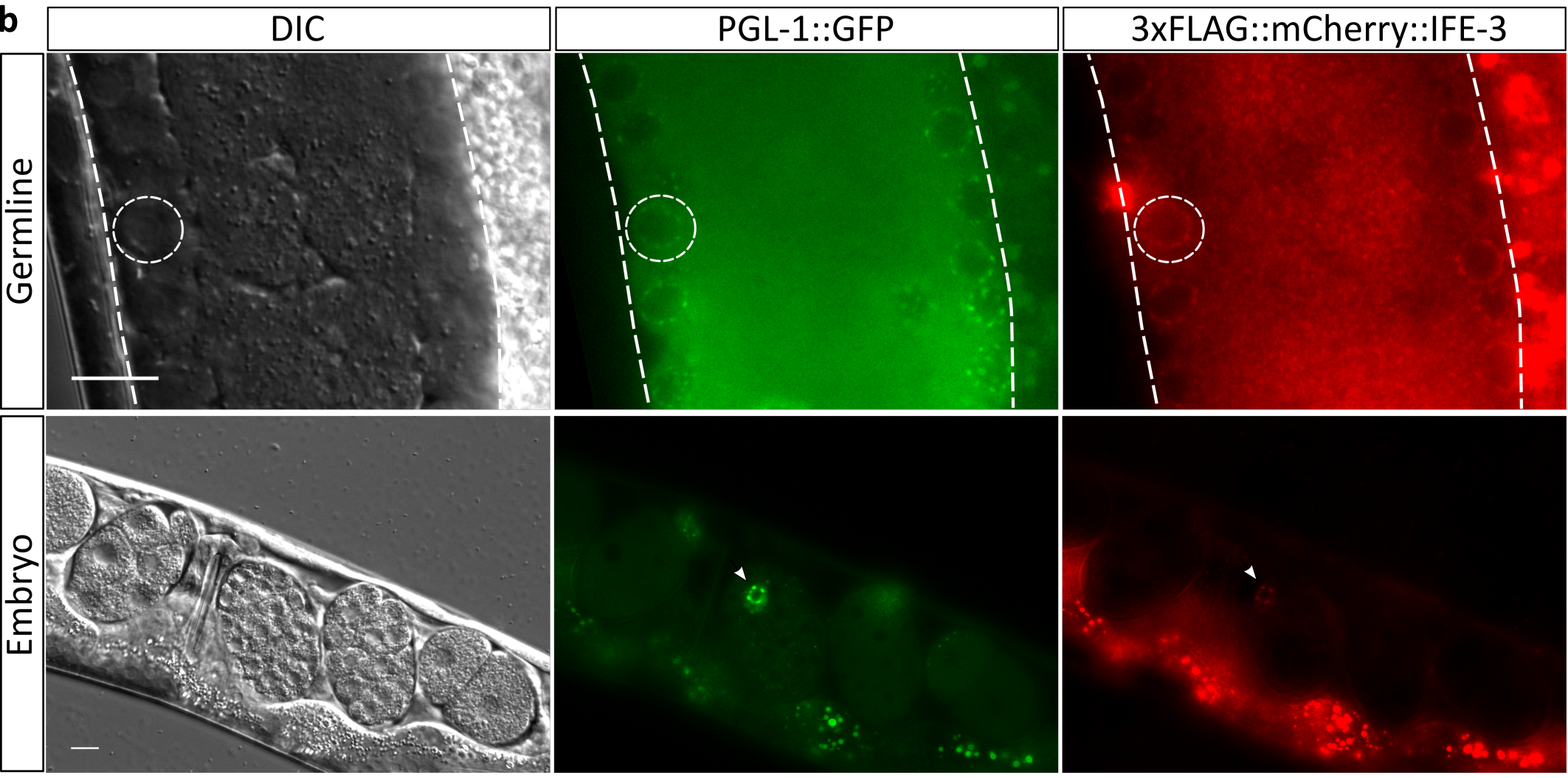


Supplemental Figure 2

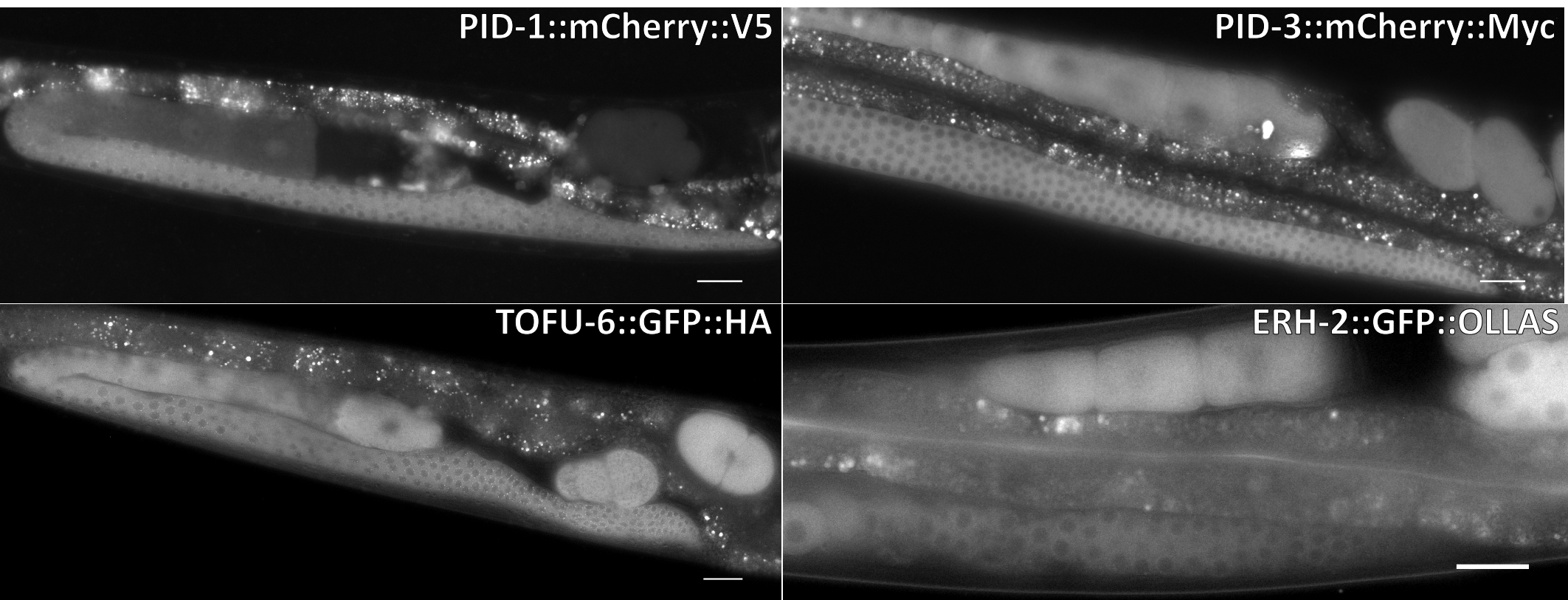
a



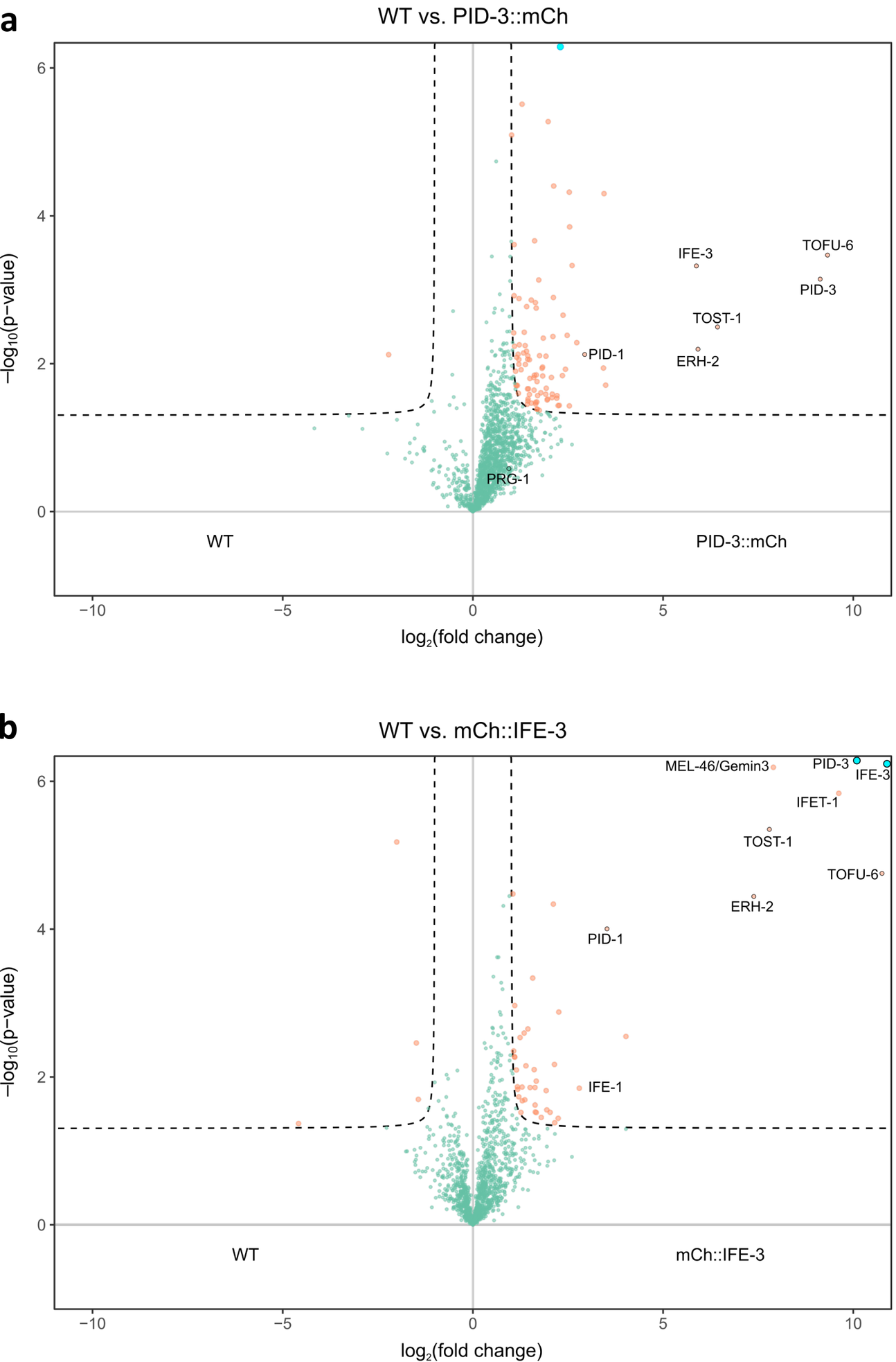
b



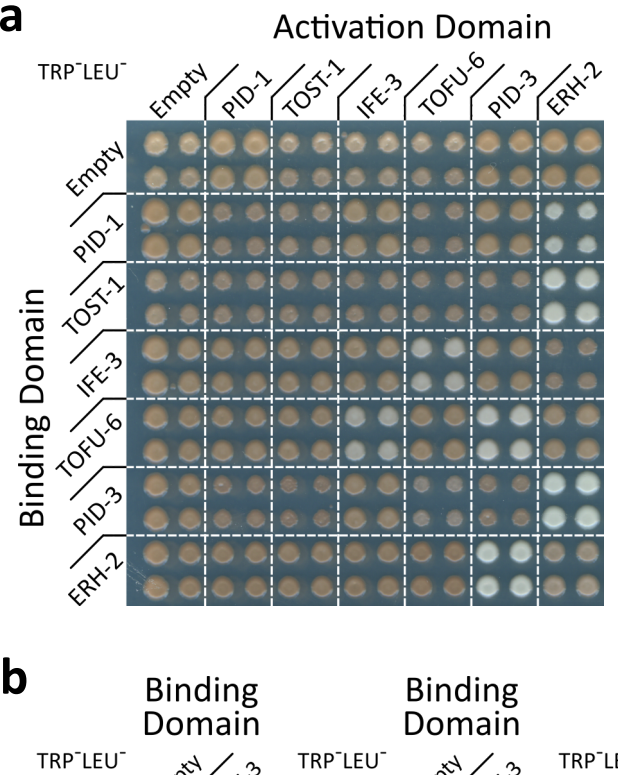
c

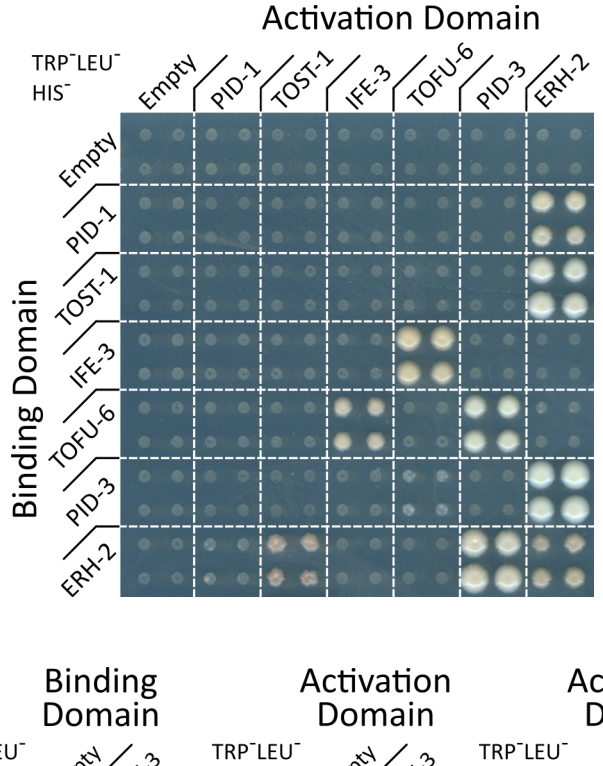


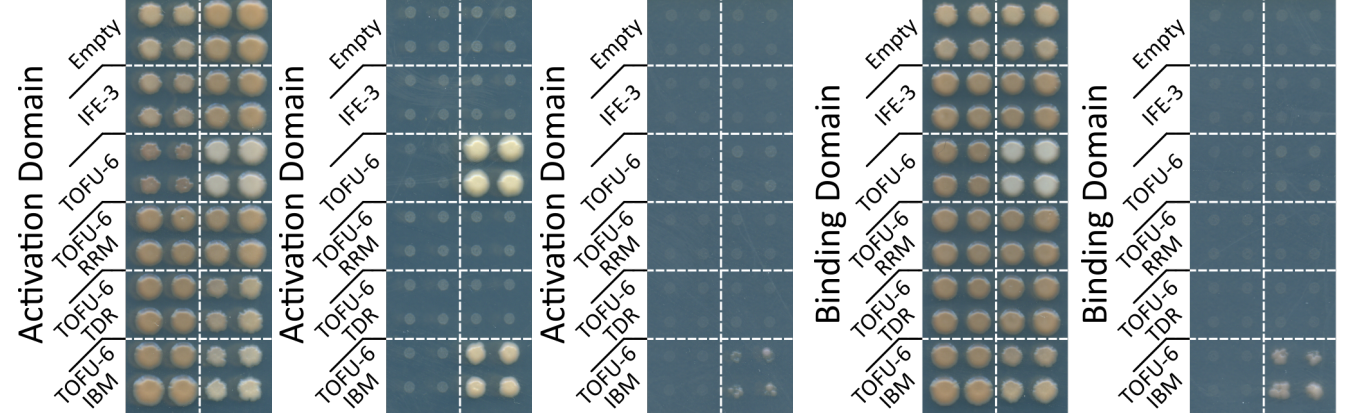
Supplemental Figure 3

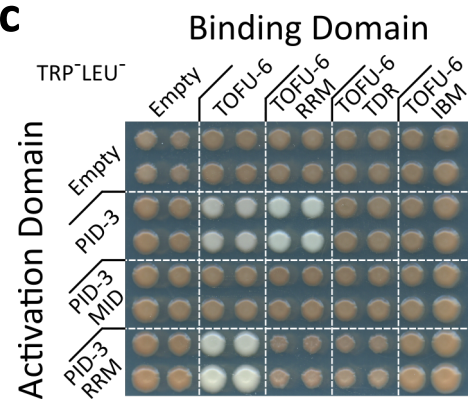


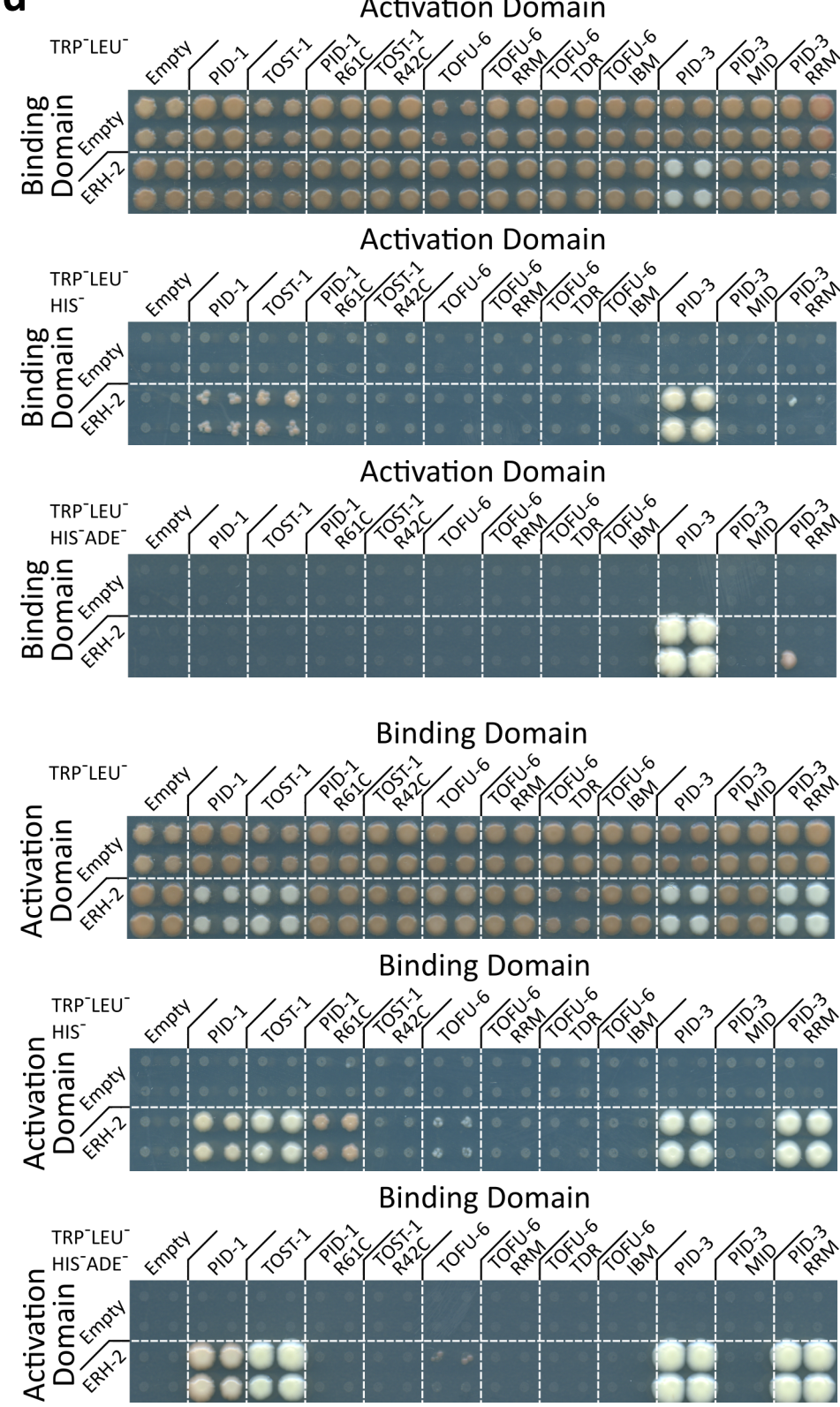
Supplemental Figure 4



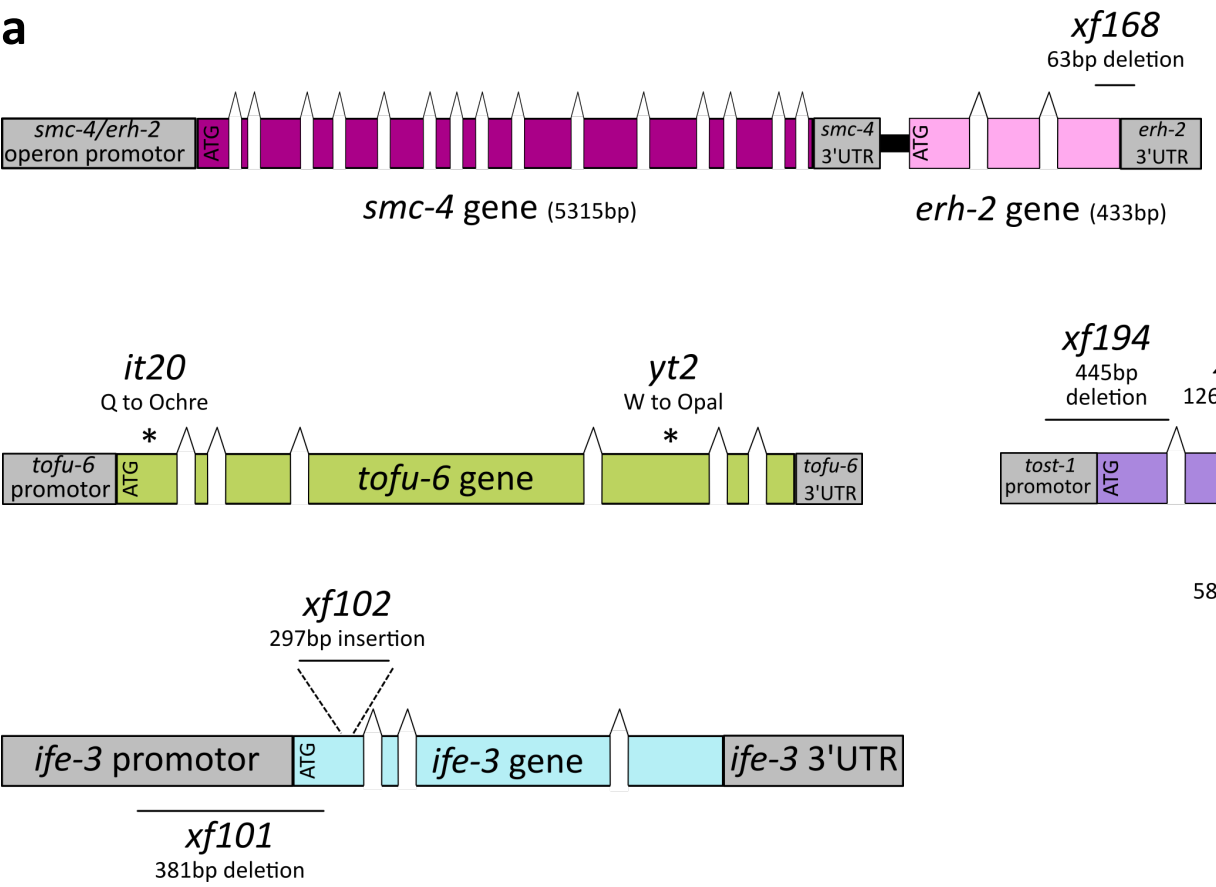


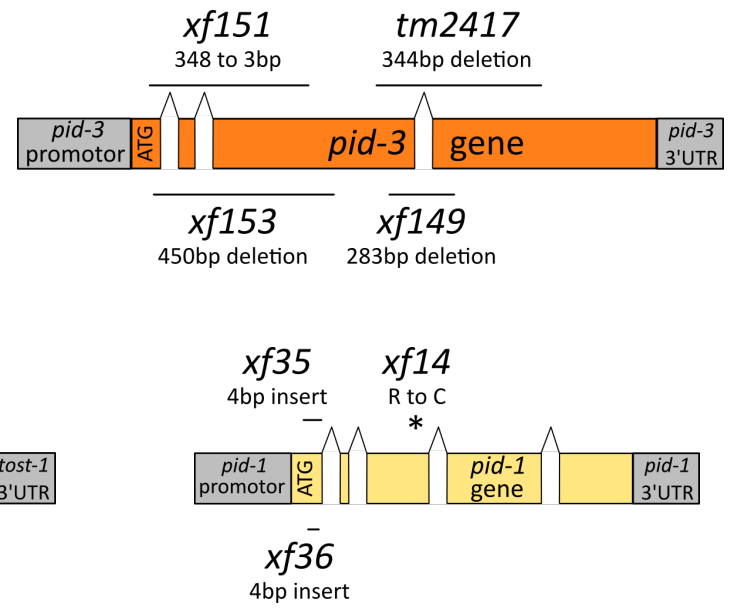


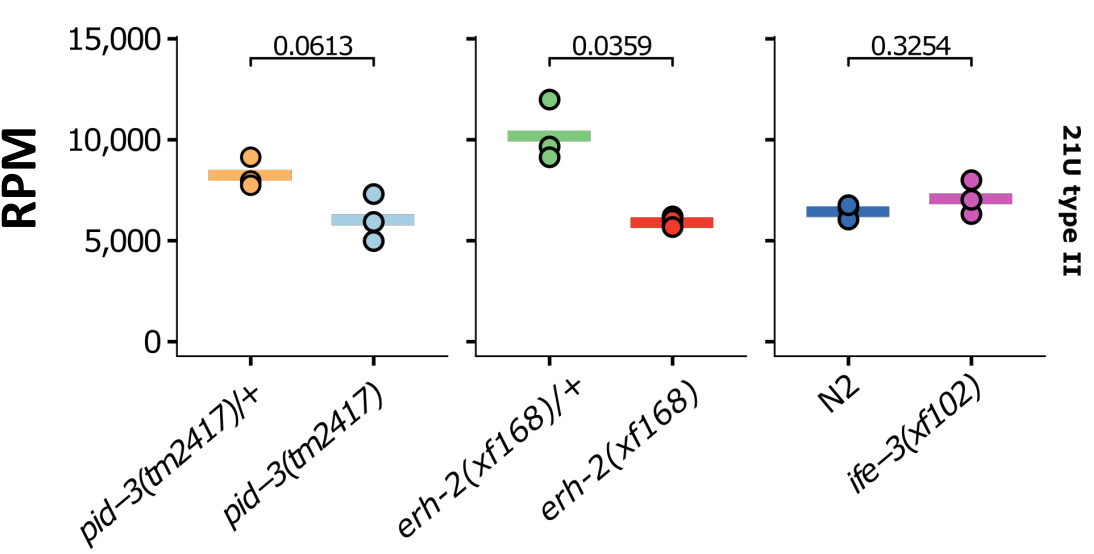


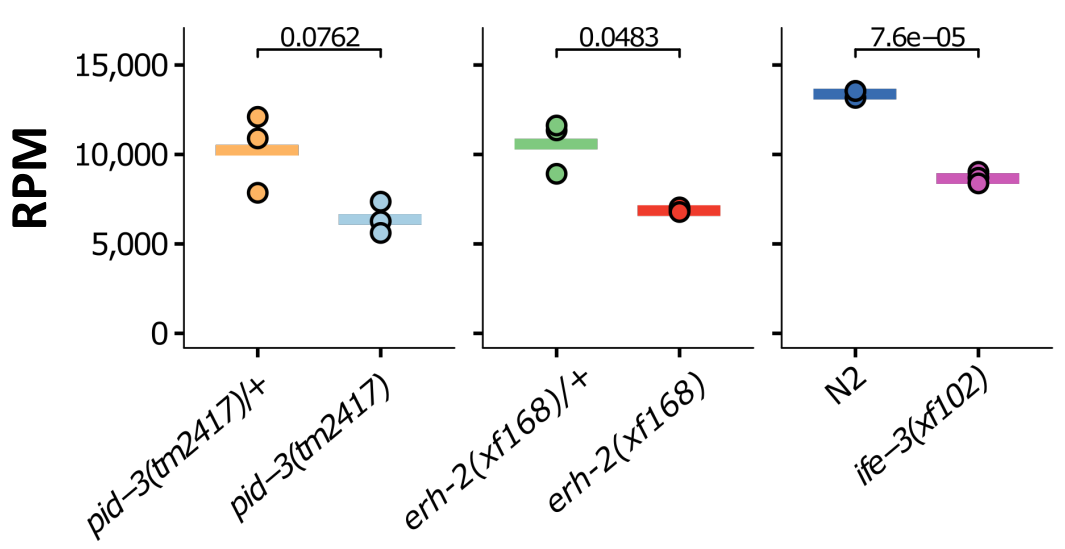


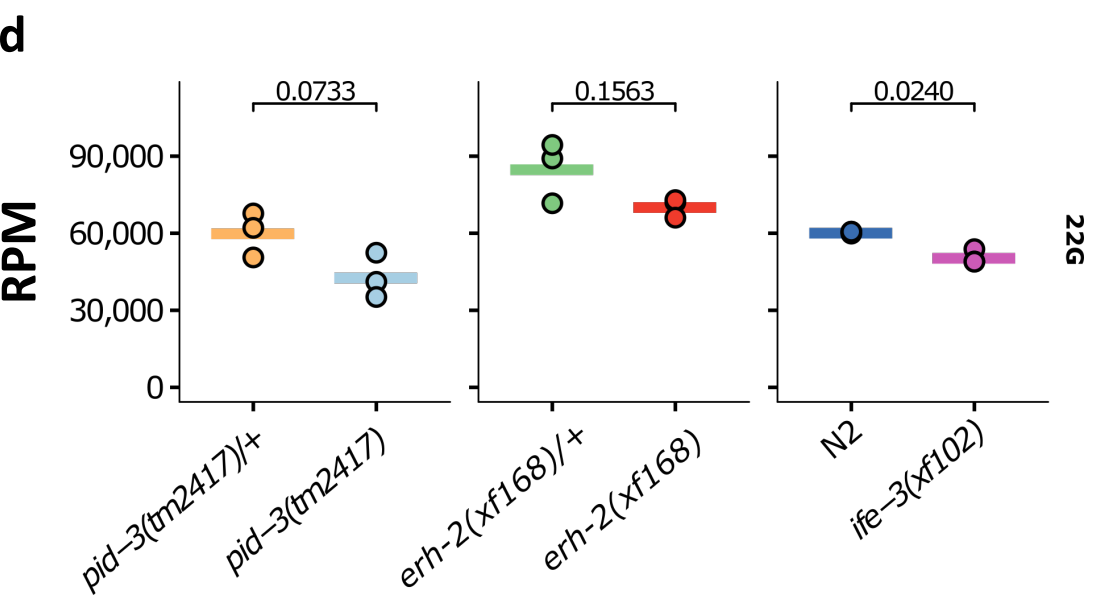
Supplemental Figure 5



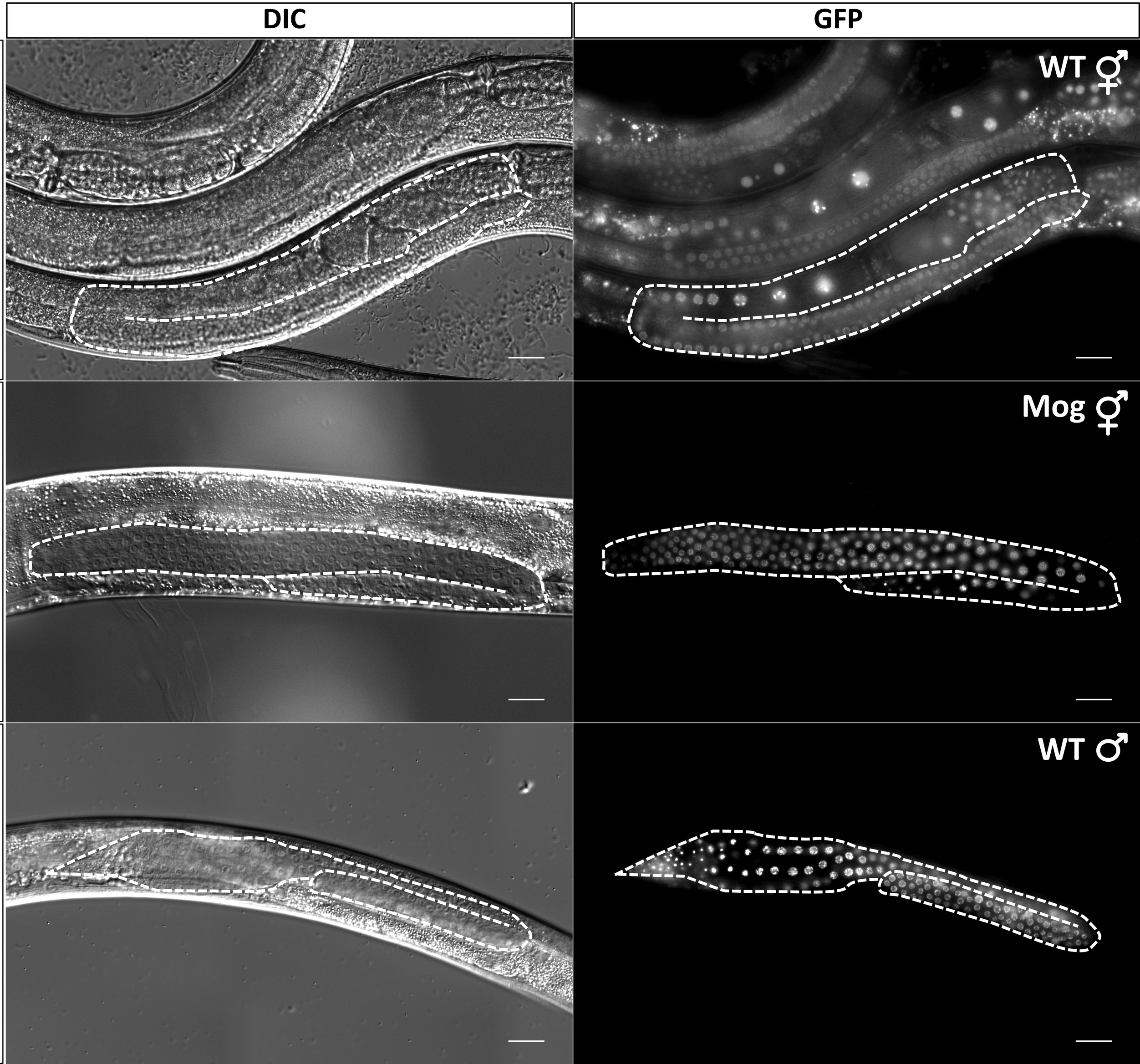








Supplemental Figure 6

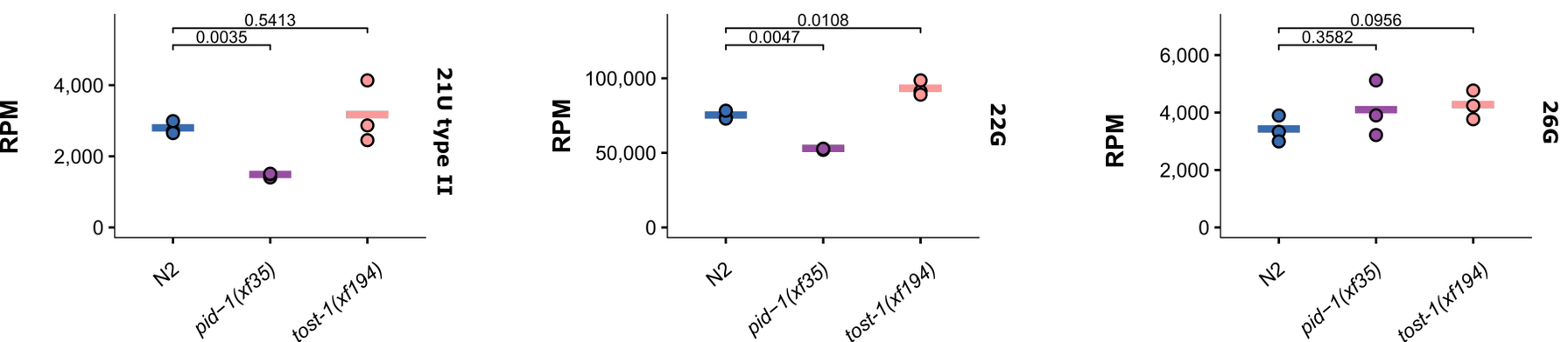


Supplemental Figure 7

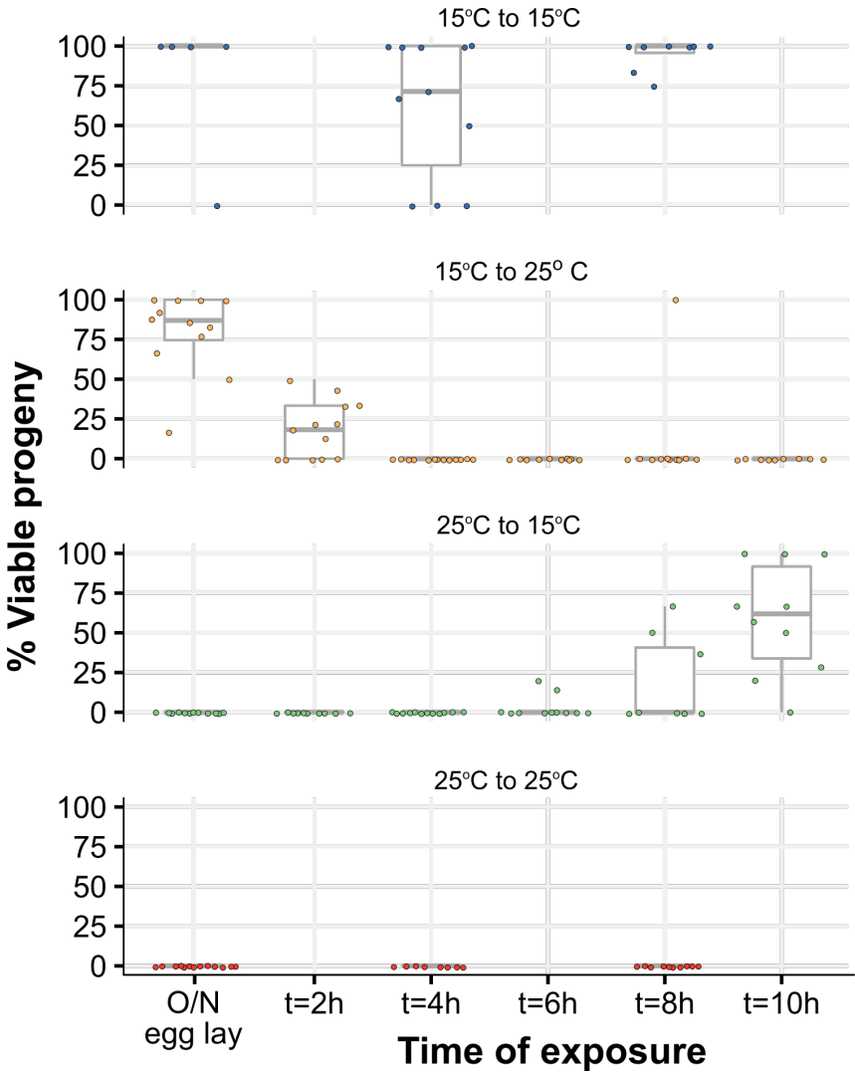
a

```
CLUSTAL multiple sequence alignment by MUSCLE (3.8)


CjpPID-1          MSAAREFSHITLVSTPFKAKIDTDARKSKLEIEKDLNTACKCADRFNYNANLHRKVTLSD
CePID-1           MSAKREFSHITLASTPFKKRIDQNSLKTDSDIEKDTNIAHKCAERFNYNTNLHRKVTLSD
CbrPID-1          MSSFREFSHITLSTTPFKVRLDQNAMKTDSDIEKDKNMAYKCAERFNYNTNLHRKVTLSD
CrePID-1          MSADREFSHITLSSTPFKYRMDQQAKKNDSDVEKDRNIAIKCAERFNYNTNLLRKVTLSD
CbnPID-1          MSAHREFSHITLSSTPFKMRMDQEAMKSETDIEKDRNIAYKCADRFNYNTNLLRKVSLSD
PpaTOST-1         MTTSSQECSPAVSPVPQIRSFPHRIIRPLA--DKEAIQRNFVTLPNRAANFSNRNLTLNE
CjaTOST-1         -MCGVQYEVVALN-------------------GRQHNQKIEAILEAKYEEACSRKITLNE
CjaTOST-2         -MCGVQYQVVALN-------------------GRQHNQKIEAILEAKYEEACSKKITLNE
CeTOST-1          MTTAIQYQVIPLA-------------------DREHSNKVASILENKYKTATNKRITLNE
CreTOST-1         MTSGIQYQIIPLA-------------------DREHNNKVASLLETKYKTAIQKRVTLNE
CbgTOST-1         MASNVQYEILSLV-------------------EGEKNNKVVSILENKYKSPA-KRVTLNE
CbnTOST-1         MATAVQYRIAELP-------------------KGDHSNKVASILKNKYEHSSKKRITLNE
                       :     :                     :           .      ..::*.:

CjpPID-1          RFELAAFGFEMKAKPNVSGNDLVQTAPD-YIYRKERRQESGINSPHTAGMSPQNPLPAGR
CePID-1           RFELAALGYEMKAKPRTIIEKHNDCDEFHFIYRKEKKNDYGTGSPLSAGLSLSNPLPAGR
CbrPID-1          RFELAAFGFEMKAKSH-TPVDTKDEDTLDYIYRKEKKQDY-AASPLPAGMSPTNPLPAGR
CrePID-1          RFELAALGFEMKSKPSSYLDKNQESEEFDYIYRKEKKQDY-NASPLPAGMSPTNPLPAGR
CbnPID-1          RFELAALGFELKGKPKSNSEKIEKSENLDYIYRKEKRQDF-AGSPLPVGMSPSNPLPAGR
PpaTOST-1         RFGIIERGYYLE-----SHPDDIPCE------LEEHRLV--NLVPLTKQATLEEVMAELA
CjaTOST-1         RFGVNEKGYTLQADSAVSEPTAQSSKSQVFLTVHERNLI--SKKNFIENYLAKHPTTEVA
CjaTOST-2         RFGVNEKGYTLQAGSAVSESTDQSSKSQVFLTVHERNLI--SKKNFIENYLAKHPTTEVA
CeTOST-1          RFGVLEKGYTIQA----TEPETVKNDSDVFFVVHDRNLN--SKKEFVENYLAKNPIVNSP
CreTOST-1         RFGVLEKGYALQA----TDPNVTKKGADVFFVVHDKNLI--SKKEFIENYVAKNPNTRAQ
CbgTOST-1         RFGVLEKGYMLHA----TEPTETKNAADVYLFVHEKNLI--SKKEFIQNYLAKNPIVDLS
CbnTOST-1         RFGVLEKGYCLEA----DEPTHIEKDSDVFLVVYDRNLI--SKKEFIENYLAKNPAAPKS
                  ** :   *: :                     :..

CjpPID-1          GFISPPAHAATRNSFVN---RRGSPHKHSSP----VAVIRPPIFDNIPSSFV
CePID-1           GFLS-PAIQNTSNQFTFSGSPRITPQKHTPVSA--NHKPARSIFDDIPSNIA
CbrPID-1          GFLS-PAIHATSNTFAYSGHSKTTPQKQSTSAT--TRKPIKSIFDDIPSNIE
CrePID-1          GFLS-PAIQATNSAFPFAGSPRSTPQKYSNSPA--PRQPIRSLFDDIPANIA
CbnPID-1          GFLS-PAIQAAGAAFGF-GSPKVTPQKSMTNAS--PRKPIRSIFDDIPTNIA
PpaTOST-1         E----ATAKRNESKKKRTSVSPNPCESTASSP--------------------
CjaTOST-1         E----EEDKENAHVSKNV--KKSSVTKDVFNPNVYVGDFAKAILKTM-----
CjaTOST-2         E----EEDKENAHVSKNV--KKSSVTKDVFNPNVYVGDFAKAILKTM-----
CeTOST-1          EKVFEEEDKENAPTKKAVLKPSSTDEKKLTSSDIYIQEITRSILQS------
CreTOST-1         EVEV-EEDKENAPSKKSAQNTTSAAKKEITTEE-YLLDITKSILKPI-----
CbgTOST-1         EENV-EEDKENVSSRKST--TPAVSKKETTSED-YVRDIAKSILKAI-----
CbnTOST-1         EEVV-EEDKENAPSKKST--PKSVAQKEFTTED-YVRDITKSILKPI-----
                                      .
```

b



c



Supplemental Figure 8