

Supplemental Data to:

Tdrd1 acts as a molecular scaffold for Piwi proteins and piRNA targets in zebrafish

Hsin-Yi Huang¹, Saskia Houwing^{1,2,γ}, Lucas J.T. Kaaij^{1,γ}, Amanda Meppelink¹, Stefan Redl³, Sharon Gauci⁴, Harmjan Vos⁵, Bruce W. Draper⁶, Cecilia B. Moens⁷, Boudewijn M. Burgering⁵, Peter Ladurner³, Jeroen Krijgsveld^{4,8}, Eugene Berezikov¹ and René F. Ketting^{1,*}

¹Hubrecht Institute-KNAW & University Medical Centre Utrecht, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

²Present address: Kimmel Center for Biology and Medicine at the Skirball Institute, NYU School of Medicine, 540 First Avenue, New York, NY 10016, USA

³Institute of Zoology, Technikerstrasse 25, A-6020 Innsbruck, Austria

⁴Biomolecular Mass Spectrometry and Proteomics Group, Bijvoet Centre for Biomolecular Research, Utrecht Institute for Pharmaceutical Sciences and Netherlands Proteomics Center, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

⁵Department of Physiological Chemistry, University Medical Center Utrecht, Utrecht, The Netherlands

⁶Molecular and Cellular Biology, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA

⁷Howard Hughes Medical Institute and Fred Hutchinson Cancer Research Center, P.O. Box 19024, Seattle, WA 98109, USA

⁸Present address: EMBL, Genome Biology Unit, Meyerhofstrasse 1, 69117, Heidelberg, Germany.

^γ Contributed equally to this work. * correspondence: r.ketting@hubrecht.eu

INVENTORY

Materials and Methods.

Figure S1: Tdrd1 expression data.

Figure S2: Cytosolic localisation of Tdrd1 in PGCs.

Figure S3: *tdrd1* mutant morphology analysis.

Figure S4: Details of transmission EM pictures displayed in Figure 2E.

Figure S5: *tdrd1* mutant analysis on nuage in relation to mitochondria and nuclear membrane.

Figure S6: *tdrd1* mutant piRNA analysis.

Figure S7: Arginine dimethylation on Zili.

Figure S8: Characteristics of Tdrd1 associated RNAs.

References.

MATERIALS AND METHODS

Zebrafish strains and genetics

Zebrafish were kept under standard conditions. The *tdrd1*^{fh244+/-} mutant allele zebrafish was generated by TILLING and obtained from Zebrafish International Resource Center (ZIRC). Animals carrying *tdrd1*^{fh244+/-} were out crossed against wild-type fish (TL) or vas::EGFP transgenic fish (Krovel and Olsen 2004) and subsequently in-crossed to obtain *tdrd1*^{fh244-/-} offspring. For genotyping, the DNA was extracted from caudal fin tissue amputated from anesthetized fish. The primers used to amplify and re-sequence the allele are: fh244_03: 5'-GAA AAA CCT AAG GAG TCA AAA GCT G-3' and fh244_04: 5'-GGCAGAGTGTCTATGCTTGGATAAC-3'. The lesion induces a truncation after amino acid E175. This residue precedes the epitope used for immunization.

Western blot and immuno staining analysis

Western blot and immuno stainings were done as described before (Houwing et al. 2007). Tdrd1 antibodies were raised in rabbits with the synthetic peptide H₂N-RRP ATG PSS LSP RGP C- CONH₂. Antisera were subsequently purified against the synthetic peptide (Eurogentec). Embryos were fixed in 4 % PFA at RT for 3hours. Mouse anti-GFP B-2 (Santa Cruz) was used.

Immuno precipitation

Immuno precipitations were done with Ziwi(Girard et al. 2006), Zili(Houwing et al. 2008), Tdrd1 affinity purified antibodies and SYM11 (Millipore) and ASYM24 (Millipore). Tissues in IP lysis buffer were sonicated for 2 minutes at 4 °C and

centrifuged for 10 minutes at 16000 rpm at 4 °C. Supernatant was used for IP. One IP contains 30 μ l Dynabeads (Invitrogen), 3 or 6 testes, and Ziwi antibody 1:200, Zili antibody 1:250 or Tdrd1 antibody 1:100 in a total volume of 500 μ l. IP lysis buffer and IP wash buffer are described before (Houwing et al. 2008), except the NaCl concentration in IP wash buffer was raised to 500 mM. Beads, antibody and lysate were incubated for 3 hours at 4 °C while rotating. Beads were used for Trizon LS (Invitrogen) RNA isolation or eluted with crack buffer for Western blot or Mass Spectrum analysis. RNA was dephosphorylated and 5' labelled with ³²P and run on 12 % denaturing polyacrylamide gels. Gels were exposed to phosphor-imager screens, which were scanned on a BAS-2500 imager (Fuji).

Reverse Transcriptase PCR

cDNA synthesis and RT-PCR were performed as described before (Houwing et al. 2007).

Primers used for amplifying Tdrd1 fragments are: Tdrd1-F1: 5'- GTA TAA AGT CTC TAA CAG ACA CCT GG -3' and Tdrd1-R1: 5'-CAG GTA ATA CGA CTC ACT ATA GGG GGG GAA TTT CAG TGG AG-3'; Tdrd1-F2: 5'- GTA AAG AGC ACT TCC CTG TA -3' and Tdrd1-R2: 5'-CAG GTA ATA CGA CTC ACT ATA GGG TGG TTC CTC GTG CTT-3'.

Q-PCR

Juvenile gonads (3wpf) were isolated from incrosses of three-week-old *tdrd1*^{fh244+/-} fish with vas::EGFP transgene. Gonads were treated with TrypLE Express (GIBCO) for 30 minutes at 28 °C to dissociate the cells. Fetal calf serum was added to 20% to stop the

reaction. Cells were spun down at 1500rpm for 1 minute and washed once with PBS. Resuspended cells were kept in PBS on ice before sorting. The cell suspension was further filtered and sorted using a MoFlo Fluorescence Activated Cell Sorting machine (Dakocytomatino). Total RNA was isolated from sorted high GFP cells. Q-PCR was done as described before (Houwing et al. 2008). Sequences of primers for I-1 element: I-1-F: 5'- GAG TTG AGG ATT GGA AGA TCA TT-3' and I-1-R: 5'- CTT GAG TTT TAG CAA TAG AAA ATC TAA ATC-3'.

In Situ Hybridizations

In situ hybridization was performed as described before (Houwing et al. 2007). Two *tdrd1* fragments were PCR amplified using the same primers as for RT-PCR with an T7 polymerase promoter added to the reverse primer. The PCR product was used as a template for probe synthesis with T7 polymerase.

Peptide pull down

Peptide pull downs were done with biotinylated synthetic peptides with concentration of 2 mM. Peptide sequences: R68(s): H₂N-GEM PVR FGR(s) GIT QSI AAK(BiotinC6) - CONH₂ ;R68(-): H₂N-GEM PVR FGR GIT QSI AAK(BiotinC6)-CONH₂; R163(s): H₂N-GSS LVS MFR(s) GLG IEP GK(BiotinC6)-CONH₂; R163(-): H₂N-GSS LVS MFR GLG IEP GK(BiotinC6)-CONH₂; R221(s): H₂N-EES ISF LGR(s) GFT GFG RAK(BiotinC6)-CONH₂; R221(-): H₂N-EES ISF LGR GFT GFG RAK(BiotinC6)-CONH₂; R228(s): H₂N-GRG FTG FGR(s) AAM PHM TVK(BiotinC6)-CONH₂; R228(-): H₂N-GRG FTG FGR AAM PHM TVK(BiotinC6)-CONH₂; R700(s): H₂N-EEL VTT FSR(s) VAG PMG

MRK(BiotinC6)-CONH₂; 3a5(s): H₂N-AR(s)T KQT ARK STG GKA PRG

GK(BiotinC6)-CONH₂. 1 μ l of peptide was incubated with 30 μ l Dynabeads M-280 Streptavidin in 500 μ l IP lysis buffer for 1 hour at 4 °C while rotating. Supernatant was removed. The beads were incubated with 500 μ l lysate from 3 testes for 2 hours at 4 °C while rotating and washed extensively. Lysate and washing buffers are identical to those used in IPs. Beads were eluted with crack buffer for Western blot analysis.

GST-tudor protein production and pull down

PCR products for individual Tdrd1 tudor domains were cloned into pGEX-KG vector containing GST tag at the 5' end of the multiple cloning site. GST-fusion tudor constructs were transformed into BL21 bacteria and induced with 1 mM IPTG for 3 hours at 30 °C. Bacteria pellets were resuspended in lysis buffer (150 mM NaCl, 50 mM Tris pH 7.5, 1 mM DTT, 1 % NP40, and protease inhibitor (Complete Mini EDTA-free, Roche)) and sonicated 1 minute at 4 °C. Fusion proteins were purified with glutathione agarose beads (Sigma-Aldrich) from the supernatant.

LC-MS/MS

Proteins were digested in-gel with trypsin, and peptides were extracted for subsequent identification by LC-MS/MS. Nanoflow liquid chromatography was performed on an Agilent 1100HPCL binary solvent delivery system (Agilent Technologies, Waldbronn, Germany) with a thermostated wellplate autosampler coupled to an LTQ-Orbitrap mass spectrometer (Thermo Electron, Bremen, Germany). Peptides were trapped at 5 μ l/min in 100 % A (0.1 M acetic acid in water) on a trapping column (30 mm x 100 μ m packed in-

house with Aqua C18, Phenomenex, Torrance, CA) for 10 minutes. After flow-splitting down to around 100 nl/min, peptides were transferred to the analytical column (200 mm x 50 μ m packed in-house with Reprisil-Pur C18 AQ, Dr. Maisch GmbH, Ammerbuch, Germany) and eluted with a gradient of 0-40 % B (80 % Acetonitrile/0,1 M Acetic Acid) in a 60-minute gradient. Nano spray was achieved using a coated fused silica emitter (New Objective, Cambridge, MA) (o.d., 360 μ m; i.d., 20 μ m, tip i.d. 10 μ m).

The LTQ-Orbitrap mass spectrometer was operated in data-dependent mode, automatically switching between MS and MS/MS. The two most intense peaks were selected for collision-induced dissociation (CID) in the linear ion trap at normalized collision energy of 35 %. Full scan MS spectra were acquired with a resolution of 60,000 at 400 m/z after accumulation to a target value of 500000.

All MS/MS spectra were converted to DTA files using Bioworks 3.3 (Thermo, San Jose) and searched against the IPI Zebrafish protein database (release 3.35) using MAS-COT (Version 2.2.01, Matrix Science, London, UK). Cysteine carbamidomethylation and methionine oxidation were used as fixed and variable modification, respectively. In selected cases, mono- and di-methylation of Arginine, or mono, di, and tri-methylation of Lysine was used as variable modifications. A peptide mass tolerance of 5 ppm and fragment mass tolerance of 0.8 Da were selected, using Trypsin as the proteolytic enzyme allowing one missed cleavage. All data were loaded into Scaffold (Version 02.01.00, Proteome-Software, Portland, OR) to probabilistically validate peptide and protein identifications. Peptide and protein identifications were accepted when reaching a minimum of 95 % probability, with a minimum of two identified peptides per protein.

Deep sequencing

Long RNA was poly(A)-tailed using poly(A) polymerase followed by ligation of a RNA adaptor to the 5'-phosphate of the RNAs. First-strand cDNA synthesis was performed using an oligo(dT)-linker primer and M-MLV-RNase H- reverse transcriptase. The cDNA was PCR-amplified for 16 cycles according to the instructions of Illumina/Solexa. cDNA was finally purified using the Macherey & Nagel NucleoSpin Extract II kit and was sent for a 50 cycle sequencing run on an Illumina/Solexa platform. piRNAs were cloned using both 5' and 3' adaptor ligations, and sequenced with 36 cycles.

Sequence analysis

The four-base adaptor sequences were trimmed from generated data using custom scripts. Resulting inserts were mapped to the zebrafish genome (Zv8 assembly) using the megablast program (Zhang et al. 2000), allowing mismatches in the reads starting from nucleotide 19 and considering the longest possible matches as true mapping positions. The mismatched 3' ends of the reads were trimmed, keeping track of the identity of trimmed nucleotides for later counting. Genomic annotations of mapped reads were retrieved from the Ensembl database (release 56) using Perl API provided by Ensembl (Hubbard et al. 2007). Read counts for the different annotated classes are listed in Table S3. Sequences will be made available at GEO upon publication. In the analyses transposon-derived piRNAs are defined as reads with a length between 24 and 29 nucleotides. Heatmap was constructed with basic R programming. The transposons used were filtered on the presence of a strandbias between Ziwi and Zili piRNA populations.

The map was sorted on Tdrd1 associated piRNAs. Colouring indicates relative strand ratios comparing the four different piRNA pools.

TEM

Trunks of 3 week old zebrafish were fixed with half-strength Karnovsky fixative (pH 7.4) (Karnovski 1965) and postfixed with 1% OsO₄ in 0.05M cacodylate buffer (pH 7.4), dehydrated in acetone series and embedded in Spurr's low viscosity resin (Spurr 1969). Semi-thin, 1- μ m-thick sections were cut with a histo Butler diamond knife (Diatome) and mounted on glass slides. After drying, the sections were stained with methylen blue Azur II mixture according to Richardson (Richardson et al. 1960) for light microscopy.

Ultrathin sections (70nm in thickness) were placed on copper grids, stained with uranyl acetate followed by lead citrate, and examined with an electron microscope (ZEISS Libra 120 energy filter electron microscope). Image acquisition and analysis were performed using a 2 k Vario Speed SSCCD camera (Droendle), the iTEM software (TEM imaging platform, Olympus) and Adobe Photoshop.

Immunogold EM

Ovaries and testes of wild type zebrafish were dissected and fixed in ice cold 4% PFA in 0.1 M PBS pH7.4 for 3 hours. Three-week-old *tdrd1*^{fh244+/-} incrossed fish were genotyped and decapitated before fixation as described above. After fixation, samples were washed 3 times for 30 minutes in PBS. Samples were further dehydrated in a standard alcohol series and embedded in LR White medium resin (London Resin Company Ltd).

Polymerization was performed at 60°C for 48 hours in an oxygen free atmosphere. Semi-

thin sections were cut on a Leica 2040 autocut and ultra-thin sections (90 nm) on a Leica ultramicrotome UCT.

Ultra-thin sections were mounted on pioloform coated gold slot grids, humidified, washed in PBS and blocking was performed using 1% BSA in 0.1 M PBS and 1 % BSA + 0.1% FG in 0.1 M PBS for 25 min each. The grids were incubated with primary antibody in 0.1 M PBS with 1% BSA and 0.1% FG at 4°C over night. After washing they were incubated in goat anti-rabbit secondary antibody, conjugated to 10 nm or 20 nm gold particles, at a dilution of 1:100 for 1 hour at room temperature. After washing and postfixation with 2.5 % GA in 0.1 M PBS, they were counterstained with saturated uranyl acetate and lead citrate. Samples were examined with a Zeiss Libra 120 Energy Filter Transmission Electron Microscope. Zili and Tdrd1 antibodies used for immunogold were affinity purified. Concentrations used for immunogold stainings are: Zili 1:2000, Tdrd1 1:400.

SUPPLEMENTAL FIGURE LEGENDS.

Figure S1. Tdrd1 expression data. **(A)** Transcripts from *tdrd1* were visualized by *in situ* hybridisation (purple) on adult testis (left panel) and ovary (right panel). Sections were counterstained with H&E. Scale bars are 100 μ m. **(B)** RT-PCR analysis for *tdrd1* expression on indicated samples. **(C)** Immunohistochemistry for Tdrd1 on testis. Scale bars are 10 μ m in the right panel and 50 μ m in the left panel. **(D)** Immunohistochemistry for Tdrd1 (red) and GFP (green) on embryos between 3 and 7dpf, containing a *vasa::EGFP* transgene marking the PGCs. Scale bars are 10 μ m. **(E)** Tdrd1 immunogold labelling on wild-type adult ovaries. Scale bar is 2 μ m. White arrow-heads indicate nuage.

Figure S2. Cytosolic localisation of Tdrd1 in PGCs. A single confocal section through the nucleus of a *vasa::EGFP* transgenic PGC, stained for Tdrd1 (Red) and DNA (DAPI, blue).

Figure S3. *tdrd1* mutant morphology analysis. **(A)** Zili immunogold labelling on three-week-old wild-type ovaries. **(B)** Zili immunogold labelling on three-week-old *tdrd1* mutant ovaries. Scale bars are 2 μ m in A and B. Black arrow-heads indicate nuage. **(C)** *Vasa in situ* hybridization (purple) on gonads from wild-type and *tdrd1* mutant animals at different developmental stages. Sex was judged by the presence or absence of juvenile oocytes. Wpf: weeks post fertilization. SB: swim-bladder (gonads are close to the swimbladder and in some preparations the SB is visible). Scale: 100 μ m. **(D, E)** Quantification method and the quantification result of Zili immunohistochemistry on wild-type and *tdrd1* oocytes at 3 weeks. The width of the Zili peak was determined by

making a line scan through the cell, always taking the longest axis. At 20% of light saturation the width of the Zili peak was determined. This was divided by the cell width (which is unaffected by *tdrd1* mutation). A Chi-squared-test was used to calculate the indicated p-value.

Figure S4. Details of transmission EM pictures displayed in Figure 2E. M: mitochondria. NC: nucleolus. White arrow-heads point at nuclear envelope. Solid arrow-heads point at nuage.

Figure S5. Distance measurements between nuage and mitochondria and the nuclear membrane.

The distance of nuage patches to mitochondria the nuclear membrane was measured on TEM images obtained from wild-type and *tdrd1* mutant samples. Results are displayed in a bin plot. Numbers of observations are given in the table below. Significance was assessed through Chi-test analysis of wild-type versus mutant patterns:

$p=1.6 \cdot 10^{-5}$ for nuage-mitochondria distance. $p=0.42$ for nuage-nuclear envelope distance.

Figure S6. *tdrd1* mutant piRNA analysis. (A) Length profile of cloned small RNAs from wild-type and *tdrd1* mutant gonads isolated from 3-week-old animals. (B) Anti-sense/sense piRNA ratios for indicated elements. Asterisks indicate significant differences ($p < 10^{-15}$; Chi-squared test). Wild-type: blue. *Tdrdl*: red. Elements left of the vertical line have antisense enrichment in Ziwi, elements to the right have sense

enrichment in Zili (Houwing et al. 2008). (C) Comparison between changes in transcript levels and changes in piRNA abundance in wild-type versus *tdrd1* mutant animals.

(D, E) Ping-pong signal among wild-type and *tdrd1* mutant piRNAs. In (D) the overlap length is plotted against the ratio between number of reads given a particular length overlap and the total mapped piRNAs. This reveals relatively little overlapping piRNA species in the *tdrd1* mutant library. In (E) the overlap length is plotted against the ratio between number of reads given a particular length overlap and the total amount of overlapping piRNAs. This corrects for the low amount of overlaps in the *tdrd1* mutant data, and reveals a functional ping-pong mechanism in the *tdrd1* mutant, although much less efficient.

Figure S7. Arginine dimethylation on Zili. Mass spectrum identifying a Zili peptide spanning R68 in a dimethylated form. Fragments (b- and y-ions) arising from a triply-charged precursor peptide (668.2819 m/z) are indicated in the spectrum and in the identified peptide.

Figure S8. Characteristics of Tdrd1 bound RNA. (A) Genome browser views of five different transposable elements displaying TAT sequence tags (tops) and piRNAs (bottom) from testis. Blue ticks represent matches to the top genomic strand, red ticks matches to the bottom strand. (B) Sense-antisense ratios for transposable element derived TATs, Tdrd1 associated piRNAs and total testis piRNAs.

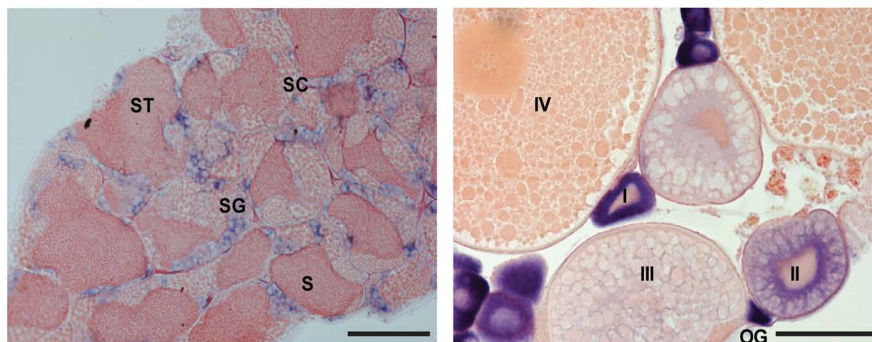
REFERENCES

- Girard A, Sachidanandam R, Hannon GJ, and Carmell MA. 2006. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442**(7099): 199-202.
- Houwing S, Berezikov E, and Ketting RF. 2008. Zili is required for germ cell differentiation and meiosis in zebrafish. *EMBO J* **27**(20): 2702-2711.
- Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, Filippov DV, Blaser H, Raz E, Moens CB et al. 2007. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell* **129**(1): 69-82.
- Hubbard TJ, Aken BL, Beal K, Ballester B, Caccamo M, Chen Y, Clarke L, Coates G, Cunningham F, Cutts T et al. 2007. Ensembl 2007. *Nucleic Acids Res* **35**(Database issue): D610-617.
- Karnovski MJ. 1965. A formaldehyde - glutaraldehyde fixative of high osmolality for use in electron microscopy. *J Cell Biol* **27**: 137A-138A.
- Krovel AV and Olsen LC. 2004. Sexual dimorphic expression pattern of a splice variant of zebrafish vasa during gonadal development. *Dev Biol* **271**(1): 190-197.
- Richardson KC, Jarret L, and Finke EH. 1960. Embedding in epoxy resins for ultrathin sectioning in electron microscopy. *Stain Technology* **35**: 313-323.
- Spurr AR. 1969. A low-viscosity epoxy resin embedding medium for electron microscopy. *J Ultrastruct Res* **26**: 31-45.
- Zhang Z, Schwartz S, Wagner L, and Miller W. 2000. A greedy algorithm for aligning DNA sequences. *J Comput Biol* **7**(1-2): 203-214.

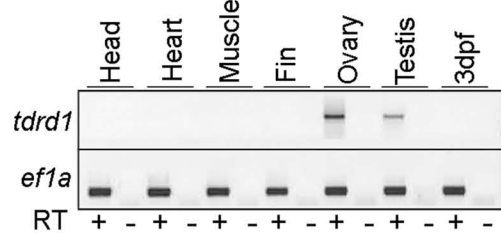
	Testis (1+2)	Tdrd1 IP long RNA	Tdrd1 IP small RNA	3 week wt	3wk tdrd1	Ziwi IP	Zili IP
miRNA	3,804,103	1,046,301	69,578	2,583,881	3,838,925	80,708	19,609
repeat piRNA	2,395,245	5,418,209	3,163,652	1,587,738	1,241,907	863,546	3,815,685
genic piRNA	170,278	923,895	587,756	198,895	201,876	246,289	864,439
other piRNA	3,467,793	6,813,346	5,547,757	1,883,575	1,693,755	1,916,994	6,641,272
snRNA	15,000	413,645	11,278	26,097	23,377	11,011	5,056
snoRNA	5,519	325,972	5,632	38,706	28,460	13,003	2,524
tRNA	1,803	13,734	15,424	5,884	7,061	13,393	3,456
rRNA	7,508	4,181,452	8,432	229,351	156,002	11,885	4,152
Total	9,867,249	19,136,554	9,409,509	6,554,127	7,191,363	3,156,829	11,356,193
unique species (%)	47	72	35	83	85	52	70
Sequenced 5 fold or less (%)	90	98	83	97	96	94	96

TableS3. Read annotation from Illumina sequencing. Note that the Tdrd1IP long RNA library 'piRNA' are for the most part not true piRNAs, but longer reads mapping to the same locations as mature piRNAs.

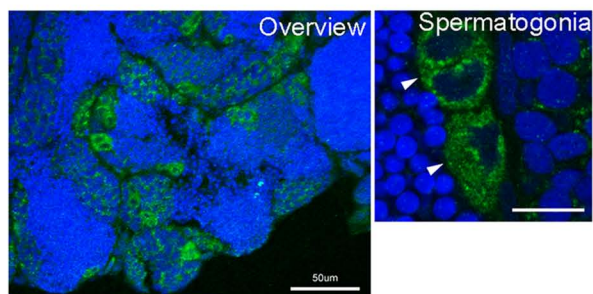
A



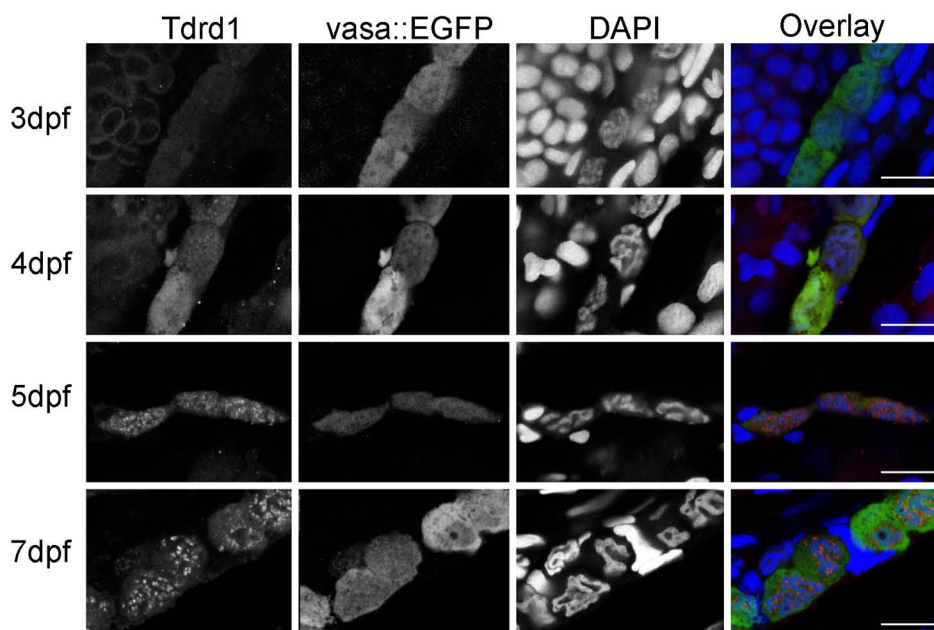
B



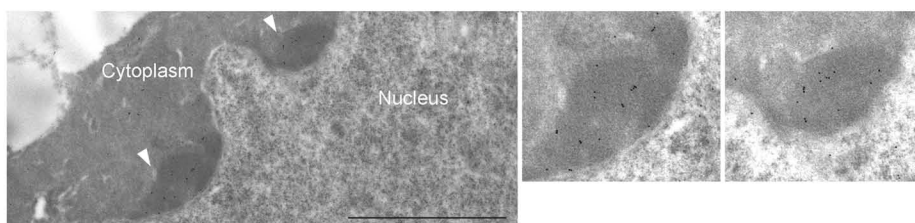
C

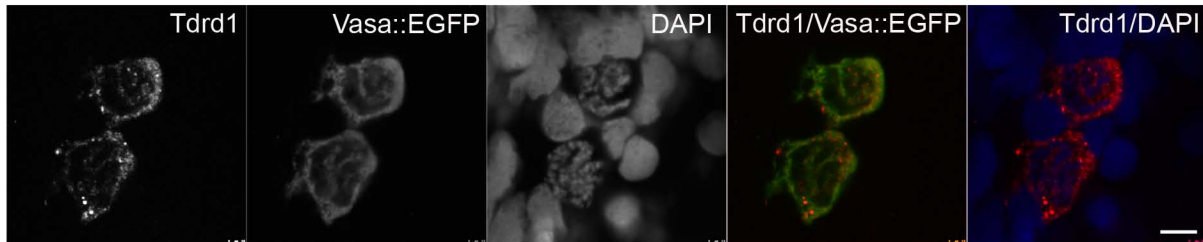


D

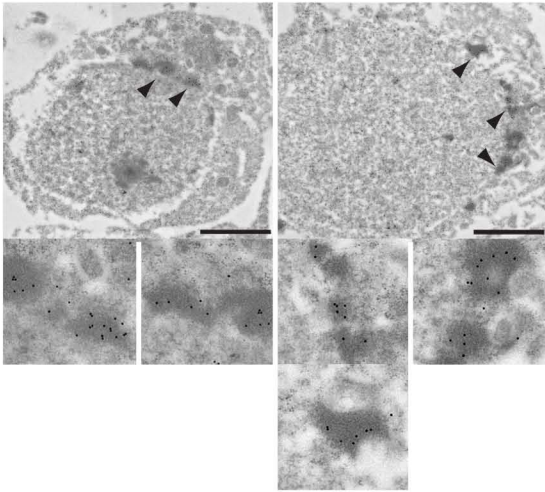


E

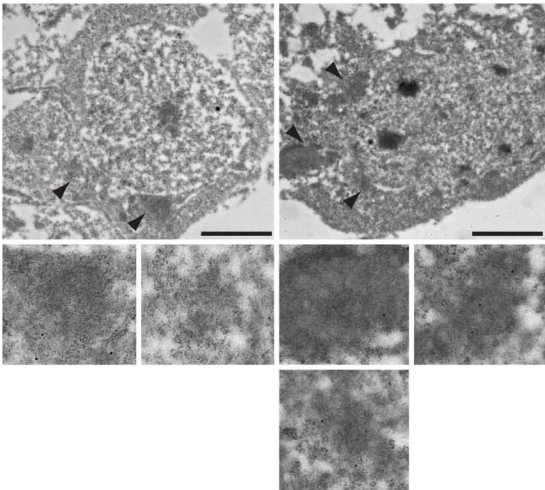




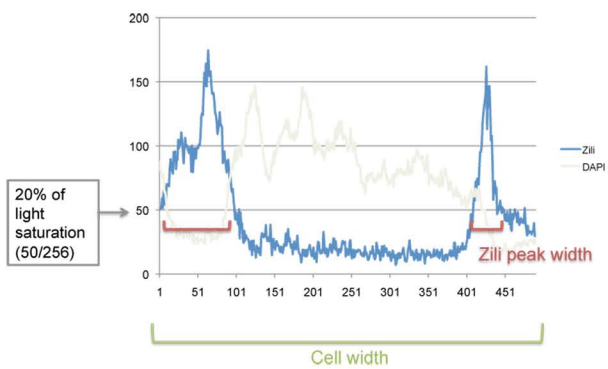
A

3wpf *tdrd1*^{fh244+/+}

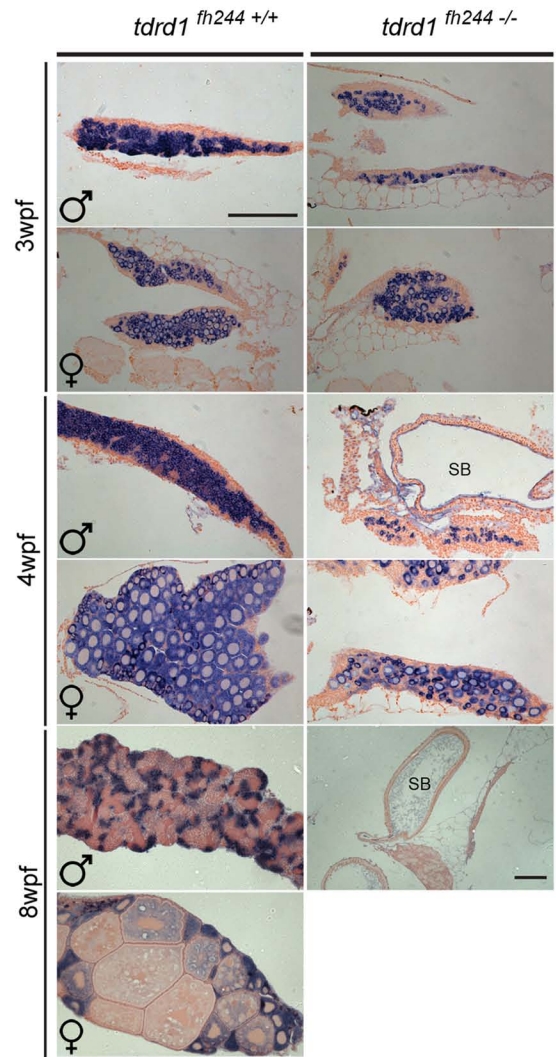
B

3wpf *tdrd1*^{fh244-/-}

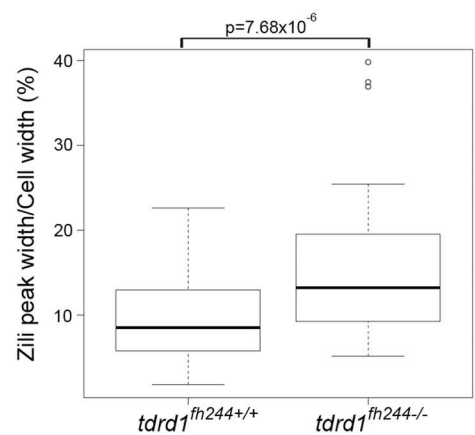
D

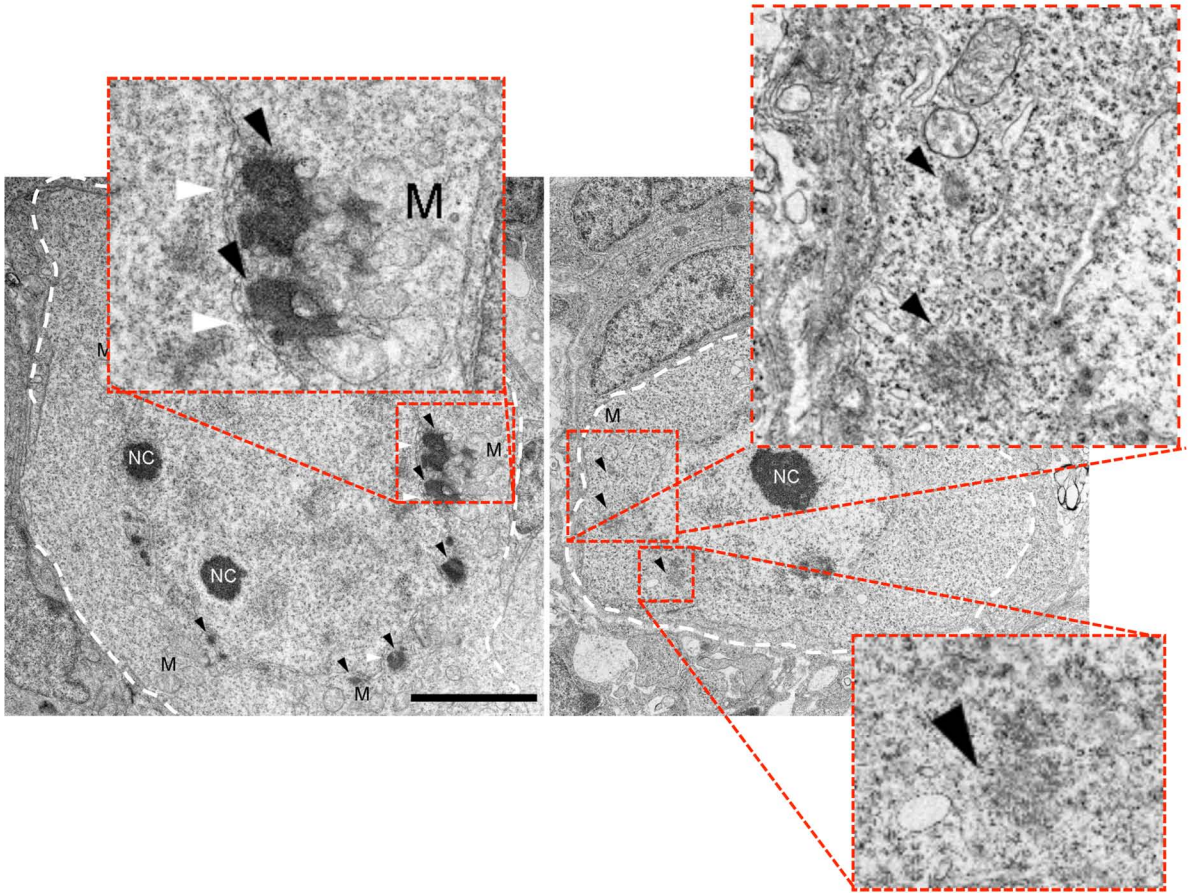


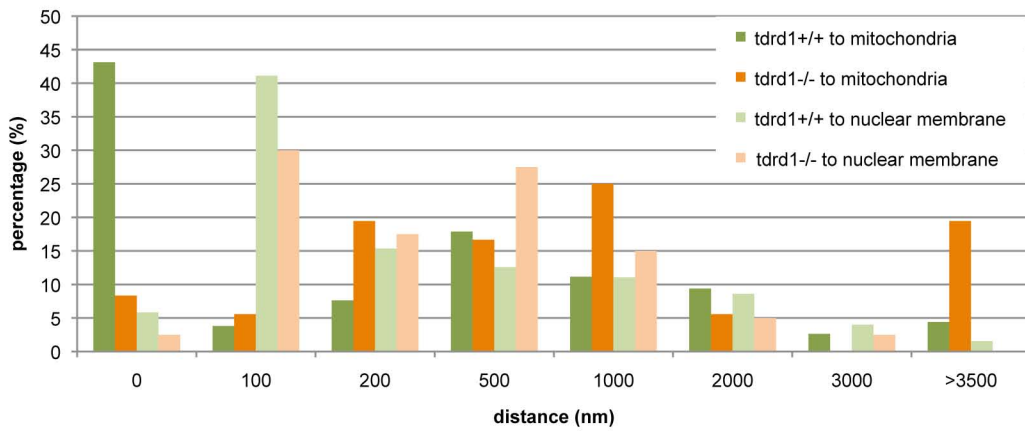
C



E

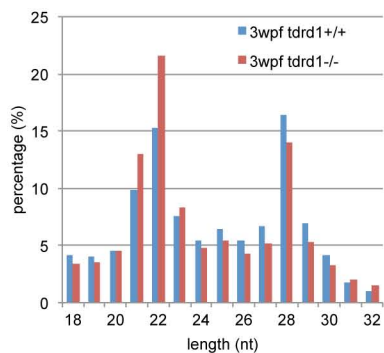




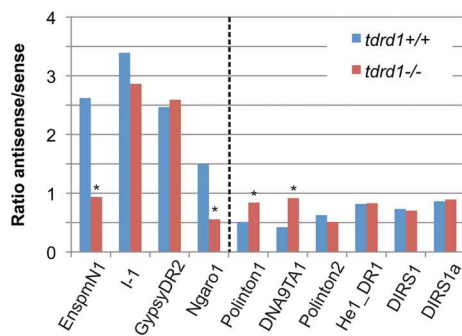


	<i>tdrd1</i> ^{+/+}	<i>tdrd1</i> ^{-/-}
Total number of cells	51	50
Total number of nuage	343	43
Average number of nuages per cell	6.725	0.86

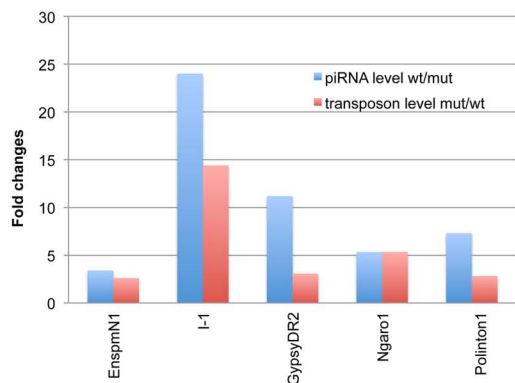
A



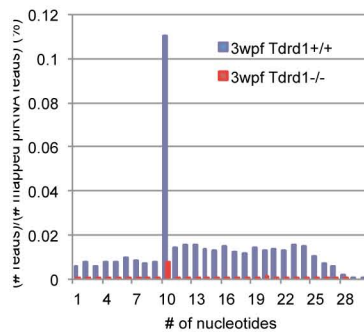
B



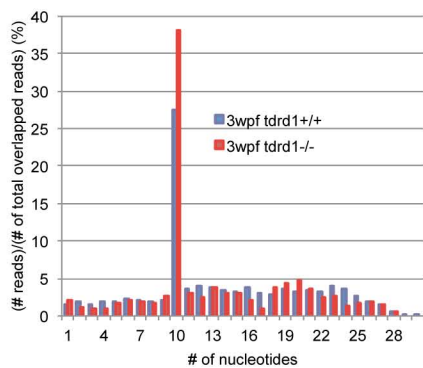
C

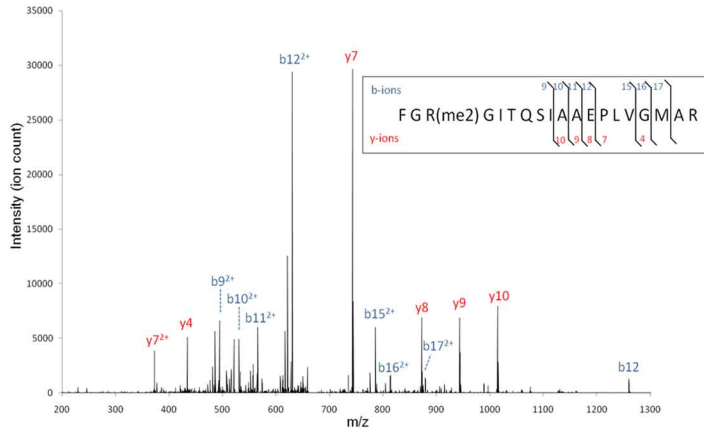


D

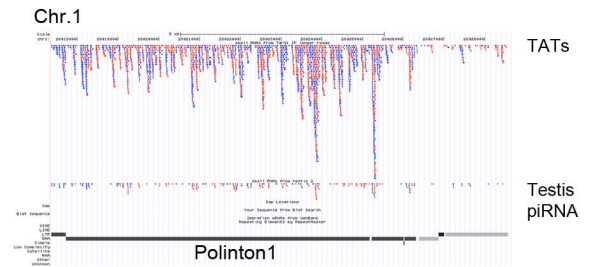
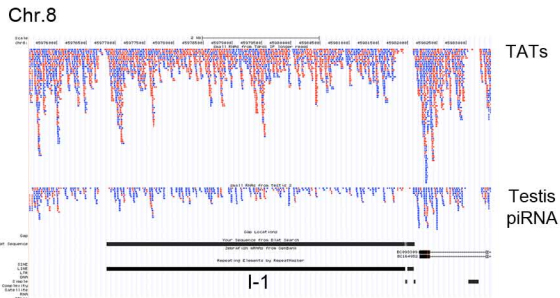
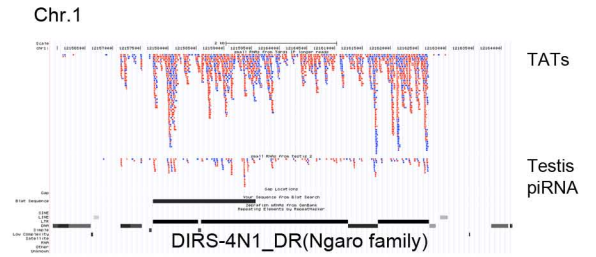
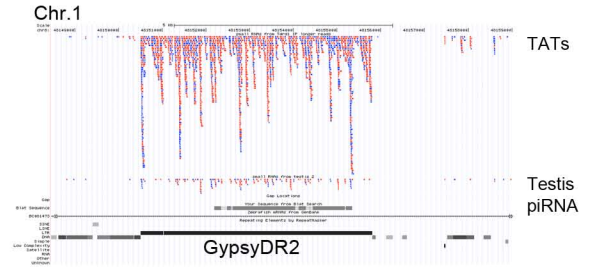
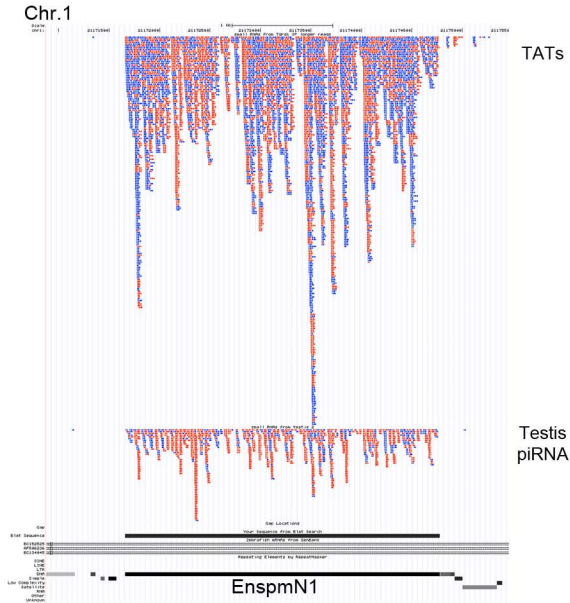


E





A



B

