A H3K27me3 reader complex couples H3K27me3 accumulation to nascent transcription of transposable elements in *Paramecium*

Thomas Balan[1], Ambre Petitalot[2], Julien Richard Albert[1], Louise Abbou[1], Guillaume Chevreux[1], Bassem Al-Sady[3], Bihter Özdemir Aygenli[4], Till Bartke[4], Raphaël Margueron[2], Sandra Duharcourt[1]

[1]Université Paris Cité, CNRS, Institut Jacques Monod, F-75013 Paris, France
[2]Institut Curie, Paris Sciences et Lettres Research University, INSERM, CNRS, F-75005 Paris, France
[3]Department of Microbiology & Immunology, George Williams Hooper Foundation, University of California San Francisco, San Francisco, California, United States of America
[4]Institute of Functional Epigenetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany

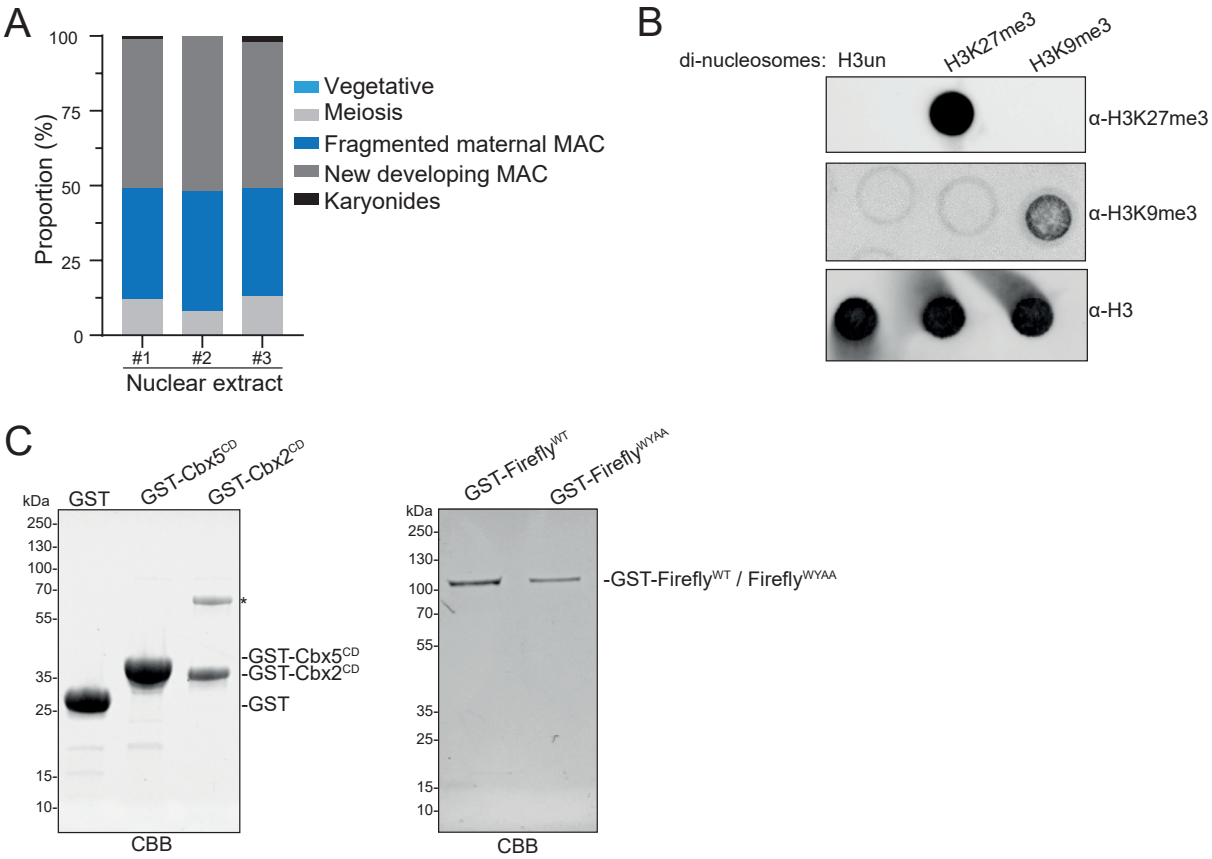
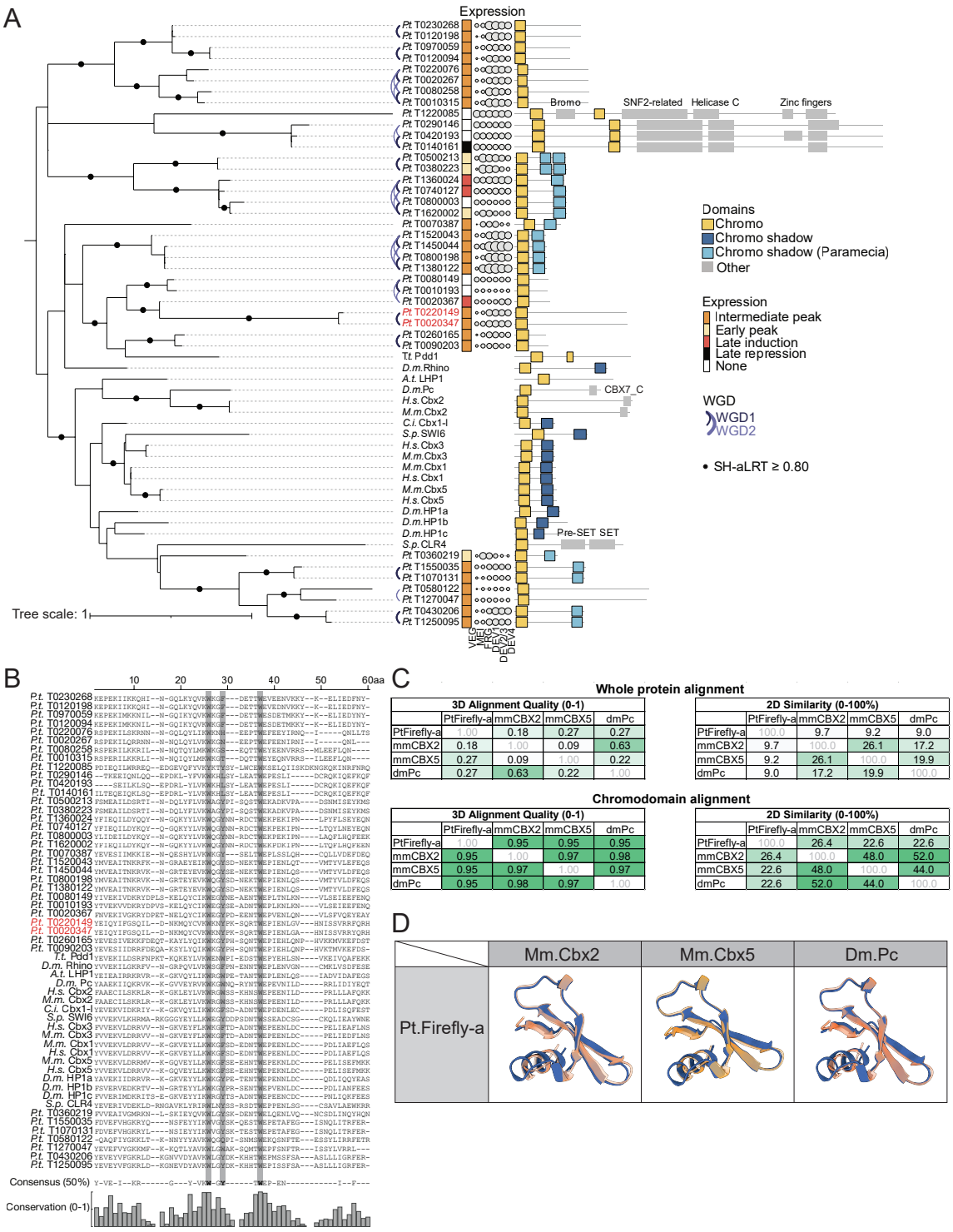


Figure S1: related to figure 1

A. Cytology of cell populations used to prepare the nuclear extracts for the nucleosome pull-downs. At least 100 cells were analyzed in each condition. Proportion of cells at each developmental stage are shown (%).

B. Dot blot analysis of the dinucleosomes used for the pull-downs. Unmodified (un), H3K27me3 or H3K9me3 dinucleosomes were assessed using anti-H3K27me3, anti-H3K9me3 and anti-H3 antibodies (Frapporti et al., 2019).

C. Coomassie blue staining of purified recombinant GST, GST-Cbx5^{CD}, GST-Cbx2^{CD} and GST-Firefly^{WT} and GST-Firefly^{WYAA}. *: contaminant



3D Alignment Quality (0-1)	PtFirefly-a	mmCBX2	mmCBX5	dmPc
PtFirefly-a	1.00	0.18	0.27	0.27
mmCBX2	0.18	1.00	0.09	0.63
mmCBX5	0.27	0.09	1.00	0.22
dmPc	0.27	0.63	0.22	1.00

2D Similarity (0-100%)	PtFirefly-a	mmCBX2	mmCBX5	dmPc
PtFirefly-a	100.0	9.7	9.2	9.0
mmCBX2	9.7	100.0	26.1	17.2
mmCBX5	9.2	26.1	100.0	19.9
dmPc	9.0	17.2	19.9	100.0

3D Alignment Quality (0-1)	PtFirefly-a	mmCBX2	mmCBX5	dmPc
PtFirefly-a	1.00	0.95	0.95	0.95
mmCBX2	0.95	1.00	0.97	0.98
mmCBX5	0.95	0.97	1.00	0.97
dmPc	0.95	0.98	0.97	1.00

2D Similarity (0-100%)	PtFirefly-a	mmCBX2	mmCBX5	dmPc
PtFirefly-a	100.0	26.4	22.6	22.6
mmCBX2	26.4	100.0	48.0	52.0
mmCBX5	22.6	48.0	100.0	44.0
dmPc	22.6	52.0	44.0	100.0

Figure S2: related to figure 1

A. Phylogenetic tree analysis of chromodomain containing proteins in the *Paramecium tetraurelia* genome. The tree was constructed using the chromodomain only, using InterPro PF00385 and hmmsearch (see Methods). For comparison, chromodomain-containing proteins from other organisms were added to the analysis. Statistical support was assessed using SH-aLRT (n = 1,000). Additional conserved protein domains are shown on the right. Expression levels at different stages of the life cycle (circles of different sizes), expression clusters [1] of the coding genes and whole-genome duplication (WGD) relationships [2] are displayed. Firefly-a and Firefly-b are displayed in red.

B. Multiple alignment of the chromodomains shown in panel A. 50% consensus sequence and amino acid conservation are shown.
C. 3D predicted structures and amino-acid sequence similarity matrices of whole proteins or chromodomains only for Pt.Firefly-a, Mm.Cbx2, Mm.Cbx5 and Dm.Pc.
D. 3D predicted structure superposition of the chromodomain of Pt.Firefly-a with the chromodomain of either Mm.Cbx2, Mm.Cbx5 or Dm.Pc. Pt.Firefly-a is in blue.

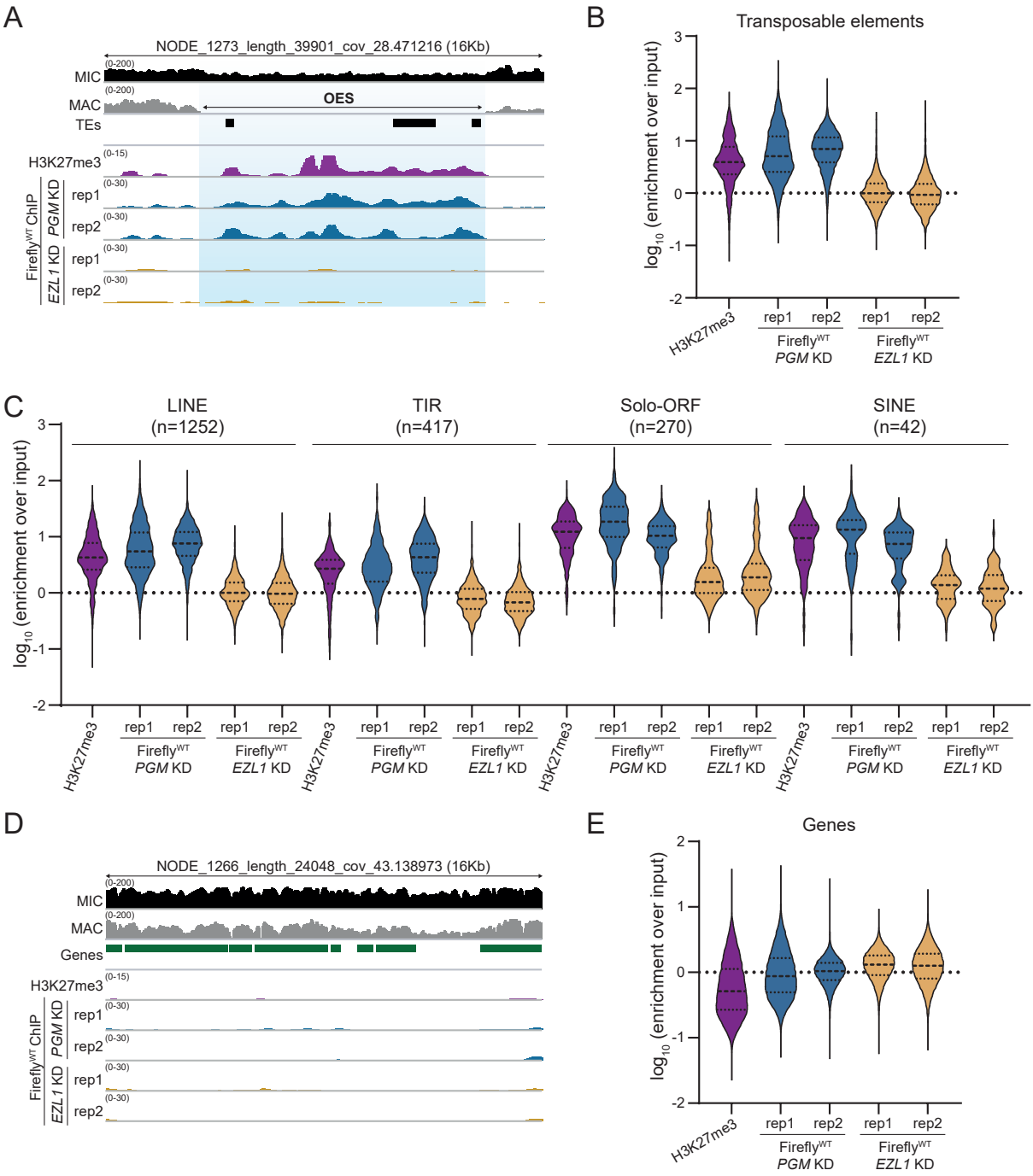


Figure S3: related to figure 1

A. Genome browser screenshot of an example of a MIC-limited region containing TEs, displaying normalized H3K27me3 ChIP enrichment in *PGM* KD and Firefly (FLAG) ChIP enrichment (2 replicates) in *PGM* and *EZL1* KDs. Same region as that shown in Figure 1F and Figure 3B. Replicates #1 are shown in Figure 1F.

B. Enrichment of H3K27me3 ChIP-seq signal over all annotated TEs [3] in *PGM* KD and of Firefly (FLAG) ChIP-seq signal in *PGM* and *EZL1* KDs (2 replicates).

C. Enrichment of H3K27me3 ChIP-seq signal over each family of TEs [3] in *PGM* KD and of Firefly ChIP-seq signal in *PGM* and *EZL1* KDs (2 replicates).

D. Genome browser screenshot of an example of a MAC-destined region containing genes, displaying normalized ChIP enrichment of H3K27me3 in *PGM* KD and of Firefly (FLAG) in *PGM* and *EZL1* KDs (2 replicates).

E. Enrichment of H3K27me3 ChIP-seq signal over genes in *PGM* KD and of Firefly (FLAG) in *PGM* and *EZL1* KDs (2 replicates).

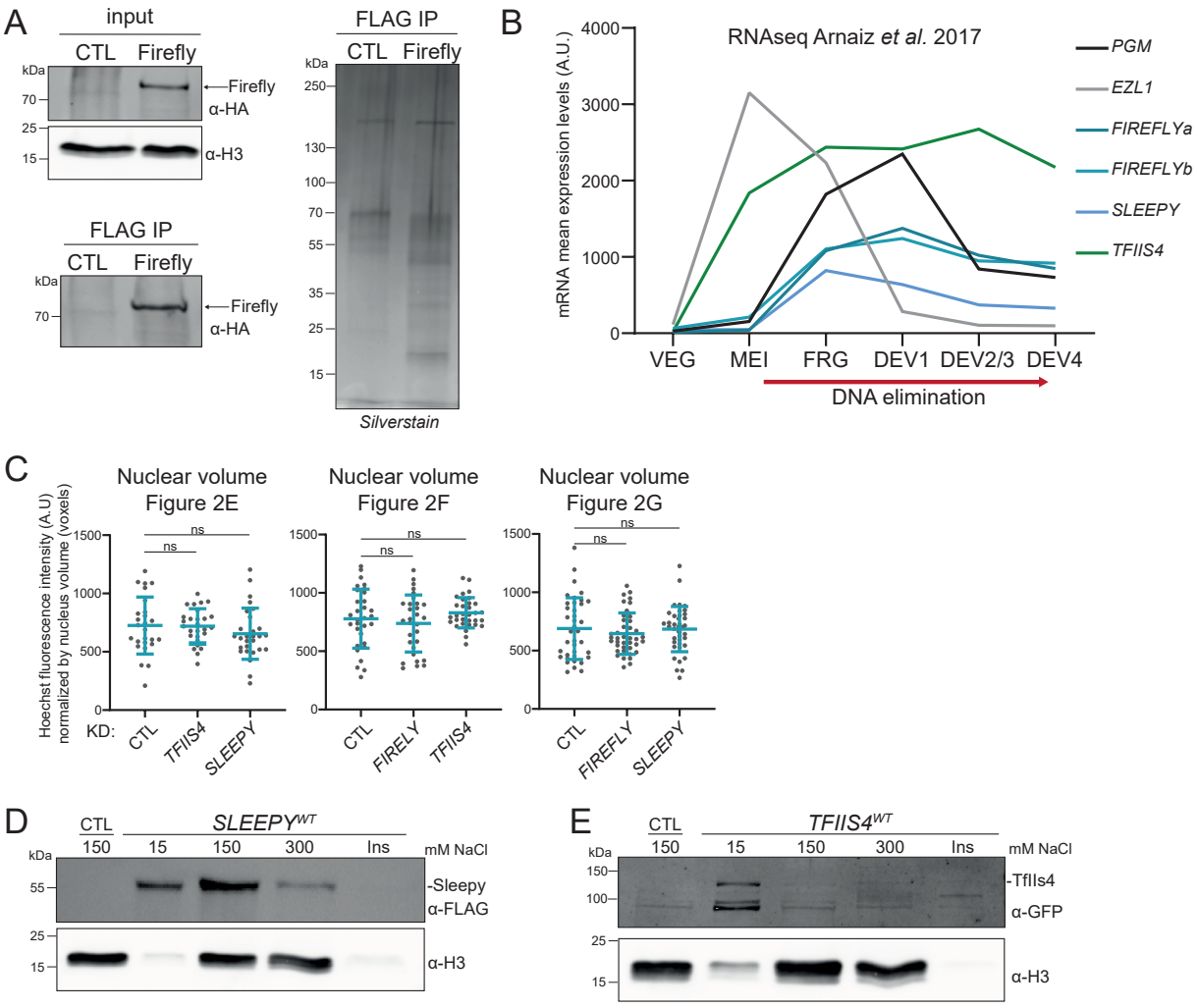


Figure S4: related to figure 2

A. (Left) Western blot analysis of nuclear extracts of wild-type cells (CTL) or cells expressing a FLAG-HA-Firefly functional protein (Firefly) before (input) or after affinity purification (FLAG IP). Anti-HA antibodies are used to detect Firefly and anti-H3 antibodies for normalization. (Right) Silver-stained gel of pulled-down (FLAG IP) proteins. Predicted MW for FLAG-HA-Firefly: 64.7kDa.

B. mRNA expression levels of the genes encoding *PGM, EZL1, FIREFLY-a, FIREFLY-b, TFIIS4* and *SLEEPY* at different developmental stages during autogamy in arbitrary units. *EZL1* belongs to the early expression gene cluster, while *PGM, FIREFLY-a, FIREFLY-b, TFIIS4* and *SLEEPY* belong to the intermediate expression gene cluster [1].

C. Boxplot of estimated nucleus (developing MAC) volume in voxels from the same data as in Figure 2E-G. Estimations of nuclear volume indicate that the cell populations were at comparable developmental stages. Bars correspond to mean +/- SD. Mann-Whitney statistical test. n.s: non-significant

D. Western blot analysis of nuclear extracts prepared with increasing salt concentrations (15 mM; 150 mM; 300 mM NaCl) and the insoluble fraction (ins) from cells expressing a *FLAG-HA-SLEEPY^{WT}* transgene at T= 10 hours. Nuclear extracts prepared at 150 mM NaCl from non-injected cells (CTL) at T= 10 hours are used as a control. Anti-FLAG antibodies are used to detect Sleepy and histone H3 antibodies for normalization. Predicted MW for FLAG-HA-Sleepy: 43.7 kDa.

E. Western blot analysis of nuclear extracts prepared with increasing salt concentrations (15 mM; 150 mM; 300 mM NaCl) and the insoluble fraction (ins) from cells expressing a *GFP-TFIIS4*^{WT} transgene at T= 10 hours. Nuclear extracts prepared at 150 mM NaCl from non-injected cells (CTL) at T= 10 hours are used as a control. Anti-GFP antibodies are used to detect TfIIs4 and histone H3 antibodies for normalization. Predicted MW for GFP-TfIIs4: 65.5 kDa.

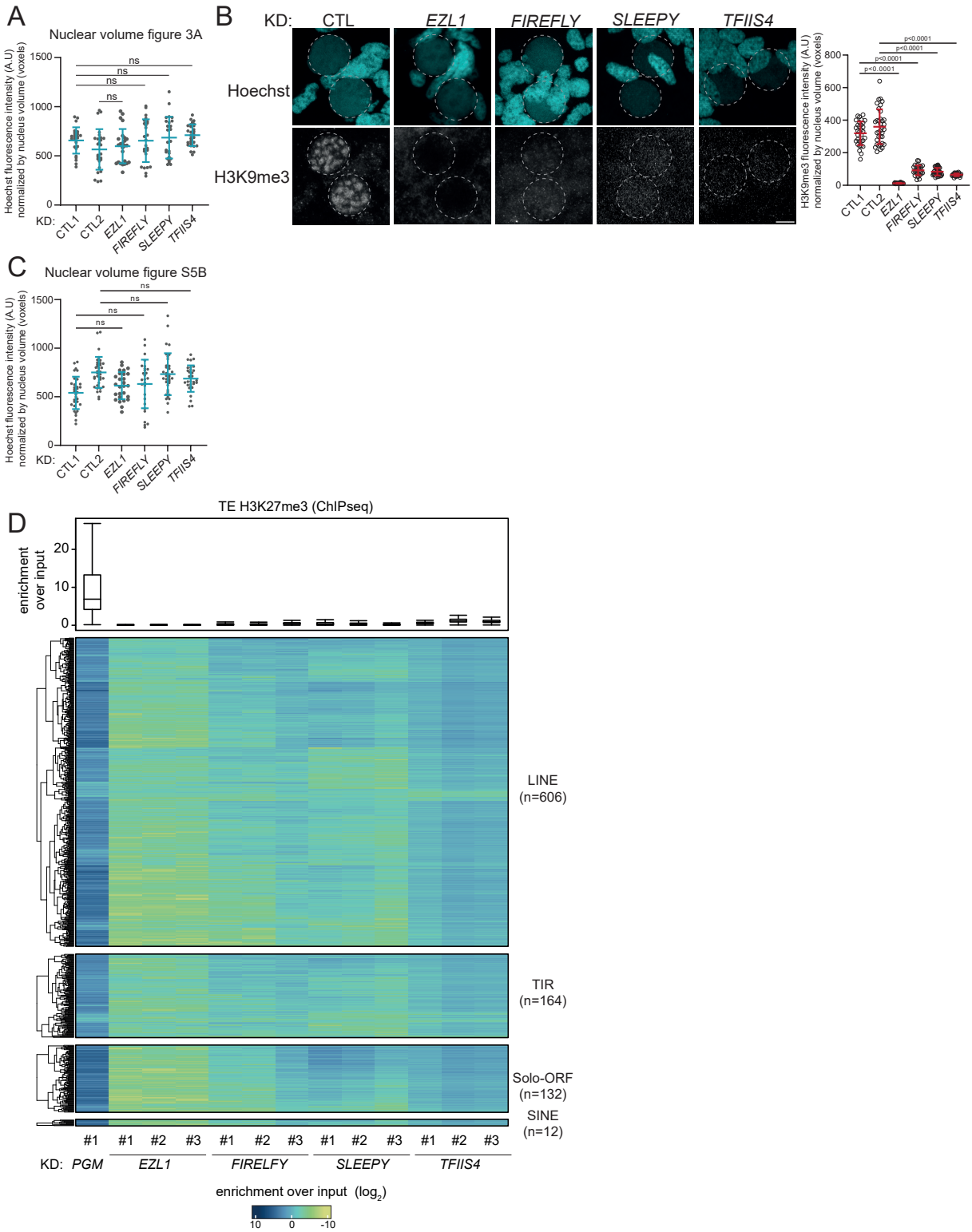


Figure S5: related to figure 3

A. Boxplot of estimated nucleus (developing new MAC) volume in voxels from the same data as in Figure 3A (H3K27me3). Estimations of nuclear volume indicate that CTL, *EZL1, FIREFLY, SLEEPY* and *TFIIS4* KD cell populations were at comparable developmental stages. Bars correspond to mean +/- SD. Mann-Whitney statistical test. n.s: non-significant

B. Immunostaining of H3K9me3 on wild-type cells at T=10 hours after the onset of autogamy upon *ICL7* (CTL1, CTL2), *EZL1, FIREFLY, SLEEPY* or *TFIIS4* KD. CTL1 is the control for *EZL1* and *FIREFLY* KD. CTL2 is the control for *SLEEPY* and *TFIIS4* KD. Developing new MACs are circled

with a dotted line. Scale bar is 5 µm. Quantification of H3K9me3 fluorescence signal in the new MAC (see Methods) in the same conditions. Bars correspond to mean ± SD. Mann-Whitney statistical tests.

C. Boxplot of estimated nucleus (developing MAC) volume in voxels from the same data as in Figure S5B (H3K9me3). Number of nuclei > 25 in each condition. Estimations of nuclear volume indicate that CTL1, CTL2, *EZL1, FIREFLY, SLEEPY* or *TFIIS4* KD cell populations were at comparable developmental stages. Bars correspond to mean +/- SD. Mann-Whitney statistical test. n.s: non-significant.

D. Heatmap and boxplot of H3K27me3 enrichment over input for TE copies as determined by ChIP-seq for the considered TE copies upon *PGM, EZL1, FIREFLY, SLEEPY* or *TFIIS4* KD. All replicates are shown. The TE copies are ordered by hierarchical clustering of DNA coverage in each family. Replicates #1 are shown in Figure 3B-C.

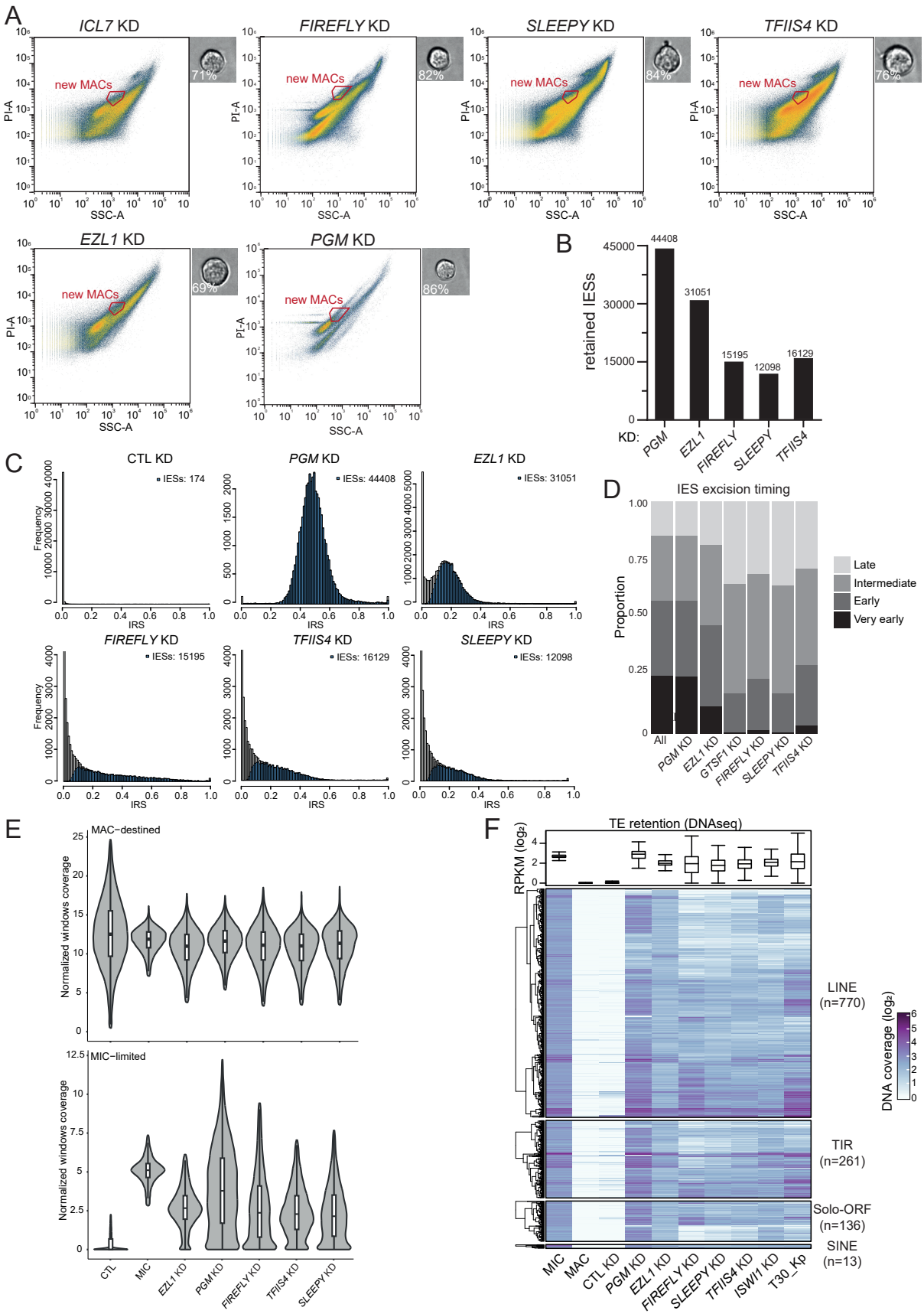


Figure S6: related to figure 3

A. Flow cytometry sorting of developing new MACs from autogamous cells (T=50 hours) subjected to *ICL7, FIREFLY, SLEEPY*, *TFIIS4*, *EZL1 or PGM* KD. The "new MACs" gate was used for

sorting. A representative image in phase contrast of the sorted population is shown with an estimation of purity.

B. Histograms of the number of retained IESs upon *PGM, EZL1, FIREFLY, SLEEPY* or *TFIIS4* KD (samples from panel A).

C. Distribution of IES retention scores (IRS) in sorted developing new MACs upon *ICL7* (CTL), *PGM*, *EZL1*, *FIREFLY*, *TFIIS4* or *SLEEPY* KD (samples from panel A).

D. IES proportions in the 4 groups of timing excision profiles defined in [4] for *PGM*, *EZL1*, *GTSF1* [5]*, FIREFLY* or *TFIIS4* KD. "All" is the random expectation for all IESs.

E. Violin plot superimposed with a boxplot of normalized coverage of MAC regions (top) or MIC-limited regions (bottom) using RPKM of 1-kb windows of the MIC assembly for the following samples: control (CTL, *ICL7* KD), MIC [3], *EZL1*, *PGM*, *FIREFLY*, *TFIIS4* and *SLEEPY* KDs. All samples except MIC are from the new MAC sorting shown in panel A. The coverage of MAC-destined regions is similarly covered in all datasets.

F. Heatmap of normalized TE DNA coverage on the MAC [6] and MIC [3] genomes and upon *ICL7* (CTL), *PGM*, *EZL1*, *FIREFLY*, *SLEEPY*, *TFIIS4* or *ISWI1* [7] KD as well as from flow cytometry sorted new MACs at T=30h after the onset of autogamy in control condition (T30_Kp) [4]. The TE copies are ordered by hierarchical clustering of DNA coverage in each family. The coverage distribution (RPKM) for all TE copies is shown as a boxplot on top (log_2 scale). The box shows the second and third quartiles. The median is displayed as a horizontal line. The whiskers run from the minimum to the maximum value.

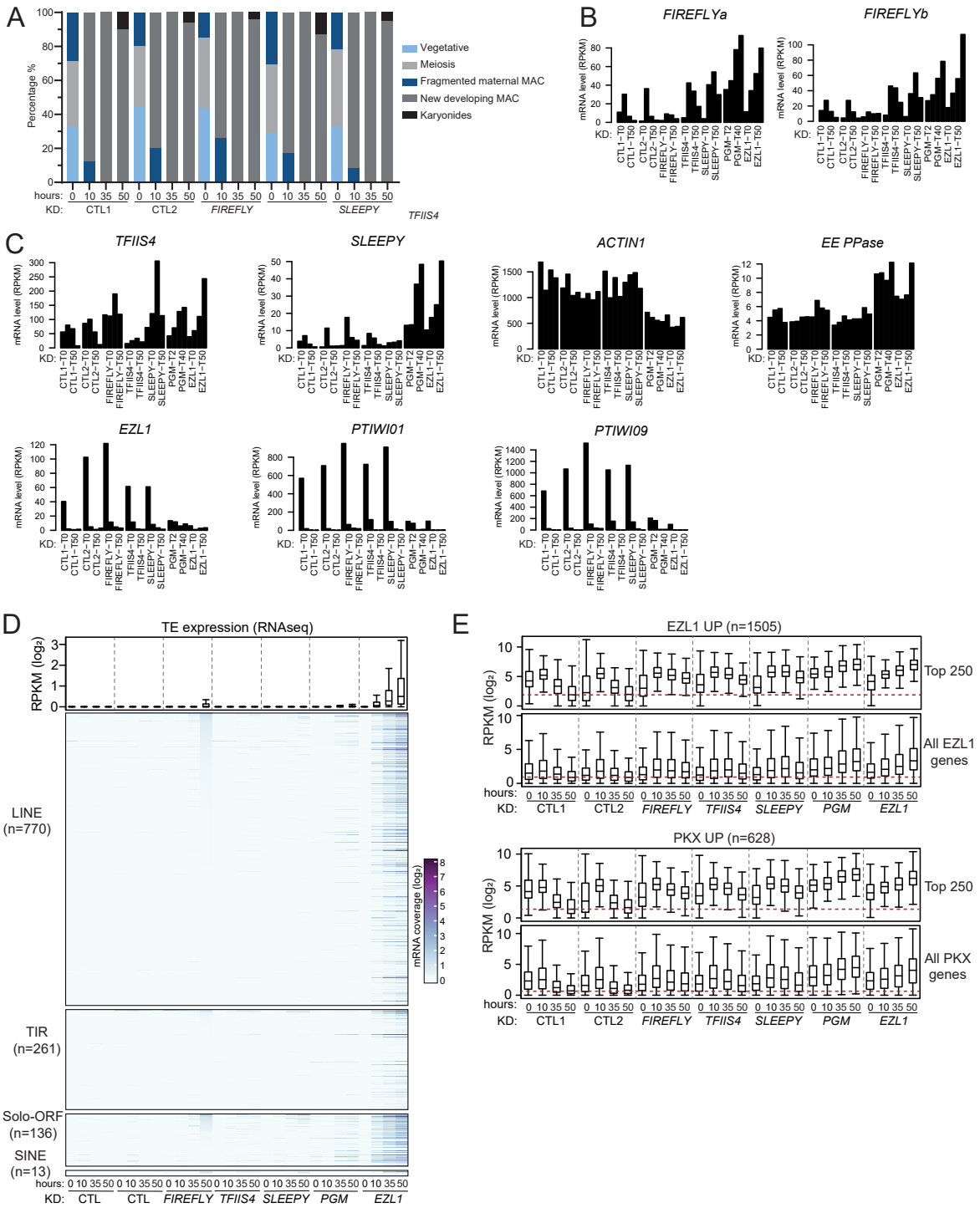


Figure S7: related to figure 3

A. Cytology of the cell populations used for RNA-seq. At least 100 cells were analyzed at each time point. Proportion of cells at each developmental stage are shown (%).

B-C. Barplots of *FIREFLY-a*, *FIREFLY-b*, *TFIIS4*, *SLEEPY*, *ACTIN1* and *EEPPase* mRNA levels upon *ICL7* (CTL1 and CTL2) and *FIREFLY*, *TFIIS4*, *SLEEPY*, *PGM* [8] or *EZL1* [9] KD over time-course RNA-seq experiment with four time points: T=0, T=10, T=35 and T=50 hours after the onset of

autogamy. Only T=0 and T=50 hours are highlighted in the legends but all time-course points are displayed.

D. Heatmap of RNA expression levels at different time points during development in each KD. TE copies are ordered by the mean RNA coverage in *FIREFLY* KD in each family (LINE n=770, TIR n=261, SOLO ORF n=136 and SINE n=13). The coverage distribution (RPKM log2) for all TE copies is shown as a boxplot. The box shows the second and third quartiles. The median is displayed as a horizontal line. The outliers are not drawn and the whiskers run from the minimum to the maximum value.

E. Boxplots of normalized gene RNA coverage for up-regulated genes upon in *EZL1* (n=1505) [9] or *PGM-KU80C-XRCC4* KDs (PKX, n=628) [8]. For each dataset, the global mRNA-seq coverage distribution (RPKM log2) is shown as a boxplot for all genes or for the top 250 most deregulated genes in the same conditions. The box shows the second and third quartiles. The median is plotted as a horizontal line. The outliers are not drawn and the whiskers run from the minimum to the maximum value.

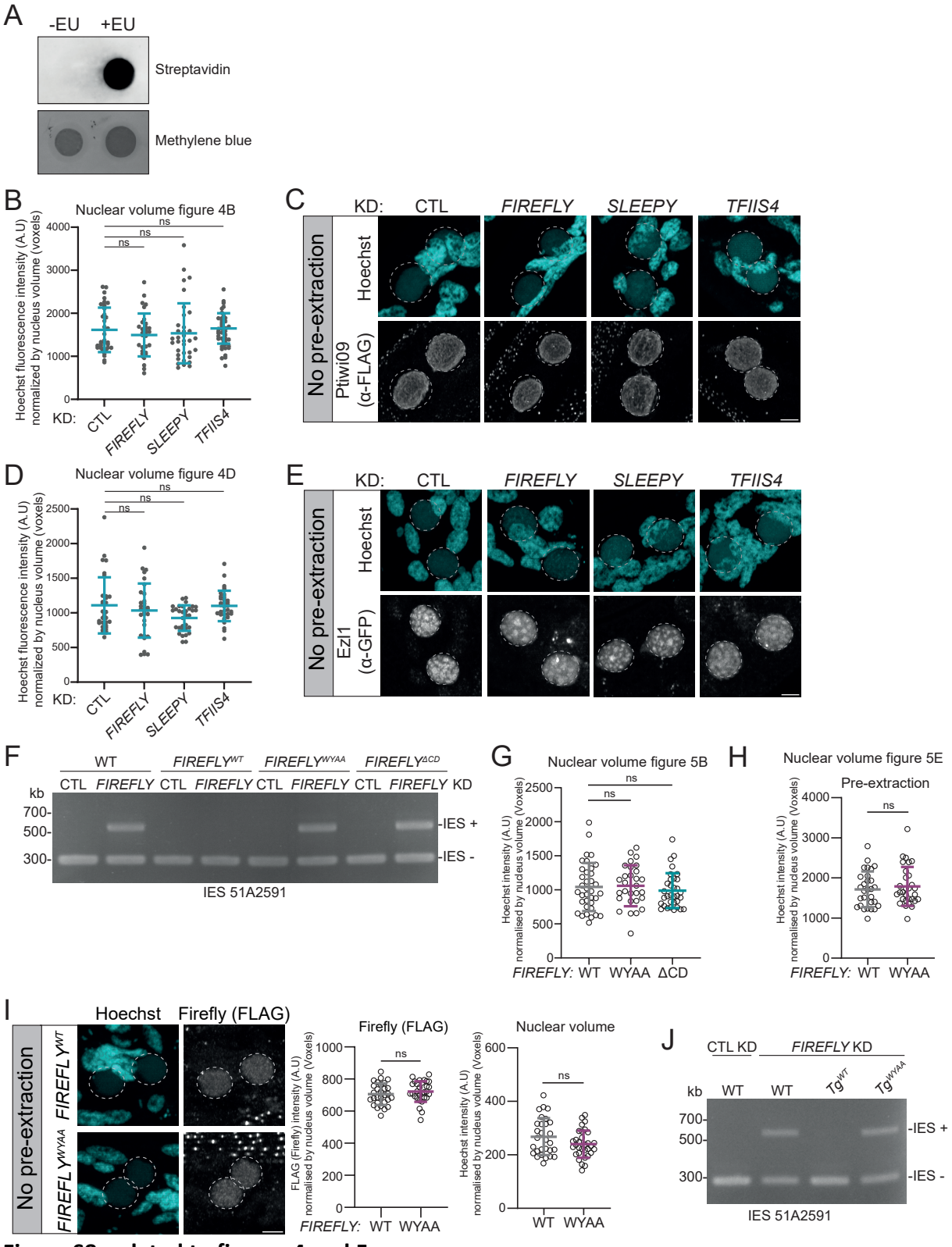


Figure S8: related to figures 4 and 5

A. Control for EU incorporation into RNA by dot blot of biotinylated RNA using HRP-conjugated streptavidin (streptavidin). Staining with methylene blue serves as a loading control.

B. Boxplot of estimated nucleus (developing MAC) volume in voxels from the same data as in Figure 4B. Estimations of nuclear volume indicate that *ICL7* (CTL), *FIREFLY*, *SLEEPY* or *TFIIS4* KD cell populations were at comparable developmental stages. Bars correspond to mean +/- SD. Mann-Whitney statistical test. n.s: non-significant

C. Immunostaining of Ptiwi09 (FLAG) in wild-type cells at T=10 hours after the onset of autogamy upon *ICL7* (CTL), *FIREFLY*, *SLEEPY* or *TFIIS4* KD when fixation is done before permeabilization (no pre-extraction) (see Methods). Scale bar is 5 µm.
D. Boxplot of estimated nucleus (developing MAC) volume in voxels from the same data as in Figure 4D. Estimations of nuclear volume indicate that *ICL7* (CTL), *FIREFLY*, *SLEEPY* or *TFIIS4* KD cell populations were at developmental comparable stages. Bars correspond to mean +/- SD. Mann-Whitney statistical test. n.s: non-significant
E. Immunostaining of Ezl1 (GFP) on wild-type cells at T=10 hours after the onset of autogamy upon *ICL7* (CTL), *FIREFLY*, *SLEEPY* or *TFIIS4* KD when fixation is done before permeabilization (no pre-extraction) (see Methods). Scale bar is 5 µm.
F. PCR analysis of the elimination of IES 51A2591 located in the A^{51} surface antigen gene using primers located outside the IES. The larger fragment corresponds to the non-excised form (IES+), the smaller fragment to the excised form (IES-). Corresponds to the experiment shown in Figure 5A-C.
G. Boxplot of estimated nucleus (developing MAC) volume in voxels from the same data as in Figure 5B. Estimations of nuclear volume indicate that $FIREFLY^{WT}$, $FIREFLY^{WYAA}$ and $FIREFLY^{\Delta CD}$ cell populations were at developmental comparable stages. Bars correspond to mean +/- SD. Mann-Whitney statistical test. n.s: non-significant
H. Boxplot of estimated nucleus (developing MAC) volume in voxels from the same data as in Figure 5E (pre-extraction). Estimations of nuclear volume indicate that $FIREFLY^{WT}$and $FIREFLY^{WYAA}$ cell populations were at developmental comparable stages. Bars correspond to mean +/- SD. Mann-Whitney statistical test. n.s: non-significant.
I. Immunostaining of cells expressing a *FLAG-HA-*$FIREFLY^{WT}$ transgene and cells expressing a *FLAG-HA-*$FIREFLY^{WYAA}$ transgene at T=10 hours after the onset of autogamy upon *FIREFLY* KD. Cells were fixed before permeabilization (no pre-extraction) (see Methods). Developing new MACs are circled with a dotted line. Scale bar, 5 µm. Quantification of $Firefly^{WT}$ and $Firefly^{WYAA}$ (FLAG) fluorescence signal in the new MAC (Methods). Estimations of nuclear volume indicate that $FIREFLY^{WT}$and $FIREFLY^{WYAA}$ cell populations were at developmental comparable stages. Bars correspond to mean +/- SD. Mann-Whitney statistical test. n.s: non-significant.
J. PCR analysis of the elimination of IES 51A2591 located in the A^{51} surface antigen gene using primers located outside the IES. The larger fragment corresponds to the non-excised form (IES+), the smaller fragment to the excised form (IES-). Corresponding to the experiment shown in Figure 5F-G.

Experiment n°	Transgene	KD gene	sexual progeny	
1- Firefly ChIP (Figure 1F-G, S3)	*3xFLAG-HA-FIREFLY* ^{WT}	*ICL7*	Alive	30
			Dead	0
		PGM	Alive	1
			Dead	29
		EZL1	Alive	0
			Dead	30
	3xFLAG-HA-FIREFLY ^{WT}	*ICL7*	Alive	30
			Dead	0
		PGM	Alive	0
			Dead	30
		EZL1	Alive	1
			Dead	29
2- Production of sexual progeny (Figure 2E, S4C) Firefly localization in SLEPEY and TFIIS4 KDs	-	*ICL7*	Alive	30
			Dead	0
	3xFLAG-HA-FIREFLY ^{WT}	*ICL7*	Alive	30
			Dead	0
		TFIIS4	Alive	29
			Dead	1
		SLEEPY	Alive	30
			Dead	0
	3xFLAG-HA-FIREFLY ^{WT}	*ICL7*	Alive	30
			Dead	0
		TFIIS4	Alive	29
			Dead	1
		SLEEPY	Alive	30
			Dead	0
3- Production of sexual progeny (Figure 2F, S4C) Sleepy localization in TFIIS4 and FIREFLY KDs	-	*ICL7*	Alive	30
			Dead	0
	3xFLAG-HA-SLEEPY	*ICL7*	Alive	30
			Dead	0
		FIREFLY	Alive	29
			Dead	1
		TFIIS4	Alive	30
			Dead	0
		SLEEPY	Alive	30
			Dead	0
	3xFLAG-HA-SLEEPY	*ICL7*	Alive	30
			Dead	0
		FIREFLY	Alive	29
			Dead	1
		TFIIS4	Alive	30
			Dead	0
		SLEEPY	Alive	29
			Dead	1
	-	*ICL7*	Alive	30
			Dead	0
	GFP-TFIIS4	*ICL7*	Alive	30
			Dead	0

4- Production of sexual progeny (Figure 2G, S4C) TfIIs4 localization in FIREFLY and SLEEPY KDs	*GFP-TFIIS4*	*FIREFLY*	Alive	29
			Dead	1
		SLEEPY	Alive	30
			Dead	0
		ICL7	Alive	30
			Dead	0
		FIREFLY	Alive	29
			Dead	1
		SLEEPY	Alive	30
			Dead	0
5-Production of viable sexual progeny (Figure 3A, S5A, B, C) H3K27me3 and H3K9me3 accumulation	-	*ICL7 CTL1*	Alive	30
			Dead	0
		ICL7 CTL2	Alive	30
			Dead	0
		FIREFLY	Alive	0
			Dead	30
		TFIIS4	Alive	0
			Dead	30
		SLEEPY	Alive	1
			Dead	29
		EZL1	Alive	0
			Dead	30
6-Production of viable sexual progeny (Figure 3B, C, S3, S5D) H3K27me3 ChIP	-	*PGM*	Alive	0
			Dead	30
		EZL1	Alive	0
			Dead	30
		FIREFLY	Alive	0
			Dead	29
		TFIIS4	Alive	0
			Dead	30
		SLEEPY	Alive	0
			Dead	30
7-Production of viable sexual progeny (Figure 3B, E, D, S6) DNAseq after KDs	-	*ICL7*	Alive	30
			Dead	0
		PGM	Alive	0
			Dead	30
		EZL1	Alive	0
			Dead	30
		FIREFLY	Alive	0
			Dead	29
		TFIIS4	Alive	0
			Dead	30
		SLEEPY	Alive	0
			Dead	30
8- Production of viable sexual progeny (Figure 4A) EU incorporation	-	*ICL7*	Alive	30
			Dead	0
		FIREFLY	Alive	0
			Dead	30
		TFIIS4	Alive	0
			Dead	30

		SLEEPY	Alive	0	
			Dead	30	
		EZL1	Alive	0	
			Dead	30	
9-Production of viable sexual progeny (Figure 4B, C, S8B, C) Ptiwi09 localization in pre-extraction conditions after KDs	-	*ICL7*	Alive	30	
			Dead	0	
	3xFLAG-HA-PTIWI09	*ICL7*	Alive	30	
			Dead	0	
		FIREFLY	Alive	0	
			Dead	30	
		TFIIS4	Alive	0	
			Dead	30	
		SLEEPY	Alive	0	
			Dead	30	
10-Production of viable sexual progeny (Figure 4D, E, S8D, E) Ezl1 localization in pre-extraction conditions after KDs	-	*ICL7*	Alive	30	
			Dead	0	
	GFP-EZL1	*ICL7*	Alive	30	
			Dead	0	
		FIREFLY	Alive	0	
			Dead	30	
		TFIIS4	Alive	0	
			Dead	30	
		SLEEPY	Alive	0	
			Dead	30	
11 - Production of sexual progeny (Figure 5A, B, C, S8F, G) FIREFLY^{WT}, *FIREFLY*^{WYAA} and *FIREFLY*$^{\Delta CD}$ complementation	-	*ICL7*	Alive	90	n=3
			Dead	0	
		FIREFLY	Alive	4	n=3
			Dead	86	
	3xFLAG-HA-FIREFLY^{WT}	*ICL7*	Alive	90	n=3
			Dead	0	
		FIREFLY	Alive	86	n=3
			Dead	4	
	3xFLAG-HA-FIREFLY^{WYAA}	*ICL7*	Alive	90	n=3
			Dead	1	
		FIREFLT	Alive	7	n=3
			Dead	83	
	3xFLAG-HA-FIREFLY$^{\Delta CD}$	*ICL7*	Alive	90	n=3
			Dead	0	
		FIREFLY	Alive	6	n=3
			Dead	84	
12 - Production of sexual progeny (Figure 5D, E, S8H, I) *FIREFLY*^{WT}, *FIREFLY*^{WYAA} SNE and pre-extraction	-	*ICL7*	Alive	30	
			Dead	0	
		FIREFLY	Alive	1	
			Dead	29	
	3xFLAG-HA-FIREFLY^{WT}	*ICL7*	Alive	30	
			Dead	0	
		FIREFLY	Alive	29	
			Dead	1	

	3xFLAG-HA-FIREFLY ^{WYAA}	*ICL7*	Alive	30	
			Dead	0	
		FIREFLY	Alive	2	
			Dead	28	
13 - Production of sexual progeny (Figure 5F, G, S8J) *FIREFLY* ^{WT} , *FIREFLY* ^{WYAA} EU incoporation	-	*ICL7*	Alive	30	
			Dead	0	
		FIREFLY	Alive	0	
			Dead	30	
	3xFLAG-HA-FIREFLY ^{WT}	*ICL7*	Alive	60	n=2
			Dead	0	
		FIREFLY	Alive	59	n=2
			Dead	1	
	3xFLAG-HA-FIREFLY ^{WYAA}	*ICL7*	Alive	60	n=2
			Dead	0	
		FIREFLY	Alive	4	n=2
			Dead	56	
14-Production of viable sexual progeny (Figure S7) RNAseq after KDs	-	*ICL7 CTL1*	Alive	30	
			Dead	0	
		ICL7 CTL2	Alive	30	
			Dead	0	
		PGM	Alive	0	
			Dead	30	
		EZL1	Alive	0	
			Dead	30	
		SLEEPY	Alive	1	
			Dead	29	
		TFIIS4	Alive	0	
			Dead	30	
		COCO	Alive	0	
			Dead	30	

Table S1: Production of sexual progeny following RNAi-mediated gene silencing of experiments used in this study. In each RNAi experiment, the number of cells that survived or died is indicated.

Name	Sequence (5' to 3')	Locus (ID)	Application	Reference
Actin_qPCR_for	TGAAGCTCCAATGAATCCAA	Actin 1-1 (PTET.51.1.G0130204)	nascent transcription qPCR	Frapporti et al. 2019
Actin_qPCR-rev	TCCTGAAGCATAGAGTGAGA	Actin 1-1 (PTET.51.1.G0130204)	nascent transcription qPCR	Frapporti et al. 2019
GAPDH_qPCR_F2	ATTTTGGTATTGTTGAGGGT	GAPDH (PTET.51.1.G0380195)	nascent transcription qPCR	Frapporti et al. 2019
GAPDH_qPCR_R2	CTCCAGTCTTTTCCACCTTT	GAPDH (PTET.51.1.G0380195)	nascent transcription qPCR	Frapporti et al. 2019
#751	CTTAGTGGGGTAGAATGAGCA	EE PPase (PTET.51.1.G1020193)	nascent transcription qPCR	Frapporti et al. 2019
#752	GACTTCTGCTTTCTTTTCTGCA	EE PPase (PTET.51.1.G1020193)	nascent transcription qPCR	Frapporti et al. 2019
#769	AGAGAGAGACTTCGTGATGA	Helicase (PTET.51.1.G1070046)	nascent transcription qPCR	Frapporti et al. 2019
#770	CAACTTGGGCATGTCAAAAT	Helicase (PTET.51.1.G1070046)	nascent transcription qPCR	Frapporti et al. 2019
#753	GGAGAGGGAAAGATAAGAGT	ST PPase (PTET.51.1.G1240023)	nascent transcription qPCR	Frapporti et al. 2019
#754	CCACTCCTTGAATTTGAGGA	ST PPase (PTET.51.1.G1240023)	nascent transcription qPCR	Frapporti et al. 2019
#551	ACAAGATTGACCAGGACTTATT	RT31010 (ms4410_NODE_3768_length_13900_cov_21.582806_RT31010_Group4_nonLTR:ClassI:LINE)	nascent transcription qPCR	Frapporti et al. 2019
#552	ATATCATTCACTCCTGCAATCT	RT31010 (ms4410_NODE_3768_length_13900_cov_21.582806_RT31010_Group4_nonLTR:ClassI:LINE)	nascent transcription qPCR	Frapporti et al. 2019
#569	TCGAAGTAGATTATGGCTGAAA	RT25765 (ms1973_NODE_322_length_47186_cov_49.739605_RT25765_Group4_nonLTR:ClassI:LINE)	nascent transcription qPCR	Frapporti et al. 2019
#570	CTTTCATCCTCCTCACTTAACT	RT25765 (ms1973_NODE_322_length_47186_cov_49.739605_RT25765_Group4_nonLTR:ClassI:LINE)	nascent transcription qPCR	Frapporti et al. 2019
#735	CCCAAGCAAAATCTGAAACA	RT43773 (ms1391_NODE_2993_length_64210_cov_44.286140_RT43773_Group2_nonLTR:ClassI:LINE)	nascent transcription qPCR	Frapporti et al. 2019
#736	CATTTTGGTGAGCACAGTTT	RT43773 (ms1391_NODE_2993_length_64210_cov_44.286140_RT43773_Group2_nonLTR:ClassI:LINE)	nascent transcription qPCR	Frapporti et al. 2019
#761	GACTATGCTGACGATCTTGT	RT48639-1 (ms6074_NODE_8863_length_4403_cov_18.113333_RT48639old_Group2_nonLTR:ClassI:LINE)	nascent transcription qPCR	Frapporti et al. 2019
#762	TTCTGATTGCCATAACACCA	RT48639-1 (ms6074_NODE_8863_length_4403_cov_18.113333_RT48639old_Group2_nonLTR:ClassI:LINE)	nascent transcription qPCR	Frapporti et al. 2019
#563	CAATTCCACCCAATATCAAACA	RT12275 (ms1017_NODE_8426_length_83307_cov_43.323395_RT12275_Group3_nonLTR:ClassI:LINE)	nascent transcription qPCR	Frapporti et al. 2019
#564	ATTAAAGGATGGTCAGAAAGCT	RT12275 (ms1017_NODE_8426_length_83307_cov_43.323395_RT12275_Group3_nonLTR:ClassI:LINE)	nascent transcription qPCR	Frapporti et al. 2019
51A2591-18	AAGTGCAACCTGTGCTGATGCTCCCGATGA	IES51A2591 (IESPGM.PTET51.1.106.284913)	PCR around IESA2591	Lhuilier-Akakpo et al. 2014
51A2591-20	AGTTCCTTTGAAAGATGTGCAAGCTCCAGA	IES51A2591 (IESPGM.PTET51.1.106.284913)	PCR around IESA2591	Lhuilier-Akakpo et al. 2014

Table S2: oligos used in this study

Sequencing	Sample	Label	ENA Accession	Reference	Number of reads	Aligned reads on the MAC		Aligned reads on the MIC	
DNAseq	KLEB	MAC	ERS452529	Lhuillier-Akakpo et al. 2014	106056122	104443929	98%	105570391	100%
DNAseq	MicGSC_BCP_AAIOSF_2_HiSeq	MIC	ERS5219632	Guérin et al. 2017	181407606	153120470	84%	179098715	99%
DNAseq	DU5_ICL7_gDNA	ICL7	ERS28489256	this study	61430914	54448970	88%	52002234	84%
DNAseq	DU6_PGM_gDNA	PGM	ERS28489257	this study	63007688	49359597	78%	54832221	87%
DNAseq	DU7_EZL1_gDNA	EZL1	ERS28489258	this study	60552522	49395892	82%	50877613	84%
DNAseq	DU8_FIREFLY_gDNA	FIREFLY	ERS28489259	this study	66666194	53502268	80%	55002835	83%
DNAseq	DU9_TFIIS4_gDNA	TFIIS4	ERS28489260	this study	60832228	50607315	83%	51816310	85%
DNAseq	DU10_SLEEPY_gDNA	SLEEPY	ERS28489261	this study	60366304	50378280	83%	50961816	84%
DNAseq	ISWI_150423_SND405_A_L001_HVA-2	ISWI1	ERS1792080	Singh et al. 2022	45711318	34250458	75%	34186389	75%
DNAseq	PTET_Nuc_T30_Kp_REGN45	T30_Kp	ERS9193172	Zangarelli et al. 2022	97332452	90725576	93%	96665100	99%
mRNAseq	PGM-T2_mRNA_CACTCA	PGM_T2	ERS14842492	Bazin-Gélis et al. 2023	81069812	77132399	95%	74844554	92%
mRNAseq	PGM-T10_mRNA_CTCAGA	PGM_T10	ERS14842490	Bazin-Gélis et al. 2023	71443788	70294491	98%	68207123	95%
mRNAseq	PGM-T30_mRNA_ATTCCT	PGM_T30	ERS14842493	Bazin-Gélis et al. 2023	80832994	75672689	94%	73986877	92%
mRNAseq	PGM-T40_mRNA_CACGAT	PGM_T40	ERS14842494	Bazin-Gélis et al. 2023	79637140	77396584	97%	76217304	96%
mRNAseq	EZL1_T0_RNA_DUHA144	EZL1_T0	ERS6679026	Miro-Pina et al. 2022	72271092	71578739	99%	70385457	97%
mRNAseq	EZL1_T10_RNA_DUHA145	EZL1_T10	ERS6679027	Miro-Pina et al. 2022	69918912	68772590	98%	68009496	97%
mRNAseq	EZL1_T35_RNA_DUHA146	EZL1_T35	ERS6679028	Miro-Pina et al. 2022	107514064	99201883	92%	104563941	97%
mRNAseq	EZL1_T50_RNA_DUHA147	EZL1_T50	ERS6679029	Miro-Pina et al. 2022	95433848	86764000	91%	92662006	97%
mRNAseq	DUHA265_ICL7_T0	ICL7_T0	ERS28489262	this study	50542382	45635210	90%	46505651	92%
mRNAseq	DUHA266_ICL7_T10	ICL7_T10	ERS28489263	this study	45188090	40424146	89%	41277434	91%
mRNAseq	DUHA267_ICL7_T35	ICL7_T35	ERS28489264	this study	49399620	43654206	88%	45066031	91%
mRNAseq	DUHA268_ICL7_T50	ICL7_T50	ERS28489265	this study	50814720	44907078	88%	46494046	91%
mRNAseq	DUHA269_TFIIS4_T0	TFIIS4_T0	ERS28489266	this study	44955080	40081213	89%	40981675	91%
mRNAseq	DUHA270_TFIIS4_T10	TFIIS4_T10	ERS28489267	this study	55971502	50546886	90%	51470149	92%
mRNAseq	DUHA271_TFIIS4_T35	TFIIS4_T35	ERS28489268	this study	48272202	43000896	89%	44341821	92%
mRNAseq	DUHA272_TFIIS4_T50	TFIIS4_T50	ERS28489269	this study	44127070	38856028	88%	40282737	91%
mRNAseq	DUHA273_SLEEPY_T0	SLEEPY_T0	ERS28489270	this study	40434730	36281627	90%	36965474	91%
mRNAseq	DUHA274_SLEEPY_T10	SLEEPY_T10	ERS28489271	this study	50119262	44959227	90%	45975915	92%
mRNAseq	DUHA275_SLEEPY_T35	SLEEPY_T35	ERS28489272	this study	46100628	41110237	89%	42454005	92%
mRNAseq	DUHA276_SLEEPY_T50	SLEEPY_T50	ERS28489273	this study	38903208	34342742	88%	35603292	92%
mRNAseq	DUHA277_ICL7-2_T0	ICL7-2_T0	ERS28489274	this study	47411806	42096984	89%	43259931	91%
mRNAseq	DUHA278_ICL7-2_T10	ICL7-2_T10	ERS28489275	this study	44956326	39824950	89%	40908773	91%
mRNAseq	DUHA279_ICL7-2_T35	ICL7-2_T35	ERS28489276	this study	28629290	24554280	86%	25883644	90%
mRNAseq	DUHA280_ICL7-2_T50	ICL7-2_T50	ERS28489277	this study	46340788	39782182	86%	41906994	90%

mRNAseq	DUHA281_FIREFLY_T0	FIREFLY_T0	ERS28489278	this study	49944358	44605862	89%	45693958	91%
mRNAseq	DUHA282_FIREFLY_T10	FIREFLY_T10	ERS28489279	this study	38370624	34585428	90%	35314445	92%
mRNAseq	DUHA283_FIREFLY_T35	FIREFLY_T35	ERS28489280	this study	49730040	43927368	88%	45431401	91%
mRNAseq	DUHA284_FIREFLY_T50	FIREFLY_T50	ERS28489281	this study	51939262	45917513	88%	47541799	92%
ChIPseq	DU26_Firefly_PGM_inp	Firefly_PGM_KD_input	ERS28489250	this study	11133128	6442310	58%	6226659	56%
ChIPseq	DU27_Firefly_PGM_rep1	Firefly_PGM_KD_rep1	ERS28489251	this study	33378482	19182467	57%	15765397	47%
ChIPseq	DU28_Firefly_PGM_rep2	Firefly_PGM_KD_rep2	ERS28489252	this study	76972152	51785689	67%	41364401	54%
ChIPseq	DU30_Firefly_EZL1_inp	Firefly_EZL1_KD_input	ERS28489253	this study	17402300	10070702	58%	9723918	56%
ChIPseq	DU31_Firefly_EZL1_rep1	Firefly_EZL1_KD_rep1	ERS28489254	this study	37415206	21120159	56%	20907651	56%
ChIPseq	DU32_Firefly_EZL1_rep2	Firefly_EZL1_KD_rep2	ERS28489255	this study	34055596	16576798	49%	16509801	48%
ChIPseq	PGM_K27_Input_DUHA306_S5_all	Input_K27_PGM	ERS28489282	this study	89330808	54335090	61%	55211367	62%
ChIPseq	PGM_K27_rep3_DUHA309_S8_all	K27_PGM_rep3	ERS28489283	this study	102040016	47573272	47%	53460538	52%
ChIPseq	EZL1_K27_Input_DUHA310_S9_all	Input_K27_EZL1	ERS28489284	this study	90497198	53121991	59%	53427900	59%
ChIPseq	EZL1_K27_rep1_DUHA311_S10_all	K27_EZL1_rep1	ERS28489285	this study	93375436	57544000	62%	57501114	62%
ChIPseq	EZL1_K27_rep2_DUHA312_S11_all	K27_EZL1_rep2	ERS28489286	this study	99288116	59939428	60%	59799413	60%
ChIPseq	EZL1_K27_rep3_DUHA313_S12_all	K27_EZL1_rep3	ERS28489287	this study	102022240	58452231	57%	58449015	57%
ChIPseq	FIREFLY_K27_Input_DUHA314_S13_all	Input_K27_FIREFLY	ERS28489288	this study	71426034	37440438	52%	37413022	52%
ChIPseq	FIREFLY_K27_rep1_DUHA315_S14_all	K27_FIREFLY_rep1	ERS28489289	this study	104649542	62400088	60%	66418266	63%
ChIPseq	FIREFLY_K27_rep2_DUHA316_S15_all	K27_FIREFLY_rep2	ERS28489290	this study	91873094	57047761	62%	60358051	66%
ChIPseq	FIREFLY_K27_rep3_DUHA317_S16_all	K27_FIREFLY_rep3	ERS28489291	this study	81092346	36597546	45%	37796363	47%
ChIPseq	TFIIS4_K27_Input_DUHA318_S17_all	Input_K27_TFIIS4	ERS28489292	this study	87102112	53064437	61%	54096834	62%
ChIPseq	TFIIS4_K27_rep1_DUHA319_S18_all	K27_TFIIS4_rep1	ERS28489293	this study	81026436	44815548	55%	53819109	66%
ChIPseq	TFIIS4_K27_rep2_DUHA320_S19_all	K27_TFIIS4_rep2	ERS28489294	this study	106480112	52027381	49%	66203645	62%
ChIPseq	TFIIS4_K27_rep3_DUHA321_S20_all	K27_TFIIS4_rep3	ERS28489295	this study	97620370	49741083	51%	62343205	64%
ChIPseq	SLEEPY_K27_Input_DUHA322_S21_all	Input_K27_SLEEPY	ERS28489296	this study	99398696	59480948	60%	60127733	60%
ChIPseq	SLEEPY_K27_rep1_DUHA323_S22_all	K27_SLEEPY_rep1	ERS28489297	this study	97300824	58098449	60%	59690480	61%
ChIPseq	SLEEPY_K27_rep2_DUHA324_S23_all	K27_SLEEPY_rep2	ERS28489298	this study	121435884	72061912	59%	74002259	61%
ChIPseq	SLEEPY_K27_rep3_DUHA325_S24_all	K27_SLEEPY_rep3	ERS28489299	this study	98261974	62117273	63%	64245803	65%

Table S3: Sequencing data and mapping statistics of the data used in this study. All the data generated in this study were deposited to ENA under the accession number PRJEB106600. For published datasets, the ENA accession numbers are provided. For each sequencing sample the ENA accession is specified followed by the number of reads sequenced and the number of mapped reads on the MAC and MIC reference genomes.

References

1. Arnaiz O, Van Dijk E, Bétermier M, Lhuillier-Akakpo M, de Vanssay A, Duharcourt S, et al. Improved methods and resources for paramecium genomics: transcription units, gene annotation and gene expression. BMC Genomics. 2017;18:483. https://doi.org/10.1186/s12864-017-3887-z

2. Aury J-M, Jaillon O, Duret L, Noel B, Jubin C, Porcel BM, et al. Global trends of whole-genome duplications revealed by the ciliate Paramecium tetraurelia. Nature. 2006;444:171–8. https://doi.org/10.1038/nature05230

3. Guérin F, Arnaiz O, Boggetto N, Denby Wilkes C, Meyer E, Sperling L, et al. Flow cytometry sorting of nuclei enables the first global characterization of Paramecium germline DNA and transposable elements. BMC Genomics. 2017;18:327. https://doi.org/10.1186/s12864-017-3713-7

4. Zangarelli C, Arnaiz O, Bourge M, Gorrichon K, Jaszczyszyn Y, Mathy N, et al. Developmental timing of programmed DNA elimination in Paramecium tetraurelia recapitulates germline transposon evolutionary dynamics. Genome Res. 2022;gr.277027.122. https://doi.org/10.1101/gr.277027.122

5. Charmant O, Gruchota J, Arnaiz O, Nowak KP, Moisan N, Zangarelli C, et al. The PIWI-interacting protein Gtsf1 controls the selective degradation of small RNAs in Paramecium. Nucleic Acids Research. 2024;gkae1055. https://doi.org/10.1093/nar/gkae1055

6. Lhuillier-Akakpo M, Frapporti A, Denby Wilkes C, Matelot M, Vervoort M, Sperling L, et al. Local effect of enhancer of zeste-like reveals cooperation of epigenetic and cis-acting determinants for zygotic genome rearrangements. PLoS Genet. 2014;10:e1004665. https://doi.org/10.1371/journal.pgen.1004665

7. Singh A, Maurer-Alcalá XX, Solberg T, Häußermann L, Gisler S, Ignarski M, et al. Chromatin remodeling is required for sRNA-guided DNA elimination in Paramecium. The EMBO Journal. John Wiley & Sons, Ltd; 2022;41:e111839. https://doi.org/10.15252/embj.2022111839

8. Bazin-Gélis M, Eleftheriou E, Zangarelli C, Lelandais G, Sperling L, Arnaiz O, et al. Inter-generational nuclear crosstalk links the control of gene expression to programmed genome rearrangement during the Paramecium sexual cycle. Nucleic Acids Research. 2023;51:12337–51. https://doi.org/10.1093/nar/gkad1006

9. Frapporti A, Miró Pina C, Arnaiz O, Holoch D, Kawaguchi T, Humbert A, et al. The Polycomb protein Ezl1 mediates H3K9 and H3K27 methylation to repress transposable elements in Paramecium. Nat Commun. 2019;10:2710. https://doi.org/10.1038/s41467-019-10648-5