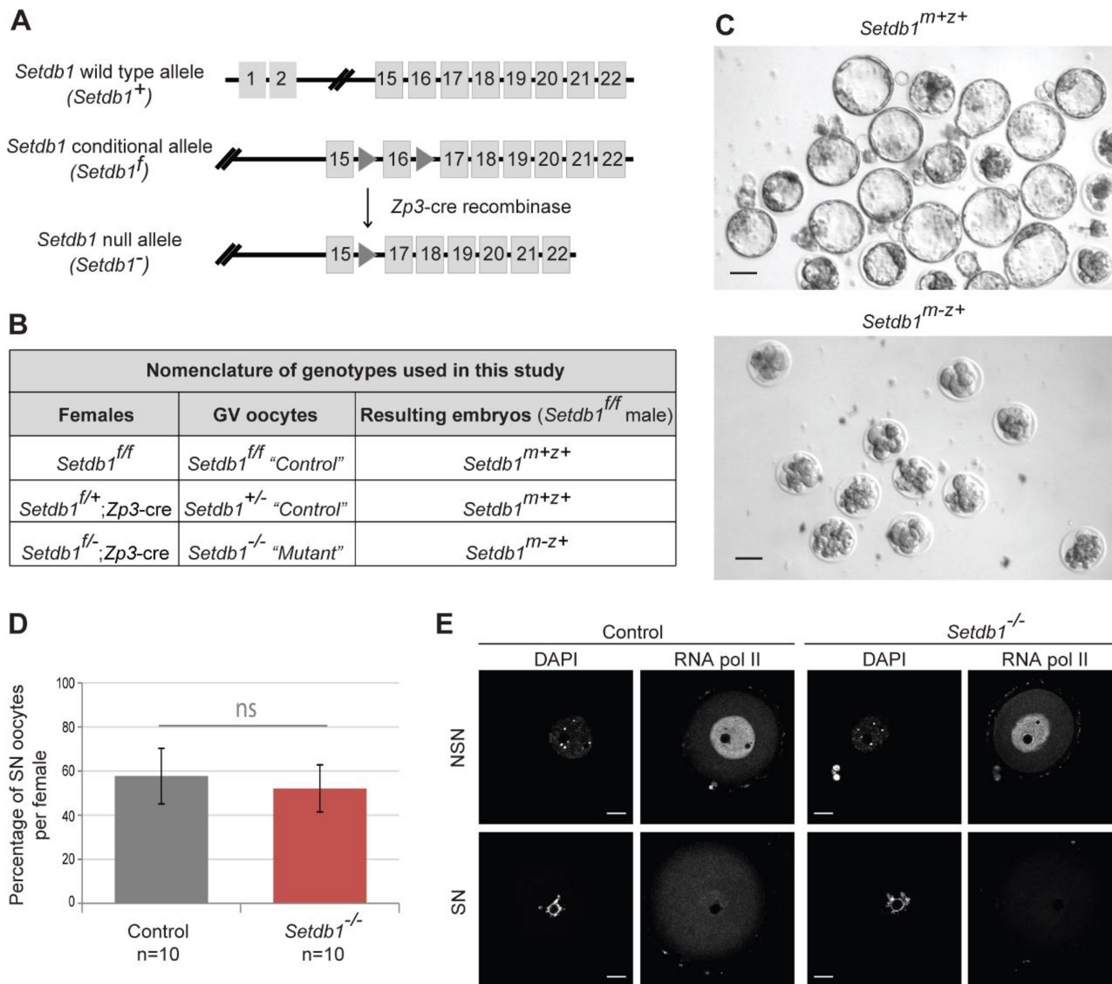
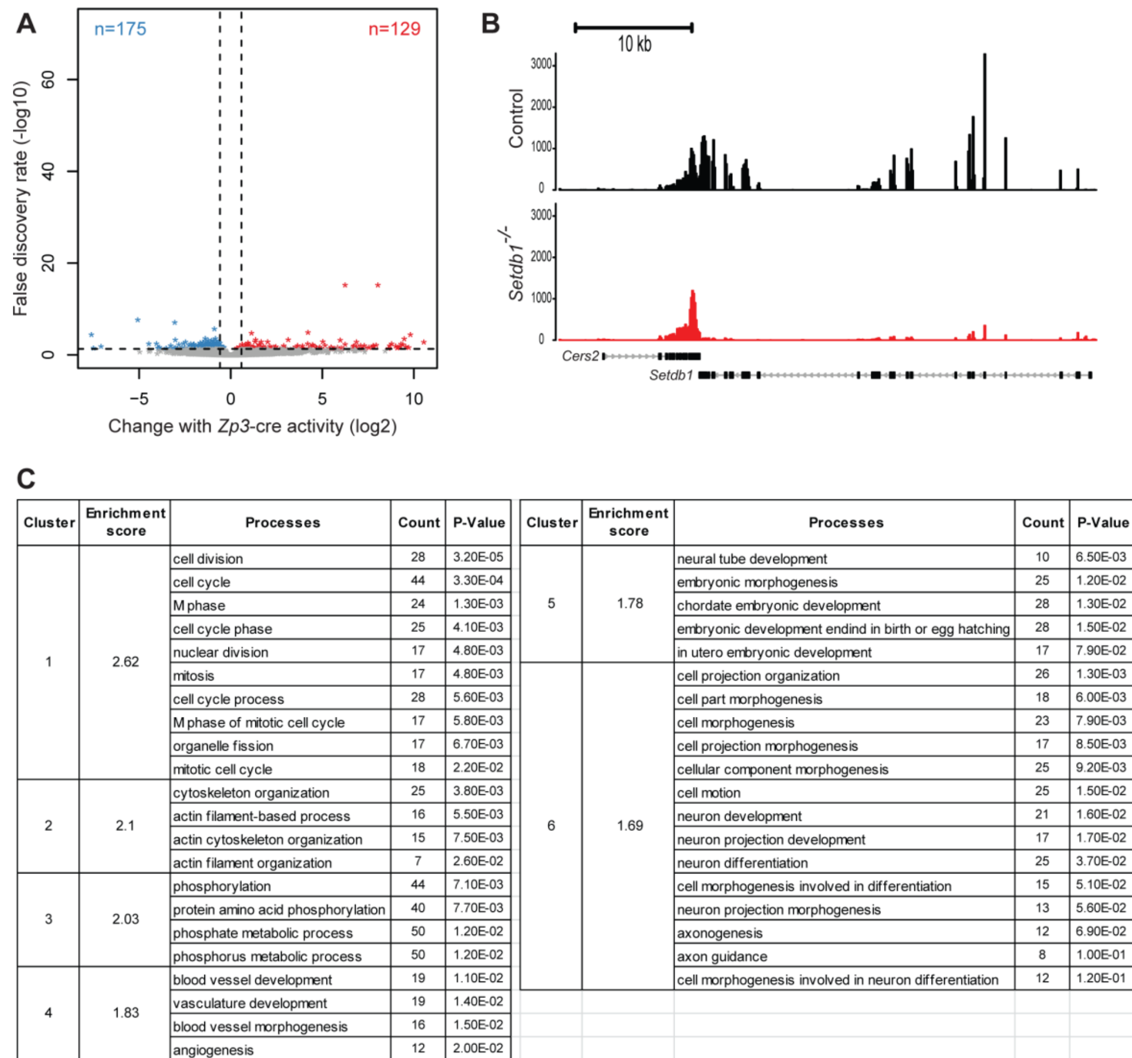


Figure S1 (related to Figure 1)

(A) Generation of *Setdb1*$^{f/f}$ mice (based on (Lohmann et al., 2010)). LoxP sites are represented as triangles. Excision of exon 16 by *ZP3*-cre mediated recombination between LoxP sites produces the *Setdb1* null allele. **(B)** Nomenclature of genotypes used in the study. **(C)** Differential interference contrast (DIC) images of E4.5 *Setdb1*$^{m+z+}$ and *Setdb1*$^{m-z+}$ *in vitro* cultured embryos. Scale bars: 50µm. **(D)** Percentage of SN-GV oocytes per female. n: number of mice analyzed. P-value was calculated using an unpaired t-test assuming equal variance; ns: non-significant. **(E)** RNA polymerase II (RNA pol II) staining on NSN and SN GV oocytes. At least 7 oocytes were analyzed for each condition. Scale bars: 20µm.

Development • Supplementary information



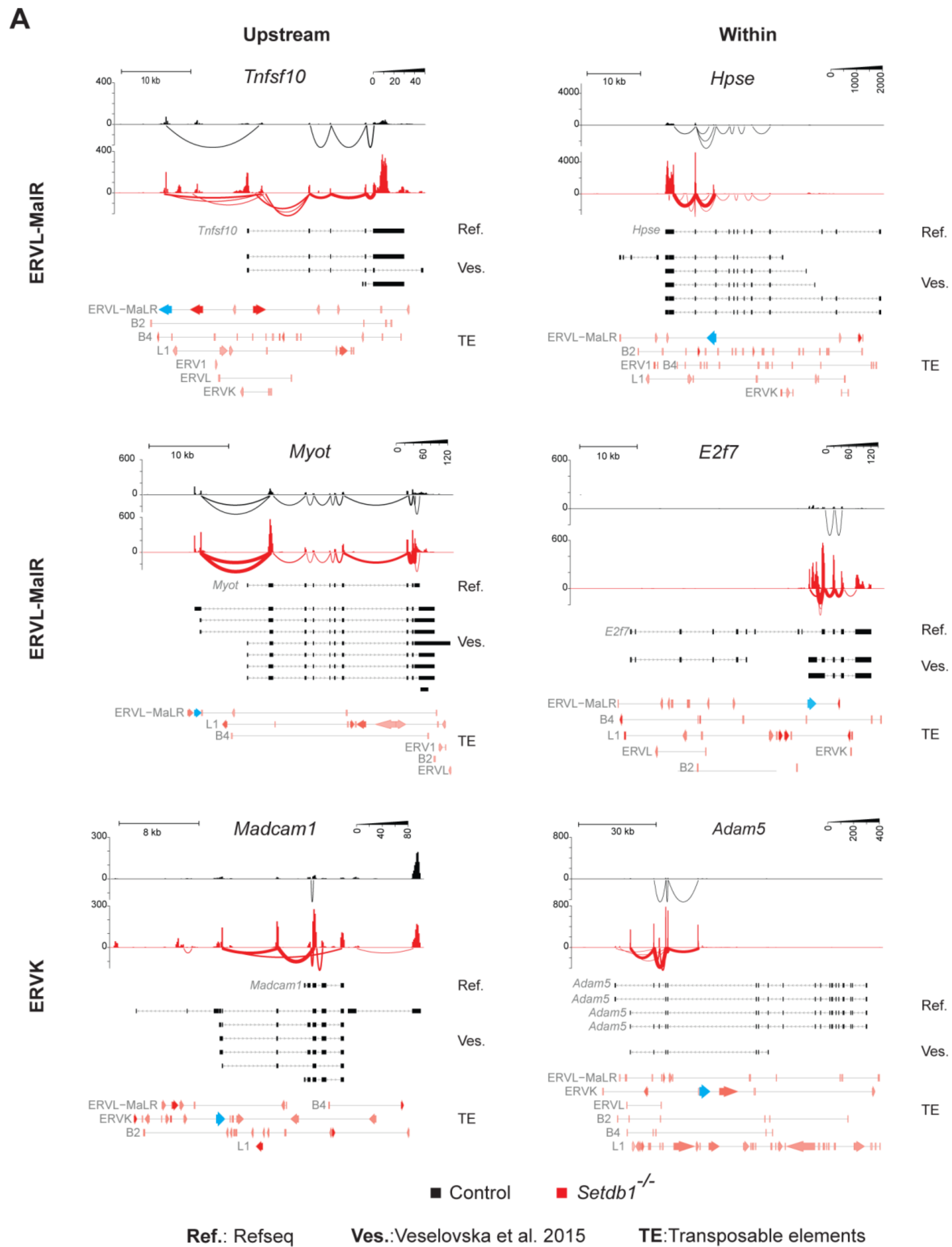
Cluster	Enrichment score	Processes	Count	P-Value
1	2.62	cell division	28	3.20E-05
		cell cycle	44	3.30E-04
		M phase	24	1.30E-03
		cell cycle phase	25	4.10E-03
		nuclear division	17	4.80E-03
		mitosis	17	4.80E-03
		cell cycle process	28	5.60E-03
		M phase of mitotic cell cycle	17	5.80E-03
		organelle fission	17	6.70E-03
		mitotic cell cycle	18	2.20E-02
2	2.1	cytoskeleton organization	25	3.80E-03
		actin filament-based process	16	5.50E-03
		actin cytoskeleton organization	15	7.50E-03
		actin filament organization	7	2.60E-02
3	2.03	phosphorylation	44	7.10E-03
		protein amino acid phosphorylation	40	7.70E-03
		phosphate metabolic process	50	1.20E-02
		phosphorus metabolic process	50	1.20E-02
4	1.83	blood vessel development	19	1.10E-02
		vasculature development	19	1.40E-02
		blood vessel morphogenesis	16	1.50E-02
		angiogenesis	12	2.00E-02

Cluster	Enrichment score	Processes	Count	P-Value
5	1.78	neural tube development	10	6.50E-03
		embryonic morphogenesis	25	1.20E-02
		chordate embryonic development	28	1.30E-02
		embryonic development endind in birth or egg hatching	28	1.50E-02
		in utero embryonic development	17	7.90E-02
6	1.69	cell projection organization	26	1.30E-03
		cell part morphogenesis	18	6.00E-03
		cell morphogenesis	23	7.90E-03
		cell projection morphogenesis	17	8.50E-03
		cellular component morphogenesis	25	9.20E-03
		cell motion	25	1.50E-02
		neuron development	21	1.60E-02
		neuron projection development	17	1.70E-02
		neuron differentiation	25	3.70E-02
		cell morphogenesis involved in differentiation	15	5.10E-02
		neuron projection morphogenesis	13	5.60E-02
		axonogenesis	12	6.90E-02
		axon guidance	8	1.00E-01
		cell morphogenesis involved in neuron differentiation	12	1.20E-01

Figure S2 (related to Figure 4)

(A) Volcano plot showing differentially expressed genes comparing *Setdb1*$^{f/+}$; *Zp3*-cre versus *Setdb1*$^{f/-}$ GV-oocytes, FDR<0.05. Blue: down-regulated genes; red: up-regulated genes. **(B)** Genome browser view of *Setdb1* locus. **(C)** Clustering of GO-terms. Count: number of genes in the cluster.

Development • Supplementary information



Development • Supplementary information

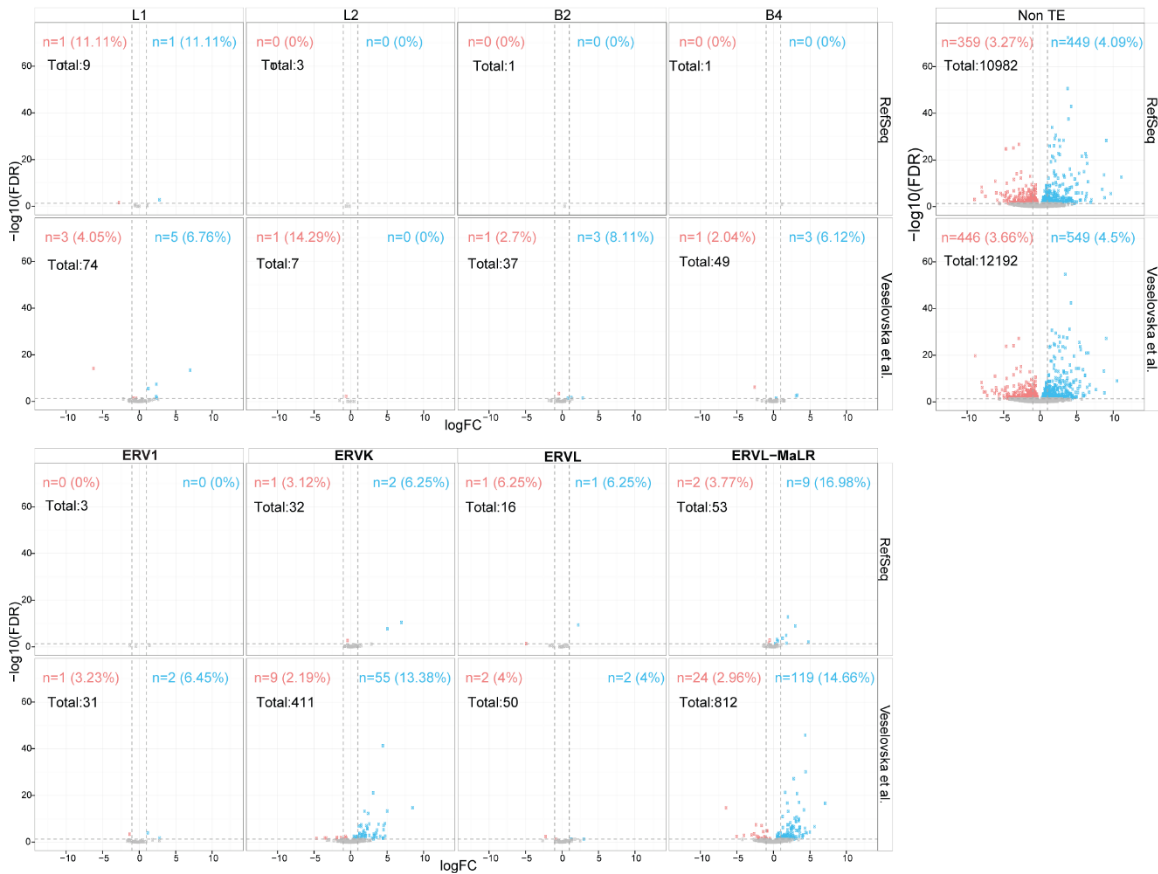


Figure S3 (related to Figure 5)

(A) Genome browser views illustrating expression along *Tnfsf10, Myot, Hpse, E2f7, Madcam1 and Adam5* loci (uniquely mapped reads per 100 base pair bins, library-size normalized) and expression at splice-junctions (uniquely mapped reads per splice-junction, library-size normalized) between ERVL-MaLR or ERVK elements and exons in control and *Setdb1*$^{-/-}$ oocytes. Splicing events occurred either with ERVs positioned upstream of or within the ORF (colored in blue), resulting in formation of chimaeric transcripts encoding for intact or likely aberrant proteins, respectively. **(B)** Volcano plots for multi-exonic transcripts with transcriptional start sites in different families of ERVs, using genome annotations based on RefSeq (top) and Veselovska et al. (Veselovska et al., 2015) (bottom).

Development • Supplementary information

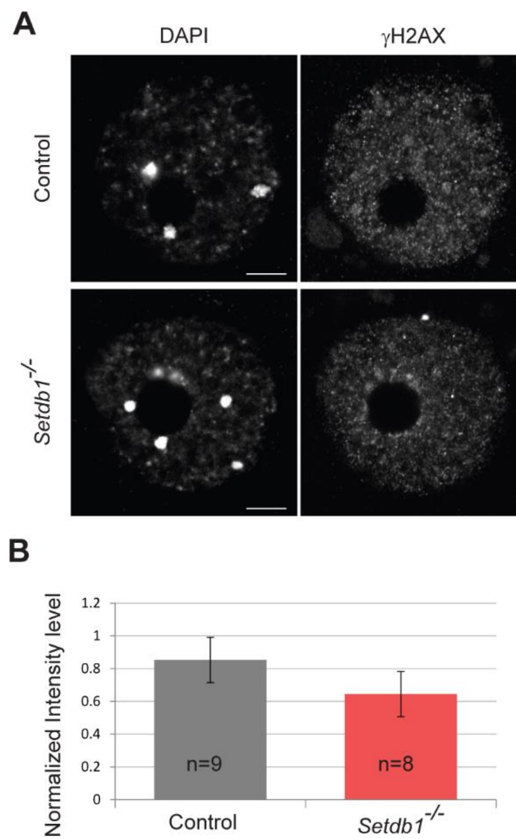


Figure S4

(A) γH2AX and DAPI staining on control and *Setdb1*$^{-/-}$ GV-oocytes. Scale bars: 20μm. **(B)** Three dimensional quantification of γH2AX levels in oocytes, normalized to nucleosomal staining. n:number of oocytes analyzed.

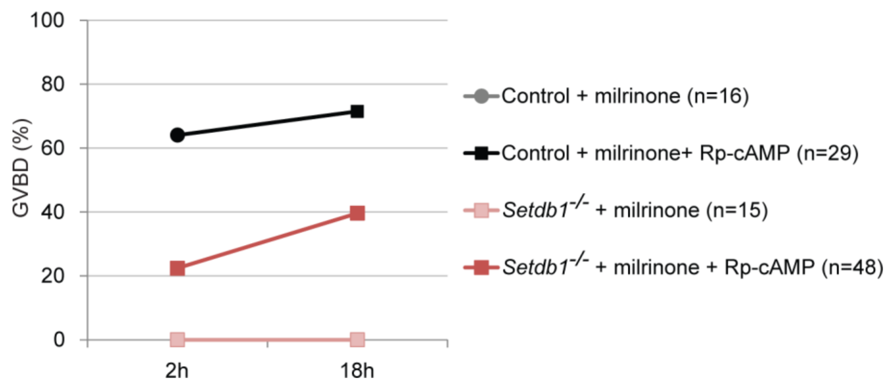
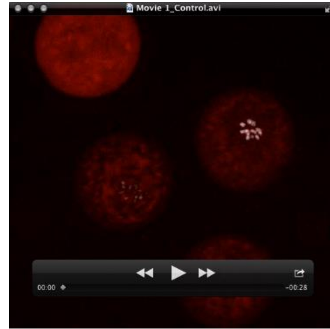


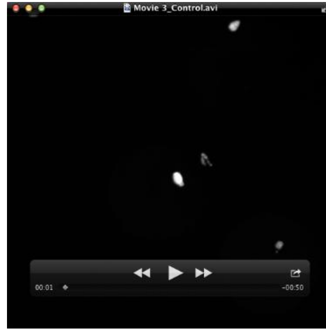
Figure S5 (related to Figure 6)

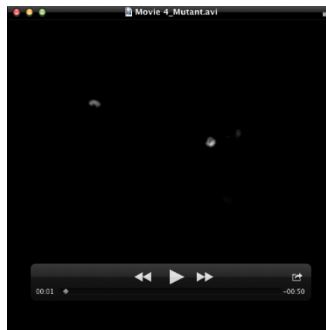
GVBD 2h and 18h after addition or in absence of Rp-cAMP in presence of milrinone.
n: number of oocytes analyzed.



Movies 1 and 2 - related to Figure 2.

GV-oocytes were micro-injected with mRNA transcripts encoding for H2B-mCherry (grey) and α-tubulin (red). Meiotic maturation was followed by live-imaging using a spinning disc confocal microscope, at 37˚C degree with a 5% CO_2 environment. Oocytes were imaged every 15 minutes.





Movies 3 and 4 - related to Figure 7.

MII-oocytes were micro-injected with mRNA transcripts encoding for H2B-mCherry (grey) followed by intra cytoplasmic sperm injection. First cleavage was followed by live-imaging using a spinning disc confocal microscope, at 37˚C degree with a 5% CO_2 environment. Embryos were imaged every 15 minutes.

Table S1 – related to Figures 4 and 5

Differential expression data of genes and splice-junctions based on RefSeq annotation. The table allows a comparison of expression of splice-junctions to chimaeric transcripts identified by Peaston et al. (Peaston et al., 2004). The table provides also the names of genes present in GO-term gene clusters (see Fig. S2C).

Click here to Download Table S1

Table S2 – related to Figure 5.

Differential expression data of ERVs grouped according to repFamily annotation (RepBase).

Click here to Download Table S2

Table S3 – related to Figure 5.

Differential expression data of ERVs grouped according to repName annotation (RepBase).

Click here to Download Table S3

Table S4 – related to Figure 5.

Numbers of splice-junctions between different types of ERVs and exons in control and *Setdb1* mutant oocytes.

Click here to Download Table S4

Table S5 – related to Figure 5.

Differential expression data of transcripts and ERVs based on Veselovska et al. (Veselovska et al., 2015) annotation. The table provides a match between data based on the Veselovska and RefSeq annotations. The table provides also the names of genes present in GO-term gene clusters (see Fig. S2C).

Click here to Download Table S5

Supplementary methods

RNA sequencing. We performed expression profiling on pools of 16 denuded GV-oocytes isolated per mouse. We used oocytes from 4 *Setdb1*$^{f/+}$*; Zp3*-cre mice and 2 *Setdb1*$^{f/-}$ mice as controls and oocytes from 4 *Setdb1*$^{f/-}$; *Zp3*-cre mice as mutant. RNA was isolated using the RNeasy Micro Kit (Qiagen). cDNA was generated and amplified from 1.2ng with the NuGen ovation RNA-seq System V2 (Part no. 7102; NuGen). 50ng of the resulting SPIA cDNA was fragmented and sequencing libraries were prepared using TruSeq DNA Sample Preparation Kit (low-throughput protocol) (Part no. 15005180 Rev. C; NuGen). Libraries were pooled equimolarly and sequenced for 50 cycles on an Illumina HiSeq 2000 instrument using RTA 1.13.48 for basecalling. Demultiplexing and fastq generation was performed with bcl2fastq (bcl2fastq-1.8.3).

Read alignment and expression quantification. The *M. musculus* genome assembly (GRCm38/mm10 Dec. 2011), RepeatMasker repeat annotation (downloaded on 7 March 2012), RefSeq gene models (downloaded from UCSC on 4 February 2016), and the oocyte transcriptome annotation (downloaded on 29 February 2016 from the web page of the paper (Veselovska et al., 2015)) were used as a basis for all analyses. Illumina adaptor sequences (GATCGGAAGAGCACACGTCTGAACTCCAGTCAC) were removed from the 3' end of RNA sequencing reads and spliced alignment of reads to the genome was created using STAR of version 2.5.0a (Dobin et al., 2013) with parameters –outFilterMultimapNmax 300 –outMultimapperOrder Random -outSAMmultNmax 1 -outSAMmapqUnique 255 –alignSJoverhangMin 8, tracking up to 300 matches and choosing only one random match for reads with multiple alignments, gapped alignments with overhang shorter than 8 bp were removed.

For quantification of mRNA uniquely aligned reads were summed per transcripts using QuasR package (version 1.10.1) (Gaidatzis et al., 2015) and for genes with multiple transcripts, the transcript with the maximal average expression over all samples was selected as the representative.

For quantification of expression of repetitive elements all reads, including multimappers, were summed per repeat names or per repeat families. For Fig. 5E only uniquely mapped reads were used (here corresponding to reads with minimum mapping quality 255).

For display in heatmaps, RPKM values were calculated (Mortazavi et al., 2008) (and log2 transformed using formula log2(RPKM + psc) – log2(psc) where pseudo-count psc was set to 0.1. Comparison between RefSeq transcriptome annotation and oocyte transcriptiome annotation was done using *cuffcompare* from Cufflinks suite (Trapnell et al., 2012).

Development • Supplementary information

Differential expression analysis. Differentially expressed genes and repeat elements were identified using R and edgeR version 3.12.0 (McCarthy et al., 2012), by fitting a two-factor model of the form “cre + genotype”, with cre corresponding to expression status of the *Zp3* cre-recombinase (“expressed” or “not expressed”), and genotype to the *Setdb1* genotype (“+/-“ or “-/-“). The cre factor was included into the model to absorb the effects of expressing *Zp3*-cre, independent of the genotype of Setdb1. Only genes with at least 3 reads per million in at least two samples were included in the analysis (11,366 for RefSeq annotation and 14,954 for oocyte specific annotation). P-values for differential expression were calculated using log-likelihood tests. Differentially expressed genes or repeat elements were defined by FDR<0.05. For Fig. S3B, RefSeq transcripts and oocyte transcripts from Veselovska et al. (2015) were classified as being initiated from a transposable element (TE) if the 5’ end of the transcript overlapped with the TE on the same strand. GO term analysis and clustering were performed using DAVID (Huang da et al., 2009a, b).

Analysis of splicing events. SJ.out.tab files produced by STAR for each sample were loaded into R environment and combined into a matrix containing the number of reads supporting each observed splicing event in each sample. Splicing events which were not supported by at least 5 reads in at least 2 samples and very long splicing events (longer than half of the average transcript length, ~47 Kbp) were removed from the analysis. Differential expression analysis was carried out similarly to the analysis for genes and repeats. Splicing events were classified “Upstream” with respect to a RefSeq transcript if the intron started upstream of the TSS and ended within the transcript. Splicing events were classified as “Within” if both intron start and end were within a transcript, and as “Downstream” if the intron started within and ended downstream of the transcript. Splicing events were classified as TE-associated if intron starts overlapped TE on any strand.

Supplementary references

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics. *29(1)*, 15-21.

Gaidatzis, D., Lerch, A., Hahne, F., and Stadler, M.B. (2015). QuasR: quantification and annotation of short reads in R. Bioinformatics *31(7)*, 1130-1132.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res *37*, 1-13.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res *40*, 4288-4297.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods *5*, 621-628.

Peaston, A.E., Evsikov, A.V., Graber, J.H., de Vries, W.N., Holbrook, A.E., Solter, D., and Knowles, B.B. (2004). Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. Developmental cell *7*, 597-606.

Veselovska, L., Smallwood, S.A., Saadeh, H., Stewart, K.R., Krueger, F., Maupetit-Mehouas, S., Arnaud, P., Tomizawa, S., Andrews, S., and Kelsey, G. (2015). Deep sequencing and de novo assembly of the mouse oocyte transcriptome define the contribution of transcription to the DNA methylation landscape. Genome biology *16*, 209.