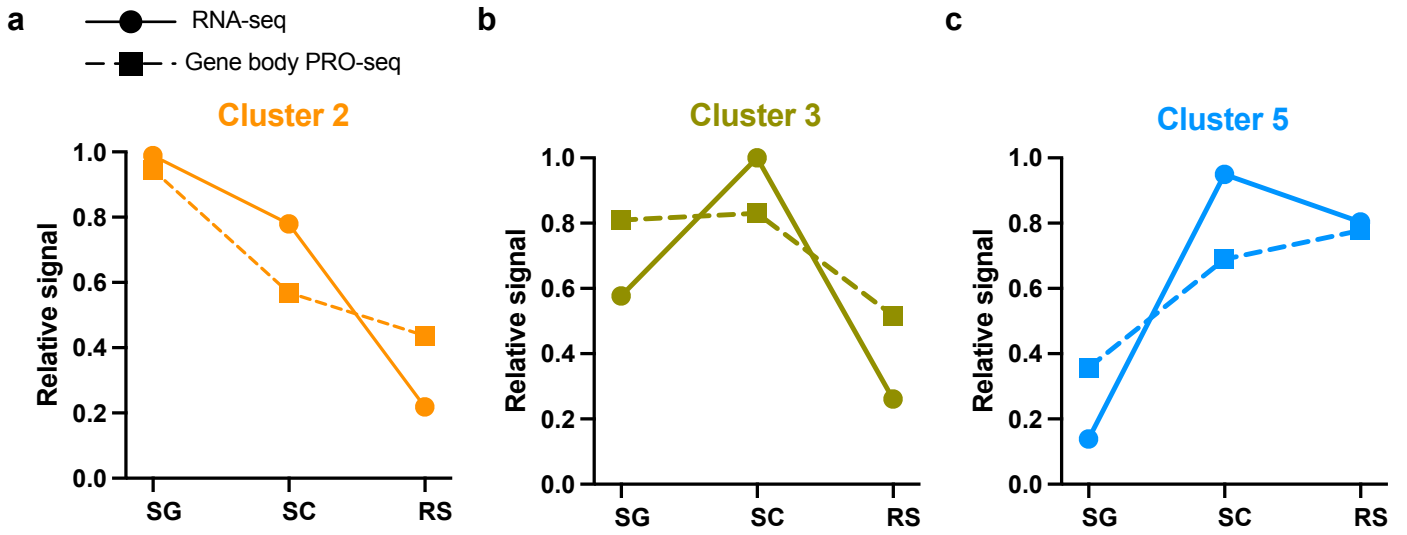


Supplementary Fig. 1. Gene ontology terms and mRNA gene lengths for clusters of differentially expressed genes.

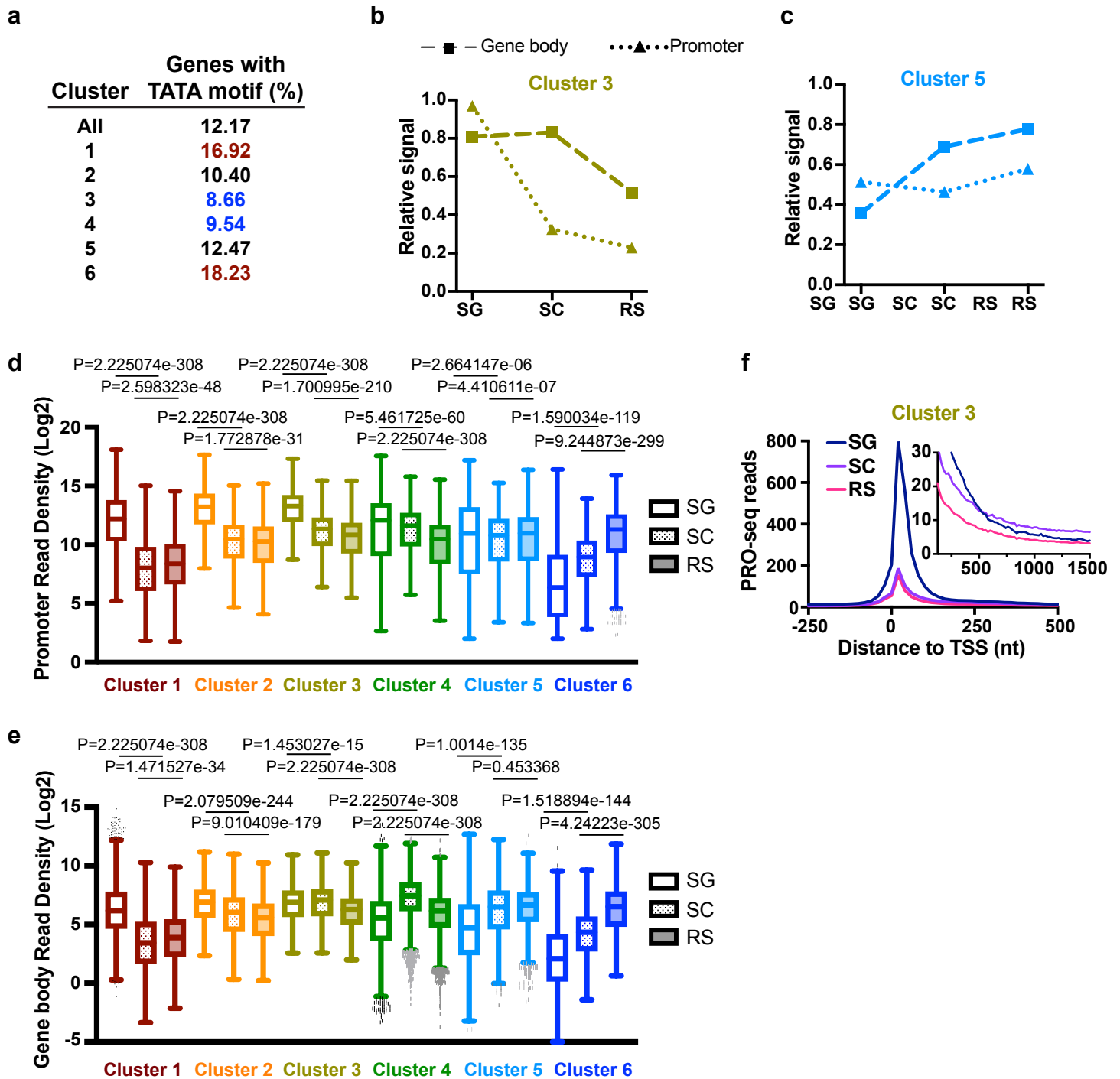
a-c. Gene ontology terms and significance values for cluster 2 (**a**), cluster 3 (**b**), and cluster 5 (**c**) generated with ClusterProfiler and consolidated using Revigo.

d. Box plot representation of the distribution of gene length (TSS to TES) for protein-coding genes in each cluster. Cluster 6 mRNAs are significantly shorter than mRNAs in all other clusters, confirming that the abundance of lncRNAs in this cluster was not biasing the length determination shown in Fig. 1k. Aside from lncRNAs, no other class or family of genes was found to be enriched in cluster 6. Line represents median, box represents 25-75th percentile, whiskers represent 1.5X interquartile range. P-values are all from Mann Whitney U unpaired, two-sided test of cluster 1 (n=3,083), 2 (n=2,533), 3 (n=3,393), 4 (n=3,267), or 5 (n=1,239) compared to cluster 6 (n=1,197).



Supplementary Fig. 2. Relative signal from RNA and PRO-seq gene body across cell types.

a-c. Plots of average relative signal from RNA-seq (solid line and circle) or PRO-seq in gene bodies (dotted line and square) for each cell type, for clusters 2 (a), 3 (b), and 5 (c).



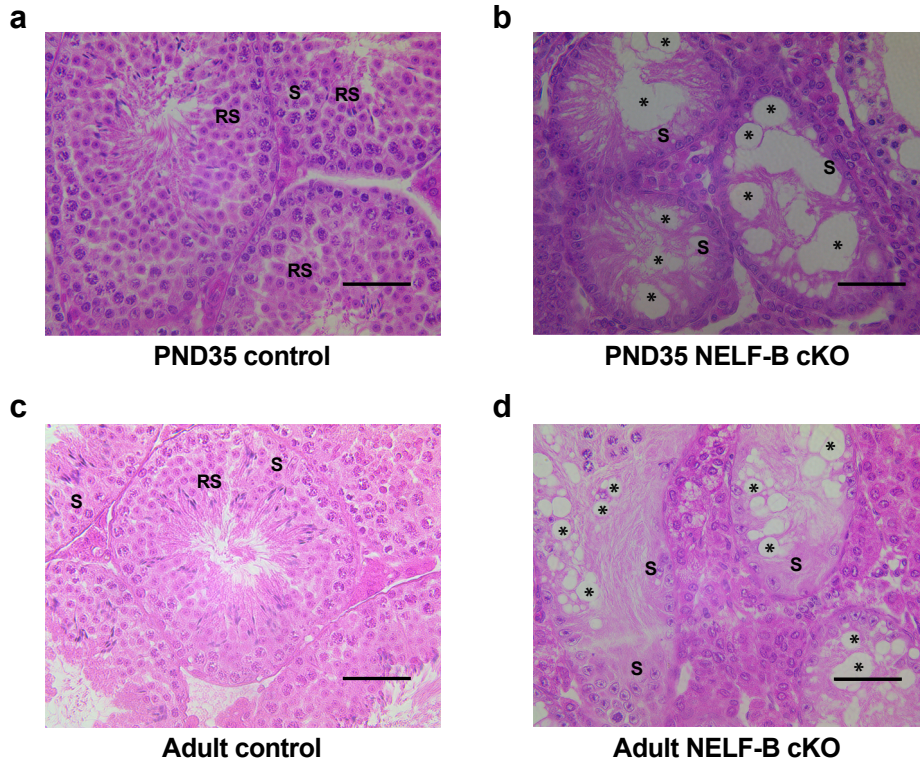
Supplementary Fig. 3 Promoter and gene body counts by cluster and cell type.

a. The proximal promoter regions (TSS-100 to TSS) for all clustered genes were evaluated for a match to the canonical TATA motif using FIMO with a threshold of $P=0.005$. The percent (%) of genes in each individual cluster is also reported. Clusters 1 and 6 have an enrichment of TATA elements, whereas Clusters 3 and 4 are depleted of TATA motifs.

b-c. Plots of average relative signal from PRO-seq within promoter (dotted line with triangles) or gene body (dashed line with squares) regions for each cell type. Shown are data for genes in clusters 3 (**b**) 5 (**c**).

d-e. Box plot representations for the distribution of reads per kb in each cluster from PRO-seq promoter, (TSS to +150, **d**) or gene body (TSS+250 to TES, **e**), for each cell type. Legend indicates SG (empty), SC (dotted), and RS (shaded). P-values are from two-sided Wilcoxon matched-pairs test.

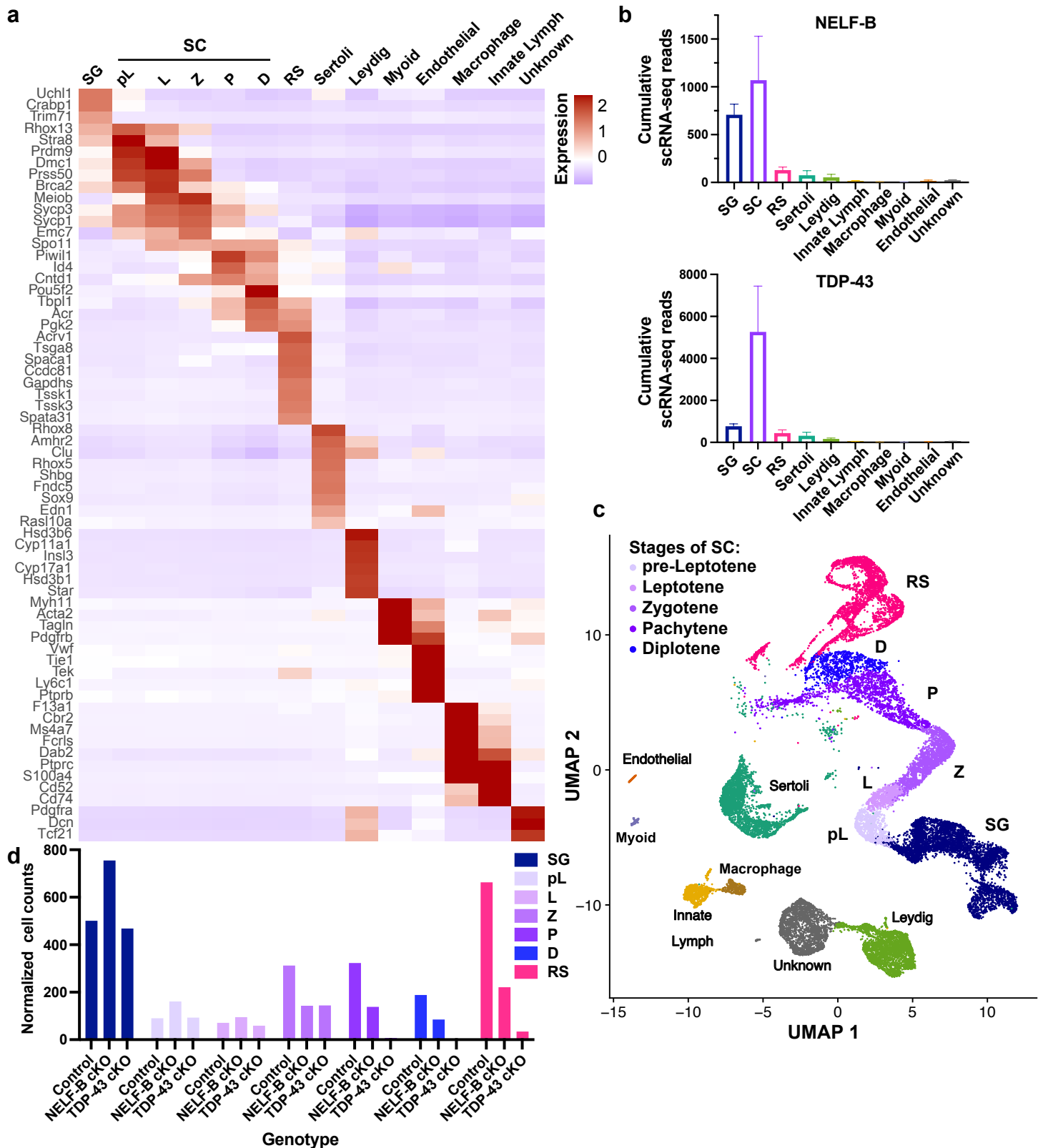
f. Metagenome plot of average PRO-seq reads in each cell type for all protein-coding genes in cluster 3. Inset y-axis is zoomed-in to show gene body signal from +125 to +1500 nt downstream of the TSS.



Supplementary Fig. 4. NELF-B is required for the successful completion of spermatogenesis.

a-b. H & E staining of testis cross sections of PND35 control (**a**) or NELF-B cKO (**b**) mice. NELF-B cKO testis show severe degeneration of the epithelium with vacuoles (asterisk) and only Sertoli cells (S) remaining in the epithelium compared to control mice which successfully generate round spermatids (RS). Scale bar, 50 μ m.

c-d. H & E staining of testis cross sections of adult 9-month old control (**c**) or NELF-B cKO (**d**) mice. There was no recovery of spermatogenesis in the adult as shown by the vacuoles (asterisk) and lack of germ cells progressing to RS. Scale bar, 50 μ m.



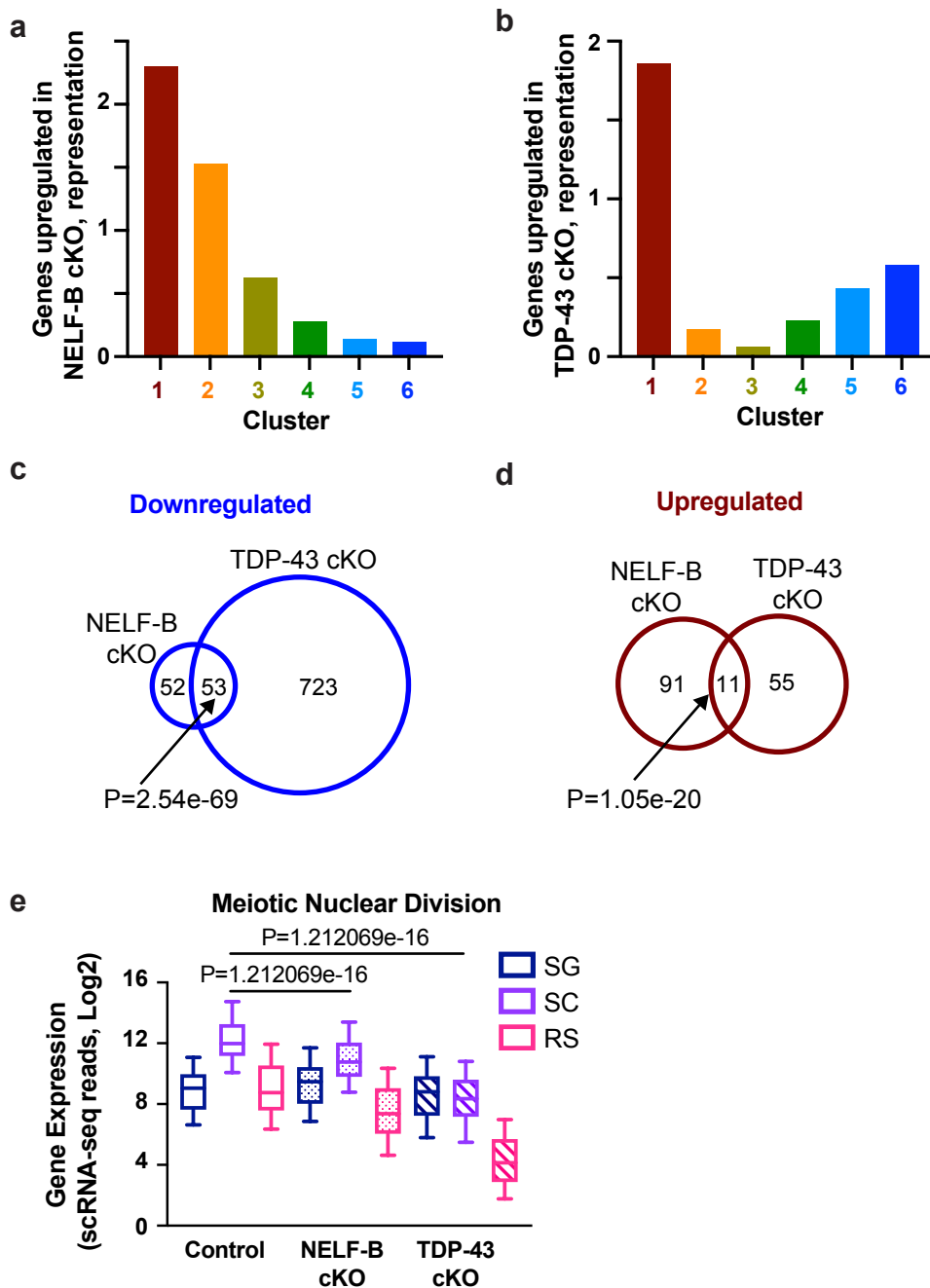
Supplementary Fig. 5 Marker genes and further characterization of scRNA-seq clusters.

a. Marker gene expression from scRNA-seq in control mice ($n=4$). Shown at left are the marker genes used to assign clusters to specific cell types. For this heatmap, signal for each gene was normalized within the row, to compare expression across cell types.

b. Gene expression of NELF-B (left) or TDP-43 (right) in control mice ($n=4$) for the indicated cell type. Shown are the mean and standard deviation of cumulative counts from scRNA-seq for each cell type.

c. UMAP representation of all 14 somatic and germ cell type clusters including the stages of SC, integrated across the 3 genotypes ($n=24,269$ cells). Normalized cell counts in control samples (SG=501, pre-Leptotene=90, Leptotene=71, Zygotene=312, Pachytene=323, Diplotene=188, RS=663).

d. Numbers of each germ cell type (including stages of SC) per genotype. Cell counts were normalized by the number of mice per genotype (Control $n=4$, NELF-B cKO $n=2$, TDP-43 cKO $n=3$).

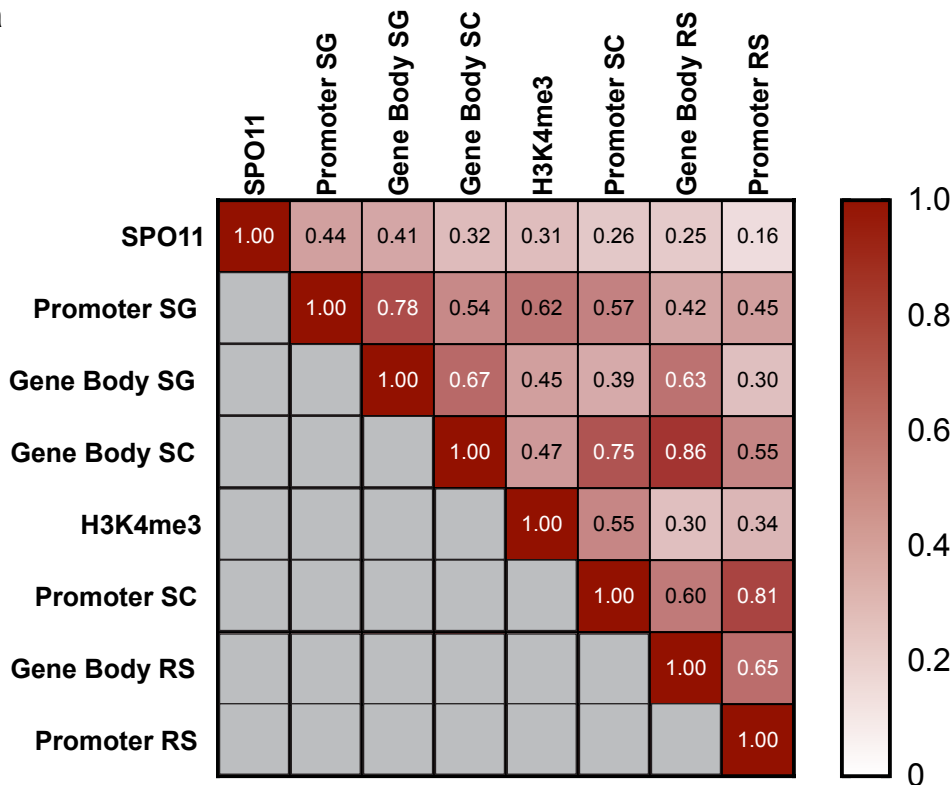
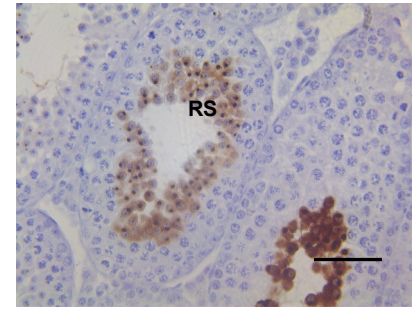
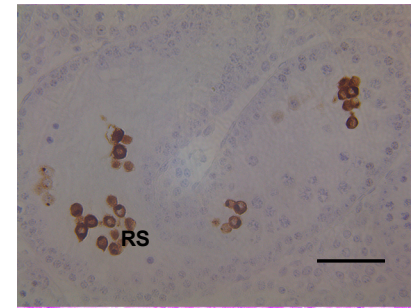


Supplementary Fig. 6 Investigation of significantly differentially expressed genes in NELF-B cKO and TDP-43 cKO mice.

a-b. Representation of upregulated genes from NELF-B cKO (**a**) or TDP-43 cKO mice (**b**), within each cluster (from Fig. 1f). A value of 1 indicates representation expected by chance.

c-d. Overlap between genes downregulated (**c**) or upregulated (**d**) in NELF-B and TDP-43 mice. P-value for the overlap of gene lists was calculated using the hypergeometric distribution with the phyper function in R

e. Box plots show distribution of expression in scRNA-seq for all 91 Cluster 4 genes in the Gene Ontology category Meiotic Nuclear Division. Expression is shown from control, NELF-B cKO, or TDP-43 cKO mice in SG, SC, or RS cells, as indicated. Line represents median, box represents 25-75th percentile, whiskers represent 10-90th percentile. P-values from two-sided Wilcoxon matched pairs test.

a**b****PND24 control****c****PND24 NELF-B cKO**

Supplementary Fig. 7. SPO11 oligo reads indicative of DSBs correlate best with promoter PRO-seq signal in SG, and NELF-B cKO mice generate very few post-meiotic germ cells.

a. Heatmap matrix of Spearman's correlation comparisons between SC SPO11 oligo sequencing (sum TSS+500), H3K4me3 ChIP-seq (Sum TSS+500), and PRO-seq promoter and gene body windows in SG, SC, and RS across all DE genes ($n=17,078$). Data sets are ordered left to right / top to bottom based on correlation with SPO11.

b-c. Immunohistochemistry of PND24 mouse testis using an antibody to the acrosomal marker SP-10. Control PND24 testis show the luminally arranged post-meiotic round spermatids (RS) with acrosome stained (**b**). PND24 NELF-B cKO testis show very few antibody-positive RS, indicating meiotic failure (**c**). Scale bar, 50 μm .

Supplementary Table 1: List of meiotic nuclear division genes from Cluster 4

Gene name	Description
<i>Hsf2bp</i>	Heat shock transcription factor 2 binding protein
<i>Msh4</i>	Muts homolog 4
<i>Spo11</i>	Spo11 initiator of meiotic double stranded breaks
<i>Cdc20</i>	Cell division cycle 20
<i>Msh5</i>	Muts homolog 5
<i>Rad51c</i>	Rad51 paralog c
<i>Spire2</i>	Spire type actin nucleation factor 2
<i>Tex14</i>	Testis expressed gene 14
<i>Asz1</i>	Ankyrin repeat, sam and basic leucine zipper domain containing 1
<i>Dmrtc2</i>	Doublesex and mab-3 related transcription factor like family c2
<i>Chfr</i>	Checkpoint with forkhead and ring finger domains
<i>Sirt2</i>	Sirtuin 2
<i>Mov10l1</i>	Mov10 like risc complex rna helicase 1
<i>Birc5</i>	Baculoviral iap repeat-containing 5
<i>Psmc3ip</i>	Proteasome (prosome, macropain) 26s subunit, atpase 3, interacting protein
<i>Ubb</i>	Ubiquitin b
<i>Sycp3</i>	Synaptonemal complex protein 3
<i>Fzr1</i>	Fizzy and cell division cycle 20 related 1
<i>Meikin</i>	Meiotic kinetochore factor
<i>Pttg1</i>	Pituitary tumor-transforming gene 1
<i>Hormad2</i>	Horma domain containing 2
<i>4930447C04Rik</i>	Riken cdna 4930447c04 gene
<i>Mlh3</i>	Mutl homolog 3
<i>Ddx4</i>	Dead box helicase 4
<i>Rad1</i>	Rad1 checkpoint dna exonuclease
<i>Erc4</i>	Excision repair cross-complementing rodent repair deficiency, complementation group 4
<i>Meiob</i>	Meiosis specific with ob domains
<i>Bag6</i>	Bcl2-associated athanogene 6
<i>Majin</i>	Membrane anchored junction protein
<i>Mus81</i>	Mus81 structure-specific endonuclease subunit
<i>Cenpx</i>	Centromere protein x
<i>Syce1</i>	Synaptonemal complex central element protein 1
<i>Mybl1</i>	Myeloblastosis oncogene-like 1
<i>Sgo2a</i>	Shugoshin 2a
<i>Nuf2</i>	Nuf2, ndc80 kinetochore complex component
<i>Terb2</i>	Telomere repeat binding bouquet formation protein 2
<i>Aurka</i>	Aurora kinase a
<i>Sycp1</i>	Synaptonemal complex protein 1
<i>Hormad1</i>	Horma domain containing 1
<i>Ccne2</i>	Cyclin e2
<i>Eif4g3</i>	Eukaryotic translation initiation factor 4 gamma, 3
<i>Cenpc1</i>	Centromere protein c1
<i>Brdt</i>	Bromodomain, testis-specific
<i>Rad51ap1</i>	Rad51 associated protein 1
<i>Tdrd12</i>	Tudor domain containing 12

<i>Mre11a</i>	Mre11a homolog a, double strand break repair nuclease
<i>Tex12</i>	Testis expressed 12
<i>Ccnb2</i>	Cyclin b2
<i>Cep63</i>	Centrosomal protein 63
<i>Syce1l</i>	Synaptonemal complex central element protein 1 like
<i>Piwil2</i>	Piwi-like rna-mediated gene silencing 2
<i>Fancd2</i>	Fanconi anemia, complementation group d2
<i>Smc4</i>	Structural maintenance of chromosomes 4
<i>Rmi1</i>	Recq mediated genome instability 1
<i>Figl1</i>	Fidgetin-like 1
<i>Psm8</i>	Proteasome subunit alpha 8
<i>Siah1a</i>	Siah e3 ubiquitin protein ligase 1a
<i>Stag3</i>	Stromal antigen 3
<i>Zscan21</i>	Zinc finger and scan domain containing 21
<i>Ttk</i>	Ttk protein kinase
<i>Tex19.2</i>	Testis expressed gene 19.2
<i>Slx4</i>	Slx4 structure-specific endonuclease subunit homolog (s. Cerevisiae)
<i>Mael</i>	Maelstrom spermatogenic transposon silencer
<i>Brca2</i>	Breast cancer 2, early onset
<i>Tdrkh</i>	Tudor and kh domain containing protein
<i>Ago4</i>	Argonaute risc catalytic subunit 4
<i>Shc1</i>	Shc sh2-domain binding protein 1-like
<i>Hfm1</i>	Hfm1, atp-dependent dna helicase homolog
<i>Mcm5</i>	Minichromosome maintenance domain containing 2
<i>Npm2</i>	Nucleophosmin/nucleoplasm 2
<i>Fbxo43</i>	F-box protein 43
<i>Morc2b</i>	Microrchidia 2b
<i>Catsperz</i>	Cation channel sperm associated auxiliary subunit zeta
<i>Plcb1</i>	Phospholipase c, beta 1
<i>Meioc</i>	Meiosis specific with coiled-coil domain
<i>Spdy1</i>	Speedy/zingo cell cycle regulator family, member a
<i>Terb1</i>	Telomere repeat binding bouquet formation protein 1
<i>Tdrd9</i>	Tudor domain containing 9
<i>Rnf212</i>	Ring finger protein 212
<i>Esp1</i>	Extra spindle pole bodies 1, separase
<i>Hspa2</i>	Heat shock protein 2
<i>Sycp2</i>	Synaptonemal complex protein 2
<i>Cks2</i>	Cdc28 protein kinase regulatory subunit 2
<i>Mapk15</i>	Mitogen-activated protein kinase 15
<i>Ccnb1ip1</i>	Cyclin b1 interacting protein 1
<i>Eme2</i>	Essential meiotic structure-specific endonuclease subunit 2
<i>Cntd1</i>	Cyclin n-terminal domain containing 1
<i>Rad54b</i>	Rad54 homolog b (s. Cerevisiae)
<i>1700028K03Rik</i>	Riken cDNA 1700028k03 gene
<i>Ankrd31</i>	Ankyrin repeat domain 31
<i>Spata22</i>	Spermatogenesis associated 22