Supplemental Materials
Molecular Biology of the Cell

Fine et al.

Supplemental Materials

Figure S1. Systematic variance is apparent across batches and litters. Colors correspond to genotype and shapes correspond to sample age, as indicated. Dataset information refers to the incorporation of a previously published dataset (baseline) into our dataset (Prdm9). (A) Principal component 1 (PC1) versus principal component 2 (PC2) of PCA run on all samples. (B,C,D) Principal component 1 (PC1) versus principal component 2 (PC2) of PCA run on 8 dpp, 12 dpp, and 16 dpp samples respectively. Icon labels correspond to designations of litters.

Figure S2. Differential expression between $Prdm9^{-/-}$ and $Prdm9^{+/+}$ samples is detected irrespective of batch effects. Colors correspond to genotype and shapes correspond to sample age, as indicated. (A) Principal component 1 (PC1) versus principal component 3 (PC3) of PCA run on all ComBat-adjusted samples. (B) Quantification of average frequencies of spermatogenic and meiotic prophase substages represented at each time point in the samples of germ cells retrieved from $Prdm9^{+/-}$ testes. (C,D,E) Principal component 1 (PC1) versus principal component 2 (PC2) of PCA run on ComBat-adjusted data from samples at 8 dpp, 12 dpp, and 16 dpp respectively. Icon labels correspond to designations of litters.

Figure S3. *Prdm9* transcript abundance reflects the genetic mutation. (A) Log_2(TPM+1) expression of *Prdm9* at 8, 12, and 16 dpp. *** represents FDR< 0.0001. (B) Sashimi plot of transcript alignment to exons of *Prdm9* in $Prdm9^{+/+}$ (blue) and $Prdm9^{-/-}$ (red). The x-axis represents the length of the *Prdm9* gene and the y-axis represents the number of reads aligned to a given region. Horizontal, curved lines represent the mRNA reads bridging an intron, with the number of reads spanning the intron denoted in the middle of the line.

Figure S4. Sex-chromosome gene expression reflects impaired meiotic sex-chromosome inactivation (MSCI) in *Prdm9*^{/-} samples. (A) Expression of each XY gene at 8, 12, and 16 dpp, relative to mean expression at 8 dpp in *Prdm9*$^{+/+}$ samples. (B) Expression of each autosomal gene at 8, 12, and 16 dpp, relative to mean expression at 8 dpp in *Prdm9*$^{+/+}$ samples. Colors denote genotypes in A and B, as indicated.
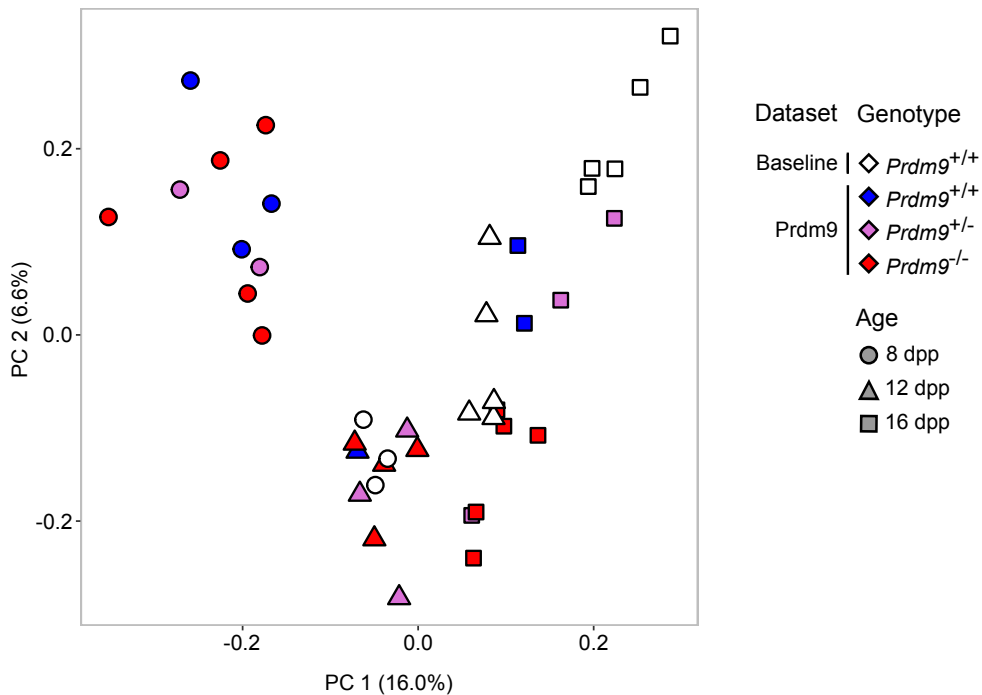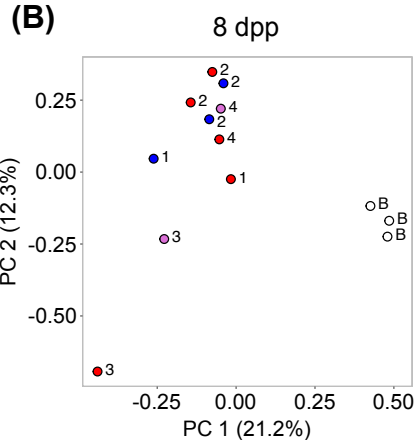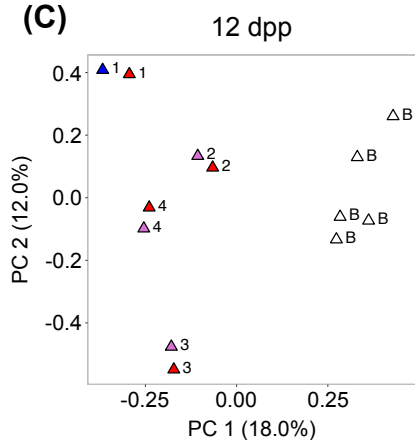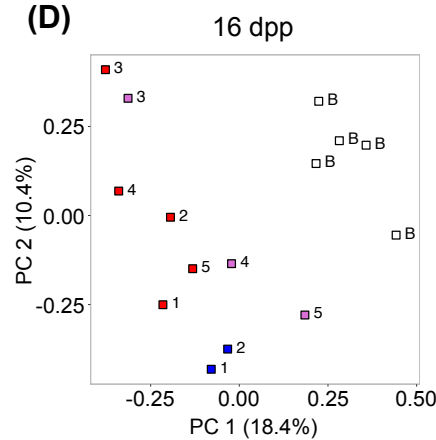
Figure S5. PMCA identifies substage-specific transcripts in wild-type and *Prdm9*$^{/-}$ samples. (A) Relative expression of transcripts assigned to each substage based on their wild-type expression patterns. (B) Relative expression of transcripts assigned to each substage based on their *Prdm9*$^{/-}$ expression patterns.
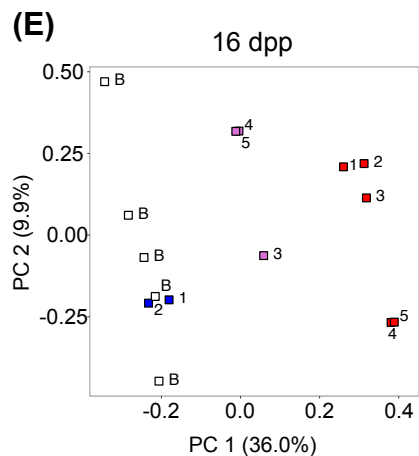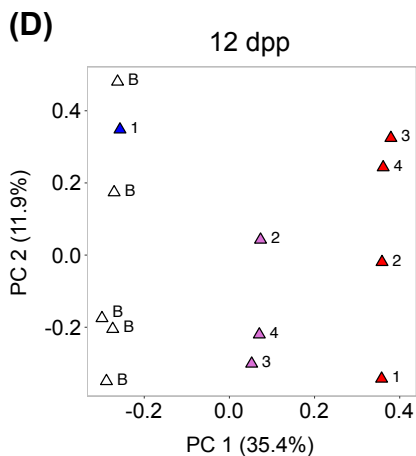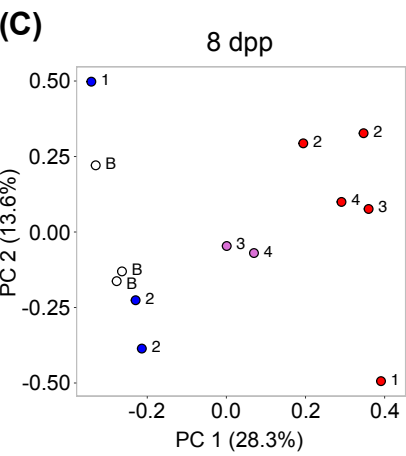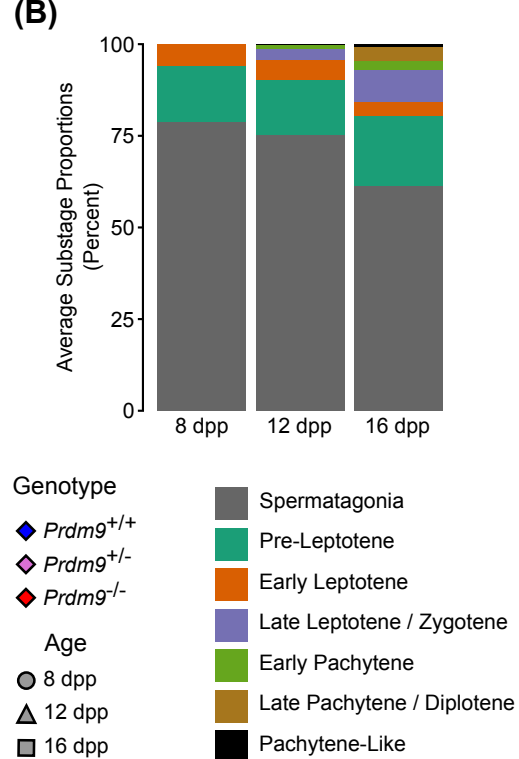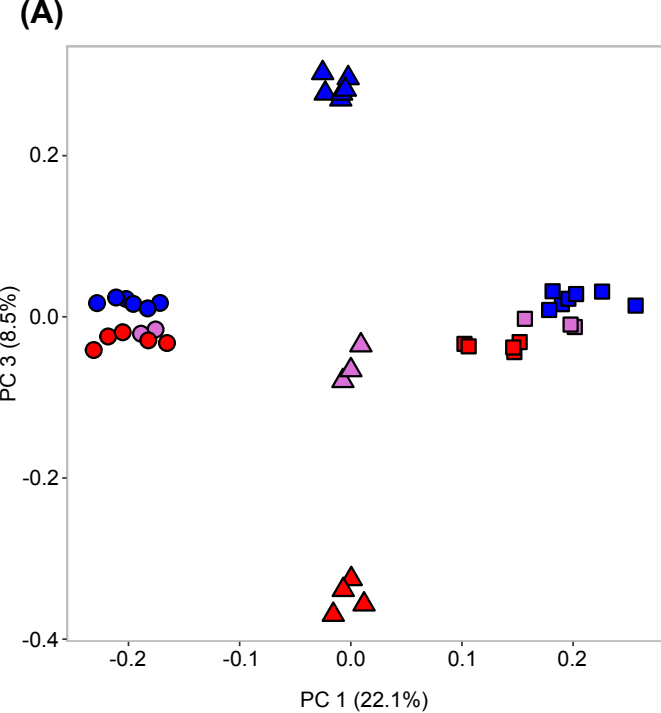
Table S1. Sample information. Age, genotype, dataset, and litter information for each sample,including cytological counts for each sample.

Table S2. Gene expression and differential analysis results. Transcript info, including transcript abundance, differential expression statistics, chromosomal location, and substage specificity. Unknown or undefined results are indicated with a "Z".

Table S3. Gene Ontology (GO) term enrichment analysis results. GO terms enriched among differentially expressed genes at 16 dpp.

Table S4. Differential expression gene counts among substage specific genes. Numbers and percentages of differentially expressed genes that had been assigned to specific substages of meiotic progression in *Prdm9*$^{+/+}$ and *Prdm9*$^{/-}$ samples.

(A)

Dataset | Genotype
Baseline | $Prdm9^{+/+}$ (open diamond)
Prdm9 | $Prdm9^{+/+}$ (blue diamond)
| $Prdm9^{+/-}$ (purple diamond)
| $Prdm9^{-/-}$ (red diamond)

Age
8 dpp (circle)
12 dpp (triangle)
16 dpp (square)

(B) 8 dpp

(C) 12 dpp

(D) 16 dpp

(A)

PC 3 (8.5%) / PC 1 (22.1%)

(B)

Average Substage Proportions (Percent)

8 dpp   12 dpp   16 dpp

Genotype
- $Prdm9^{+/+}$
- $Prdm9^{+/-}$
- $Prdm9^{-/-}$

Age
- 8 dpp
- 12 dpp
- 16 dpp

- Spermatagonia
- Pre-Leptotene
- Early Leptotene
- Late Leptotene / Zygotene
- Early Pachytene
- Late Pachytene / Diplotene
- Pachytene-Like

(C) 8 dpp — PC 1 (28.3%) / PC 2 (13.6%)

(D) 12 dpp — PC 1 (35.4%) / PC 2 (11.9%)

(E) 16 dpp — PC 1 (36.0%) / PC 2 (9.9%)

(A) *Prdm9*

Expression (Log$_2$(TPM+1)) vs Days postpartum (dpp)

Genotype
- ◆ *Prdm9*$^{+/+}$
- ◆ *Prdm9*$^{+/-}$
- ◆ *Prdm9*$^{-/-}$
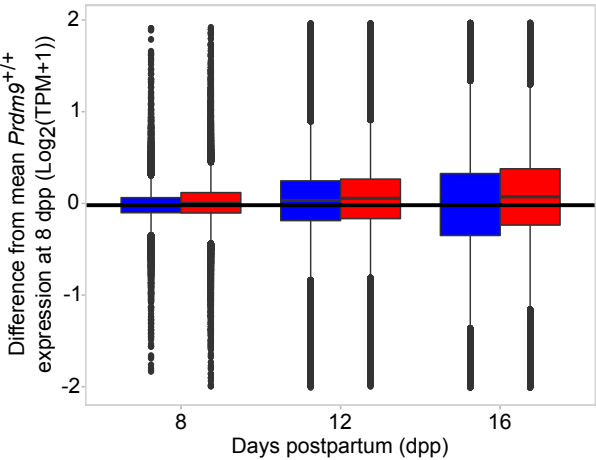
(B)

Reads mapped

Prdm9

(A)

XY-linked gene expression
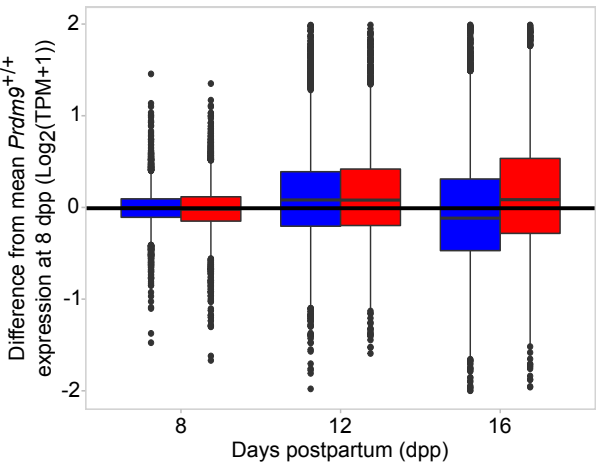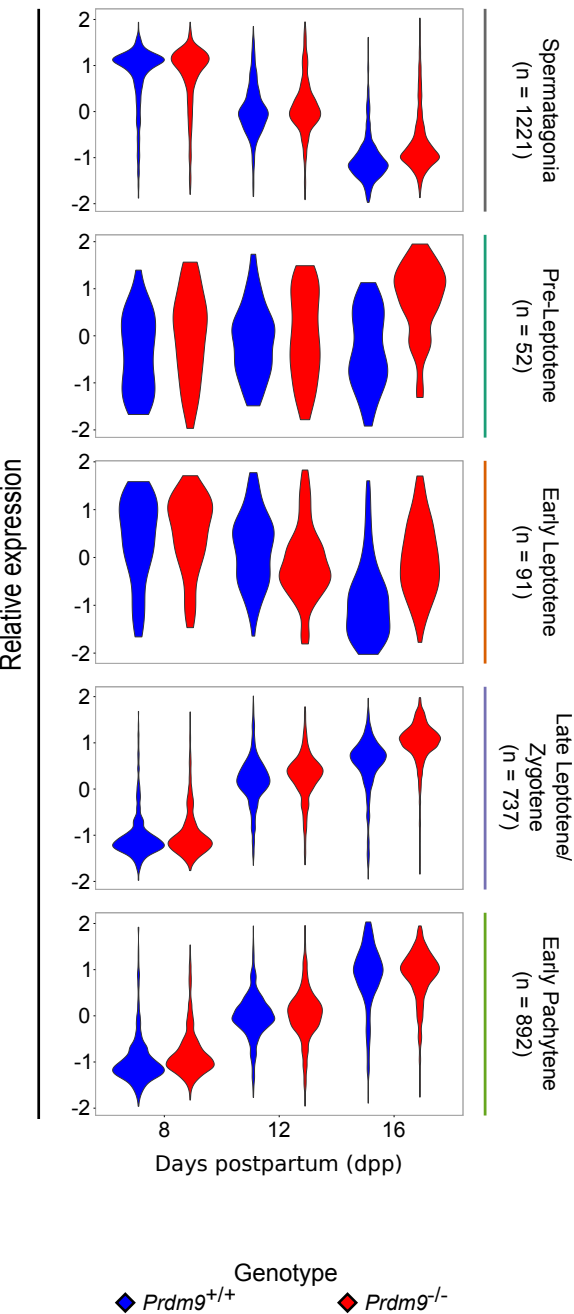
(B)

Autosome gene expression

Genotype

◆ *Prdm9*$^{+/+}$ ◆ *Prdm9*$^{-/-}$

(A) Substage specificity in wild-type

(B) Substage specificity in *Prdm9*-KO

Spermatogonia (n = 1221)

Pre-Leptotene (n = 52)

Early Leptotene (n = 91)

Late Leptotene/Zygotene (n = 737)

Early Pachytene (n = 892)

Spermatogonia (n = 412)

Pre-Leptotene (n = 134)

Early Leptotene (n = 14)

Late Leptotene/Zygotene (n = 460)

Early Pachytene (n = 479)

Pachytene-Like (n = 60)

Relative expression

Days postpartum (dpp)

Genotype

◆ *Prdm9*⁺/⁺   ◆ *Prdm9*⁻/⁻