Interrogating the Functions of PRDM9 Domains in Meiosis

Sarah Thibault-Sennett,*[,1] Qi Yu,[†,1] Fatima Smagulova,*[,2] Jeff Cloutier,[†] Kevin Brick,[†] R. Daniel Camerini-Otero,[†,3] and Galina V. Petukhova*[,3]

*Department of Biochemistry and Molecular Biology, Uniformed Services University of Health Sciences (USUHS), Bethesda, Maryland 20814 and [†]Genetics and Biochemistry Branch, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health (NIH), Bethesda, Maryland 20892

ORCID ID: 0000-0002-9596-6400 (K.B.)

ABSTRACT Homologous recombination is required for proper segregation of homologous chromosomes during meiosis. It occurs predominantly at recombination hotspots that are defined by the DNA binding specificity of the PRDM9 protein. PRDM9 contains three conserved domains typically involved in regulation of transcription; yet, the role of PRDM9 in gene expression control is not clear. Here, we analyze the germline transcriptome of *Prdm9*$^{-/-}$ male mice in comparison to *Prdm9*$^{+/+}$ males and find no apparent differences in the mRNA and miRNA profiles. We further explore the role of PRDM9 in meiosis by analyzing the effect of the KRAB, SSXRD, and post-SET zinc finger deletions in a cell culture expression system and the KRAB domain deletion in mice. We found that although the post-SET zinc finger and the KRAB domains are not essential for the methyltransferase activity of PRDM9 in cell culture, the KRAB domain mutant mice show only residual PRDM9 methyltransferase activity and undergo meiotic arrest. In aggregate, our data indicate that domains typically involved in regulation of gene expression do not serve that role in PRDM9, but are likely involved in setting the proper chromatin environment for initiation and completion of homologous recombination.

KEYWORDS PRDM9; meiosis; homologous recombination; KRAB domain

HOMOLOGOUS recombination is required for the faithful segregation of homologous chromosomes during meiosis. It assures that homologous chromosomes find each other and stay connected until the first meiotic division, when they segregate to different daughter cells. Failure to segregate properly results in aneuploidy in gametes and offspring, which accounts for 35% of miscarriages and developmental disabilities in 0.3% of live births (Hassold *et al.* 2007). The most severe recombination defects lead to meiotic arrest and consequent infertility (reviewed in Handel and Schimenti 2010).

doi: https://doi.org/10.1534/genetics.118.300565
Manuscript received November 27, 2017; accepted for publication April 13, 2018; published Early Online April 19, 2018.
Available freely online through the author-supported open access option.
Supplemental material available at Figshare: https://doi.org/10.25386/genetics.6157487.
[1]These authors contributed equally to this work.
[2]Present address: 9 Avenue du Professeur Léon Bernard, 35000 Rennes, France.
[3]Corresponding authors: Department of Biochemistry and Molecular Biology, USUHS, 4301 Jones Bridge Rd., Bethesda, MD 20814. E-mail: Galina.Petukhova@usuhs.edu; and National Institutes of Health, Bldg. 5, Room 205A, 5 Memorial Dr., Bethesda, MD 20814. E-mail: rc10d@nih.gov; rdcamerini@nih.gov

During meiotic recombination, the chromosomes undergo programmed DNA double-stranded breaks (DSBs) (reviewed in Keeney and Neale 2006). This triggers a genome-wide search for the intact homologous chromosome to be used for repair. RAD51 and DMC1 recombinases bind to single-stranded tails of the resected DSB ends and initiate recombinational repair of DSBs that results in either crossover or noncrossover products (Keeney and Neale 2006). Meiotic DSBs preferentially occur in discrete sites of the genome called recombination hotspots (reviewed in Arnheim *et al.* 2007; Buard and de Massy 2007; Lichten 2008; Clark *et al.* 2010; Paigen and Petkov 2010). In mice and humans, hotspot locations are defined by the DNA binding specificity of the PR/SET domain containing nine (PRDM9) protein that trimethylates Lysine 4 of the histone H3 (H3K4me3) at the sites where recombination may subsequently occur (Baudat *et al.* 2010; Myers *et al.* 2010; Parvanov *et al.* 2010; Grey *et al.* 2011; Brick *et al.* 2012). The DNA binding domain of PRDM9 consists of an array of zinc finger (ZF) domains that vary primarily at three amino acids that define the sequence specificity of DNA binding (reviewed in Wolfe *et al.* 2000;

Persikov *et al.* 2009). Remarkably, the ZF array is extremely polymorphic and may consist of different ZFs in different individuals, leading to a completely different distribution of recombination hotspots. When PRDM9 is absent, DSBs are formed at so-called default hotspots, the other available H3K4me3 sites within the genome, mostly at gene promoters and enhancer elements (Brick *et al.* 2012). These default sites are almost never used in wild-type mice, suggesting a dominant direct interaction between PRDM9 and the DSB formation machinery. Indeed, recent studies implicated the KRAB domain of PRDM9 in tethering the hotspot DNA to the chromosome cores, where the DSB machinery resides, through an interaction with CXXC1 (Imai *et al.* 2017; Parvanov *et al.* 2017). *Prdm9*$^{-/-}$ mice undergo meiotic arrest due to recombination failure, leading to sterility in both sexes (Hayashi *et al.* 2005). Interestingly, the loss of *Prdm9* does not universally result in sterility as *Prdm9* knockouts in some species remain fertile (Narasimhan *et al.* 2016), and the functional *Prdm9* gene has been lost in several lineages (Ponting 2011; Baker *et al.* 2017).

PRDM9 belongs to the PRDM protein family, a group that is defined by the presence of a ZF array and the PR/SET domain (reviewed in Hohenauer and Moore 2012). The PR/SET domain of PRDM9 is responsible for histone methyltransferase activity, including mono-, di- and trimethylation of the H3K4 (Hayashi *et al.* 2005; Wu *et al.* 2013), as well as H3K36 trimethylation (Eram *et al.* 2014; Powers *et al.* 2016). This domain is distantly related to the classic SET domain found in the histone lysine methyltransferases, a protein family involved in regulation of gene expression. The crystal structure of the PRDM9 PR/SET domain suggests that the methyltransferase activity of PRDM9 is autoregulated and the lone ZF located in the post-SET region is involved in this regulation (Wu *et al.* 2013).

PRDM family members either methylate their histone targets directly or recruit other histone modifiers to chromatin (Hohenauer and Moore 2012). Owing to these activities, the PRDM family is involved in a diverse array of developmental processes through the regulation of gene expression (Hohenauer and Moore 2012). PRDM9, however, is unique in that it also has the Kruppel Associated Box (KRAB) domain (Liu *et al.* 2014). This leads to inclusion of PRDM9 in a second family of transcription factors, the KRAB-ZFPs (KRAB Zinc Finger Proteins). Members of this family also contain a ZF DNA-binding array, but have a KRAB domain instead of PR/SET (Lupo *et al.* 2013).

The KRAB domain is typically involved in gene silencing through interaction with the TRIM28 protein, also known as KAP-1 (reviewed in Iyengar and Farnham 2011). TRIM28 interacts with Heterochromatin Protein 1 (HP-1) and serves as a scaffold for recruiting chromatin modifying enzymes and remodeling complexes to promoters of target genes bound by the KRAB-ZFPs (Iyengar and Farnham 2011). Interestingly, the KRAB domain of PRDM9 is missing the amino acids required for interaction with TRIM28 (Liu *et al.* 2014), and recent studies found no direct interaction between these two proteins (Patel *et al.* 2016; Imai *et al.* 2017). In fact, the KRAB domain of PRDM9 has higher similarity to the atypical KRAB domain found in the Synovial Sarcoma X chromosome breakpoint (SSX) protein family (Lim *et al.* 1998). PRDM9 and the SSX family members have another similarity: they both contain the SSX Repressive Domain (SSXRD) following the atypical KRAB domain. Although the SSXRD was shown to be a potent transcription repressor (Lim *et al.* 1998) the molecular mechanism behind this function has not been described.

Considering that PRDM9 has three domains implicated in regulation of transcription, it was originally thought of as a transcription factor (Hayashi *et al.* 2005; Matsui and Hayashi 2007; Mihola *et al.* 2009; reviewed in Nowick *et al.* 2013; Capilla *et al.* 2016). This role was attractive, because *Prdm9* is the only known speciation gene in mammals, and the conflict between regulatory mechanisms is perceived as the most likely cause of hybrid incompatibilities (Nowick *et al.* 2013; Mack and Nachman 2017). Although later studies suggested a different role of *Prdm9* in speciation (Davies *et al.* 2016; Smagulova *et al.* 2016) they did not rule out the possibility of additional mechanisms that may involve misregulation of transcription. Furthermore, it was recently reported that, during meiosis, PRDM9 binds to some active promoters (Grey *et al.* 2017), bringing back the question of the possible regulatory function of PRDM9 in gene expression. In this study we evaluate the potential role of PRDM9 in transcription and the role of transcription-related domains within PRDM9 in PRDM9 function.

We first used mRNA sequencing to demonstrate that inactivation of *Prdm9* does not lead to changes in gene expression. We then showed that PRDM9 does not appear to regulate miRNA expression as well. We next assessed the effect of the KRAB domain, SSXRD, and post-SET ZF deletions on methyltransferase activity of PRDM9 expressed in cell culture. Although deletion of SSXRD completely abolished histone methylation by PRDM9, the KRAB and post-SET ZF mutants retained methyltransferase activity. To assess whether the KRAB domain is essential for maintaining the meiotic program, we generated the KRAB domain deletion mouse strain and found that mutant mice undergo meiotic arrest resulting in sterility. We found only a trace amount of H3K4me3 at the sites of expected hotspots. However, DSB formation was evident at ~30% of PRDM9-dependent H3K4me3 sites in the KRAB mutant, raising the possibility that deletion of the KRAB domain does not completely abolish communication of PRDM9 with the DSB formation machinery.

Materials and Methods

Mice and cell lines

Cell-culture-based PRDM9 mutant analyses were performed using the GC-1 spg immortalized BALB/c cell line derived from spermatogonial cells (ATCC CRL-2053; Hofmann *et al.* 1992).

All animal procedures have been approved by the Uniformed Services University of Health Sciences (USUHS) Animal Care and Use Committee. *Hop2* knockout mice have been previously described (Petukhova *et al.* 2003). *Prdm9* knockout mice have been previously described (Hayashi *et al.* 2005) and were obtained from the Jackson Laboratory (stock number 010719). *Spo11* knockout mice (strain *deltaSpo.BC/B6*) have been described in Smagulova *et al.* (2013). C57Bl/6 was used as a wild-type strain and all mutant strains were on the C57Bl/6 background. Adult (2–6 month old) mice were used for the analyses unless specifically mentioned in the text. *Prdm9*^{K} mice were generated on a C57Bl/6 background as outlined in Supplemental Material, Figure S2 and Table S4. Targeting design and pronuclear injections were done by University of California Irvine transgenic mouse facility. Genomic DNA was sequenced to confirm the correct deletion.

RNA sequencing

Total RNA was isolated from whole testes using Qiagen RNeasy mini kit (74104). Polyadenylated RNA was purified using Oligotex beads (70022; Qiagen). Quality of the RNAs was verified on Bioanalyzer and RNA was fragmented to ~200 bp. Sequencing libraries were prepared using mRNA-Seq Sample Preparation Kit (RS-930-10; Illumina) and sequencing was performed using Illumina X2500 platform with 50 bp paired-end reads. Reads for each sample were quantified against mouse GRCm38 cDNA using kallisto0.43.0 (Bray *et al.* 2016), bootstrap = 10 was used to measure the accuracy of the quantification (which does not influence quantification). Sleuth (Pimentel *et al.* 2017) was used to perform the differential expression tests of genes and transcripts from kallisto quantifications. Genes were identified as expressed if transcripts per million reads (TPM) > 0.1 in two replicates (Everaert *et al.* 2017).

miRNA analysis

Small RNAs were isolated from whole testes using the mirVana miRNA Isolation Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. The resulting RNA (~50 μg) was heat-denatured, loaded and resolved on the 15% denaturing acrylamide gel. Two radiolabeled RNA oligos of 18 and 28 bases were used as the ladder. The gel was run for 3 hr at 45 W. The gel slice containing RNA from 19 to 27 bases was excised and RNA was eluted in 800 μl of 0.4 M NaCl overnight. RNA was precipitated with ethanol and resuspended in 6 μl of H_2O (~200 ng of RNA). Library was prepared with the NEBNext Small RNA Library Prep Set for Illumina (E7330S; New England BioLabs) following the manufacturer's instructions. The microRNA samples were sequenced using Illumina HiSequation 2500 using single-end reads. After trimming the adapter, reads were filtered using the length threshold of 17–29 bp. Reads were mapped to the mouse (mm10) genome using bowtie 0.12.9 (Langmead *et al.* 2009) with parameters of "bowtie -f -n 0 -e 80 -l 18 -a -m 5 –best –strata" in MiRDeep2 v2.0.0.7 (Friedländer *et al.* 2012), bowtie "best strata" has been shown to be one of the most sensitive and robust alignment methods for miRNA analysis (Tam *et al.* 2015). The quantification and novel miRNAs identification (used miRNAs from other species as templates) was performed by MiRDeep2 v2.0.0.7 against annotated mouse miRNAs of miRBase (Version 21) (Kozomara and Griffiths-Jones 2014). Significance testing for the differential expression analysis was performed using DESeq2 (Love *et al.* 2014). Common novel miRNAs between the four samples were defined as novel miRNAs (with TPM > 10) that have 100% sequence overlap between each other.

GFP-PRDM9 constructs

Full length B6 mouse PRDM9 following GFP was cloned into pcDNA4-Myc-His-B (ThermoFisher). A linker region of the following sequence separated GFP from PRDM9 "GGAGGCTCTGGCGGATCTGGAGGGTCTGGAGGT." The PRDM9 start codon was removed but the stop codon was maintained. Mutant constructs were generated using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs). The specific mutations for constructs in Figure 2 are as follows:

SET: Y345F; KRAB1: del(D30-P87), KRAB2: del(K27-V198V), SSXRD: del(L175-Y206); ZincF: del(H392-W413). GFP-STOP was generated by the following mutation in linker region above: G1T, G2A.

Cell culture and transfection

GC-1 cells were grown in DMEM supplemented with 10% FBS and 5% Penn-Strep at 37° in 5% CO_2. Cells were plated and allowed to grow for 15 hr before transfection using Lipofectamine 3000 (Thermo Fisher Scientific). Cells were allowed to grow in transfection solution for 15 hr before collection, and typically showed 50–60% transfection efficiency.

Chromatin immunoprecipitation followed by high-throughput sequencing

GC-1 cells were fixed directly on the plate as described in Baker *et al.* (2015). Chromatin immunoprecipitation (ChIP) for testis and GC-1 cell samples was carried out as previously described (Brick *et al.* 2012).

For single nucleosome resolution, H3K4me3 ChIP followed by high-throughput sequencing (ChIP-seq) instead of sonication the fixed nuclei were resuspended in MNase buffer (50 mmol Tris-HCl, 1 mmol $CaCl_2$, 4 mmol $MgCl_2$, 4% NP-40) at 26 μl/10^6 cells, supplemented with 1 mmol PMSF and 1× cOmplete, EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich). The Micrococcal nuclease digestion was then carried out as described in Billings *et al.* (2013). Following digestion, the sample was diluted 3:1 in ChIP buffer (0.1% SDS, 1% NP-40, 2 mM EDTA, 20 mM Tris-HCl pH 8, 150 mM NaCl) and H3K4me3 ChIP was performed as in Brick *et al.* (2012). The following antibodies were used: anti-DMC1, Santa Cruz sc8973; anti-H3K4me3, Millipore #07-473. H3K4me3 sequencing was performed with 75 or 150 bp SE reads on an Illumina NextSeq500 and SSDS sequencing was performed with 50 bp PE reads on an Illumina HiSequation 2500. H3K4me3 ChIP-sequencing reads were aligned to the mouse

genome mm10 using bwa aln (Li and Durbin 2009). Overlap of reads with either promoters or known hotspots was performed using bedtools 2.26.0 Intersect. Bigwig files were generated with deepTools 1.5.11 bam-Coverage with the following parameters: -binSize 10 -centerReads -normalizeUsingRPKM. Bigwig files were then used to generate heatmaps across promoters or hotspots using deepTools computeMatrix reference-point followed by plotHeatmap. SSDS reads were aligned to the genome using bwa aln (0.7.12) and ssDNA derived reads were identified using the single stranded DNA sequencing (SSDS) processing pipeline (Khil *et al.* 2012; https://github.com/kevbrick/SSDSpipeline.git).

Identification of DSB hotspots

Uniquely mapped fragments unambiguously derived from ssDNA (ssDNA type 1) and having both reads with a mapping quality score ≥30 were used for identifying hotspot locations (peak calling). Normalization of ChIP-seq (NCIS) was used to estimate the background fraction for each library. Peak calling was performed using MACS (v.2.1.0.20150420) (Zhang *et al.* 2008) with the following parameters: -ratio [out-put from NCIS] -g mm -bw 1000 -keep-dup all -slocal 5000. DSB hotspots within regions previously blacklisted (Smagulova *et al.* 2016) were removed and hotspot strength was calculated as described previously (Smagulova *et al.* 2016). The pipeline for calling DSB hotspots from SSDS can be obtained from (https://github.com/kevbrick/callHotspotsSSDS.git). The fraction of reads in peaks is calculated as the number of ssDNA fragments that overlap a feature (hotspots or TSSs) divided by the total number of ssDNA fragments.

Assessment of PRDM9 expression in GC-1 cells

Live cell images were taken 12 hr post transfection to confirm nuclear localization and estimate expression level of GFP-PRDM9 constructs. During the ChIP procedure, a 50 μl aliquot of sheared chromatin lysate was removed and frozen at −80°. The sample was incubated at 95° for 10 min in SDS loading buffer with DTT and resolved in a Bolt Bis-Tris Plus acrylamide gel followed by transfer onto polyvinylidene difluoride (PVDF) membrane. The following antibodies were used: rabbit anti-GFP (AB290; Abcam) and rabbit anti-CTCF (AB70303; Abcam). The blots were developed using SuperSignal West Pico PLUS Chemiluminescent Substrate.

Assessment of PRDM9 expression in mouse testis

Testes were removed and stored at −80° in RNALater (Thermo Fisher Scientific). Total RNA was extracted using RNeasy plus mini kit (Qiagen) and the cDNA was prepared with the Transcriptor First Strand cDNA Synthesis Kit (Roche). Quantitative PCR was performed using the SYBR Green PCR Master Mix (Thermo Fisher Scientific) using a 7500 AB Real-Time PCR system with ABI SDS Software. The primers used are listed in Table S5; *Gapdh* primers are from Spandidos *et al.* (2010).

Histology, chromosome spreads, and immunostaining

Testes or ovaries were dissected and stored in buffered formalin (1:10 dilution). Tissue sections with hematoxylin/eosin staining was done by American HistoLabs (Gaithersburg, MD). Meiotic chromosomes spreads were made and immunostained as described in Smagulova *et al.* (2013). The following primary antibodies were used: anti-SCP3 (Petukhova *et al.* 2003); anti-γH2AX (4411-PC-100; Trevigen). Samples were imaged using Leica fluorescent DM5500 microscope with Retiga camera and velocity acquisition system.

Data availability

Mouse strains are available upon request. Gene expression and ChIP-seq data are available at GEO with the accession number: GSE104850. Supplemental material available at Figshare: https://doi.org/10.25386/genetics.6157487.

Results

Inactivation of Prdm9 does not lead to changes in gene expression

Since the SET domain, the KRAB domain and the SSXRD are usually involved in regulation of gene expression, we assessed whether PRDM9 indeed is involved in transcription control. We carried out high throughput sequencing of mRNA (mRNA-seq) from testes of *Prdm9*$^{-/-}$ mice. As *Prdm9*$^{-/-}$ spermatocytes undergo meiotic arrest at pachytene-like stage (Hayashi *et al.* 2005), testes of the *Prdm9*$^{-/-}$ mice lack germ cells of later stages leading to a vastly different cell type composition compared to that in wild-type testes. Consequently, comparing the transcriptome between wild type and *Prdm9* knockout mice would not be meaningful. To get around this problem we used *Hop2*$^{-/-}$ mice as the *Prdm9*$^{+/+}$ control. In *Hop2*$^{-/-}$ mice, meiotic DSBs are formed, but not repaired, which leads to meiotic arrest at the stage similar to *Prdm9*$^{-/-}$, with a very similar cell type composition (Petukhova *et al.* 2003; Hayashi *et al.* 2005).

We examined differences in gene expression using two biological replicates of mRNA-seq for *Prdm9*$^{-/-}$ and *Hop2*$^{-/-}$ mice (Table S1). The correlation between *Prdm9*$^{-/-}$ and *Hop2*$^{-/-}$ expression is very strong across annotated genes that are expressed in at least one sample (TPM $>$ 0.1, Figure 1A, Spearman's R = 0.98–0.99). This indicates that the absence of *Prdm9* does not perturb the global gene expression pattern. As expected, the transcriptome of wild-type mice showed low similarity with the mutants (Figure 1A, Spearman's R = 0.69–0.70), reflecting the presence of the late meiotic and postmeiotic cells in wild type testes.

PRDM9 has been implicated in the regulation of the *Morc2b* gene expression because of altered expression level in *Prdm9*$^{-/-}$ testes (Hayashi *et al.* 2005) and in sterile hybrids between *Mus musculus musculus* and *Mus musculus domesticus* subspecies (Mihola *et al.* 2009). We found that in wild-type mice there is a 300-fold change in the *Morc2b* expression over the course of spermatogenesis, with a sharp increase

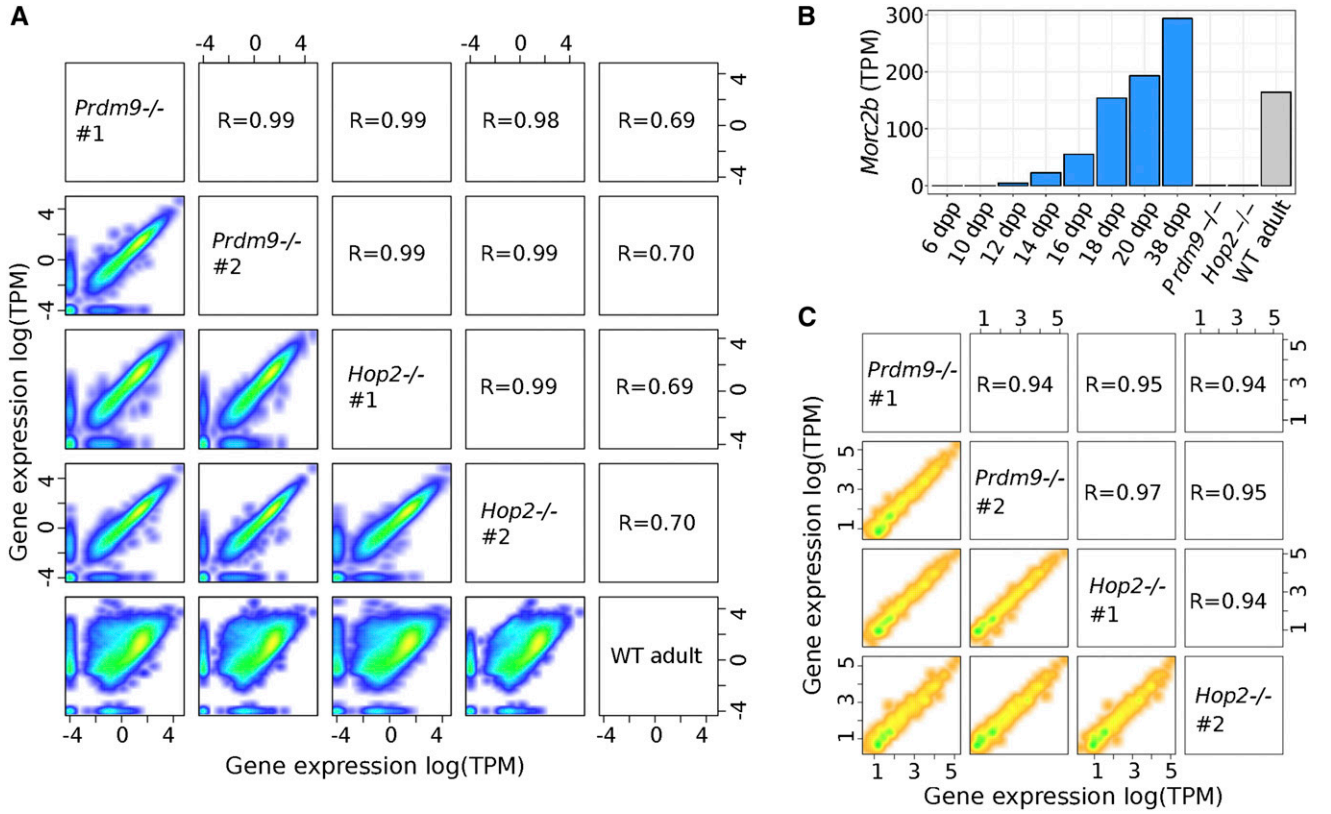

Figure 1 Inactivation of the *Prdm9* gene does not lead to changes in gene expression profiles. (A) Scatter plot of gene expression level measured by log(TPM). The numbers represent Spearman's correlation coefficient. Genes expressed at least in one sample (TPM $>$ 0.1) are shown. (B) Expression of the *Morc2b* gene in *Prdm9*$^{-/-}$ (average from two libraries), *Hop2*$^{-/-}$ (average from two libraries) and wild-type juvenile mice undergoing the first wave of meiosis (Margolin *et al.* 2014). Pachytene stage is reached by 14 days post partum (dpp) and most spermatocytes reach diplonema by 18 dpp. (C) Scatter plot of miRNA expression level measured by log(TPM). The numbers represent Spearman's correlation coefficient. Annotated miRNAs expressed at least in one sample (TPM $>$ 10) are shown.

starting at the onset of pachytene stage (Figure 1B). Therefore, even very slight changes in the testes cell population due to meiotic arrest or delay would lead to detectable changes in the abundance of the *Morc2b* transcript, explaining prior observations. Indeed, we found no significant difference (q-value = 1) in the *Morc2b* expression level between *Prdm9*$^{-/-}$ (mean TPM = 1.44, *Prdm9*$^{-/-}$ #1 TPM = 1.04, *Prdm9*$^{-/-}$ #2 TPM = 1.84) and *Prdm9*$^{+/+}$*Hop2*$^{-/-}$ mice (mean TPM = 0.38, *Hop2*$^{-/-}$ #1 TPM = 0.29, *Hop2*$^{-/-}$ #2 TPM = 0.48), indicating that *Morc2b* transcription is unlikely to be regulated by PRDM9. Overall, we did not detect any differentially expressed genes or transcripts between the *Prdm9*$^{-/-}$ and *Prdm9*$^{+/+}$*Hop2*$^{-/-}$ samples (q-value $<$ 0.05).

We next explored whether PRDM9 could play a role in regulating miRNAs, as they have been shown to play crucial roles in meiosis (Yu *et al.* 2005; Song *et al.* 2009; Royo *et al.* 2015). To compare miRNA levels we considered miRNAs that were detected at ≥10 TPM in at least one miRNA sequencing library prepared from two *Hop2*$^{-/-}$ and two *Prdm9*$^{-/-}$ samples (Table S2). We found that there is a strong overall correlation of miRNA profiles between *Prdm9*$^{-/-}$ and *Hop2*$^{-/-}$ samples (R = 0.94–0.97, Figure 1C), and that there are no differentially expressed miRNAs between the *Prdm9*$^{-/-}$ and *Hop2*$^{-/-}$ samples (adjusted P-value $<$ 0.05). In addition to annotated miRNAs we identified mature miRNAs that were not described in the miRBase (release 21), 12 of which were present in all four samples (Table S3). The majority of these novel miRNAs had very low expression level (1000 times lower than annotated miRNAs) and none of them showed differential expression between the *Prdm9*$^{-/-}$ and *Hop2*$^{-/-}$ samples (adjusted P-value $<$ 0.05). Although it remains possible that transcripts below our detection level are being influenced by PRDM9, our data indicate that inactivation of the *Prdm9* gene does not lead to significant changes in transcription, either of protein coding genes or of miRNAs.

Neither the KRAB domain nor the post-SET zinc finger are essential for H3K4me3 activity of PRDM9 in cell culture

Since we found no evidence of PRDM9 being involved in regulation of transcription, we decided to assess whether the KRAB domain and the SSXRD were required for other PRDM9 activities. We first expressed the wild type GFP-tagged PRDM9 (*Dom2* allele (Parvanov *et al.* 2010), WT in Figure 2) in the GC-1 mouse spermatogonial cell line (Hofmann *et al.* 1992) and evaluated the methyltransferase activity of

PRDM9 by H3K4me3 ChIP-seq. As expected, we found a marked enrichment of H3K4me3 reads at the known *Dom2* hotspots, but not at hotspots defined by a different *Prdm9* allele, *13R* (Brick *et al.* 2012) (Figure 2B). This enrichment was not due to expression of the endogenous *Prdm9*, because no H3K4me3 signal was detected at *Dom2* hotspots in GC-1 cells before transfection (Figure 2B, GC-1 cells carry the *Dom2 Prdm9* allele).

Following confirmation that the expressed GFP-PRDM9 was active, we used targeted mutagenesis to generate two alternative GFP-PRDM9 constructs that lacked the KRAB domain (KRAB-1 and KRAB-2, Figure 2A and Figure S1) and assessed their activity by H3K4me3 ChIP-seq. Additionally, we generated a point mutant with a single Y341F amino acid change in the active site of the PR/SET domain (SET, Figure 2A and Figure S1). This mutation was previously shown to abolish methyltransferase activity of PRDM9 (Wu *et al.* 2013) and serves in this study as a negative control. H3K4me3 is a relatively common epigenetic modification that is usually found at gene promoters and enhancer elements. H3K4me3 in these regions is not dependent on PRDM9 and therefore, we used the H3K4me3 ChIP-seq signal at promoters as an internal control for the ChIP efficiency. This allowed us to compare H3K4 trimethylation activity at hotspots across multiple wild type and mutant samples (Figure 2B).

We found that both KRAB mutants retained methyltransferase activity (Figure 2B). Relative to the H3K4me3 signal at transcription start sites, the magnitude of this signal was similar to that in cells expressing wild type PRDM9 (HS:TSS ratio, Figure 2B). Since the KRAB-2 mutant also lacks 24 out of 32 amino acids of the SSXRD we assessed whether the SSXRD is required for PRDM9 function. Unexpectedly, removing the entire SSXRD (SSXRD, Figure 2A and Figure S1) resulted in complete inactivation of PRDM9 with no trace of H3K4me3 enrichment at the *Dom2* hotspots (Figure 2B). We confirmed that the SSXRD mutant protein product was expressed at comparable to wild type level and properly localized to the nucleus (Figure S1); therefore, it is unlikely that the lack of the H3K4me3 signal was due to low level of the mutant PRDM9.

We next tested whether the post-SET zinc finger had an effect on the H3K4 methyltransferase activity of PRDM9. This lone zinc finger has been proposed to regulate the PR/SET catalytic activity by blocking access to the active site in the absence of histones (Wu *et al.* 2013). We found that deletion of this zinc finger (ZincF, Figure 2A and Figure S1) did not reduce PRDM9 activity and H3K4 trimethylation marks were still detected at expected hotspots (Figure 2B).

The KRAB domain is necessary for PRDM9 activity in vivo

After we found that deletion of the KRAB domain does not abolish methyltransferase activity of PRDM9 we wanted to investigate the effect of the KRAB domain deletion in meiosis. It was recently proposed that the KRAB domain of PRDM9 mediates tethering of the hotspot DNA to chromosome cores, where the proteins required for the DSB formation are located (Imai *et al.* 2017; Parvanov *et al.* 2017). To test this, we used CRISPR-based gene targeting to generate a knock-in mouse strain that lacks the KRAB domain of PRDM9, $Prdm9^{K/K}$ (Figure S2 and Table S4). The mutant $Prdm9^{K}$ allele matches exactly the KRAB-1 deletion mutant that was tested in cell culture and lacks the majority of the KRAB domain while retaining three N- and three C-terminal residues of the domain.

We found that both male and female $Prdm9^{K/K}$ mice were infertile. $Prdm9^{K/K}$ ovaries were greatly reduced in size and had no follicles (Figure S3). Testes size was also reduced and seminiferous tubules lacked spermatids, although primary spermatocytes were easily detectable (Figure S3). To evaluate the time of arrest more precisely, we analyzed meiotic progression by immunostaining chromosome spreads of the $Prdm9^{K/K}$ spermatocytes. We found that spermatocytes arrested at the stage resembling pachynema, but DSBs were only partially repaired and synapsis of the homologous chromosomes was incomplete (Figure 3). Overall, the phenotype of the $Prdm9^{K/K}$ mice was indistinguishable from that of the $Prdm9^{-/-}$ mutant (Figure 3). While we could not assess the protein level of PRDM9 in testes, we confirmed that the expression of the *Prdm9* KRAB deletion mutant transcript was not different from that of wild type *Prdm9* (Figure S2). As this same allele yields a protein product in GC-1 cells at the level similar to wild type *Prdm9* (Figure S1) and retains the methyltransferase activity (Figure 2B), we do not expect that PRDM9 lacking the KRAB domain is unstable *in vivo*. This makes it unlikely that the null phenotype observed in the $Prdm9^{K/K}$ mouse is due to lack of the $Prdm9^{K/K}$ protein product. Together these data indicate that the KRAB domain of PRDM9 is indispensable for meiosis.

To understand the reason for meiotic failure in the $Prdm9^{K/K}$ mice we evaluated whether the methyltransferase activity retained by the PRDM9 KRAB mutants in GC-1 cells (Figure 2B) was also detectable in the $Prdm9^{K/K}$ mice. We found only trace levels of H3K4 trimethylation at *Dom2* hotspots (Figure 4A). This suggests that the methyltransferase activity observed in cell culture was likely the result of *Prdm9* over-expression (Figure S1), although we cannot exclude the possibility that some complex meiosis-specific regulatory mechanisms gave rise to this difference.

Intriguingly, the purified PRDM9 PR/SET domain alone is capable of carrying out methyltransferase activity *in vitro* using synthetic peptides as targets (Wu *et al.* 2013). As PRDM9 lacking the KRAB domain is largely incapable of trimethylating H3K4 in testes it is possible that the KRAB domain or KRAB-mediated interactions with other proteins (Imai *et al.* 2017; Parvanov *et al.* 2017) are required for the PR/SET domain activity within the chromatin, where the access of PRDM9 to DNA or histone targets may be restricted. Dynamic interactions between PRDM9 and chromatin was previously proposed, when H3K4me3 profiling at recombination hotspots revealed nucleosome depleted

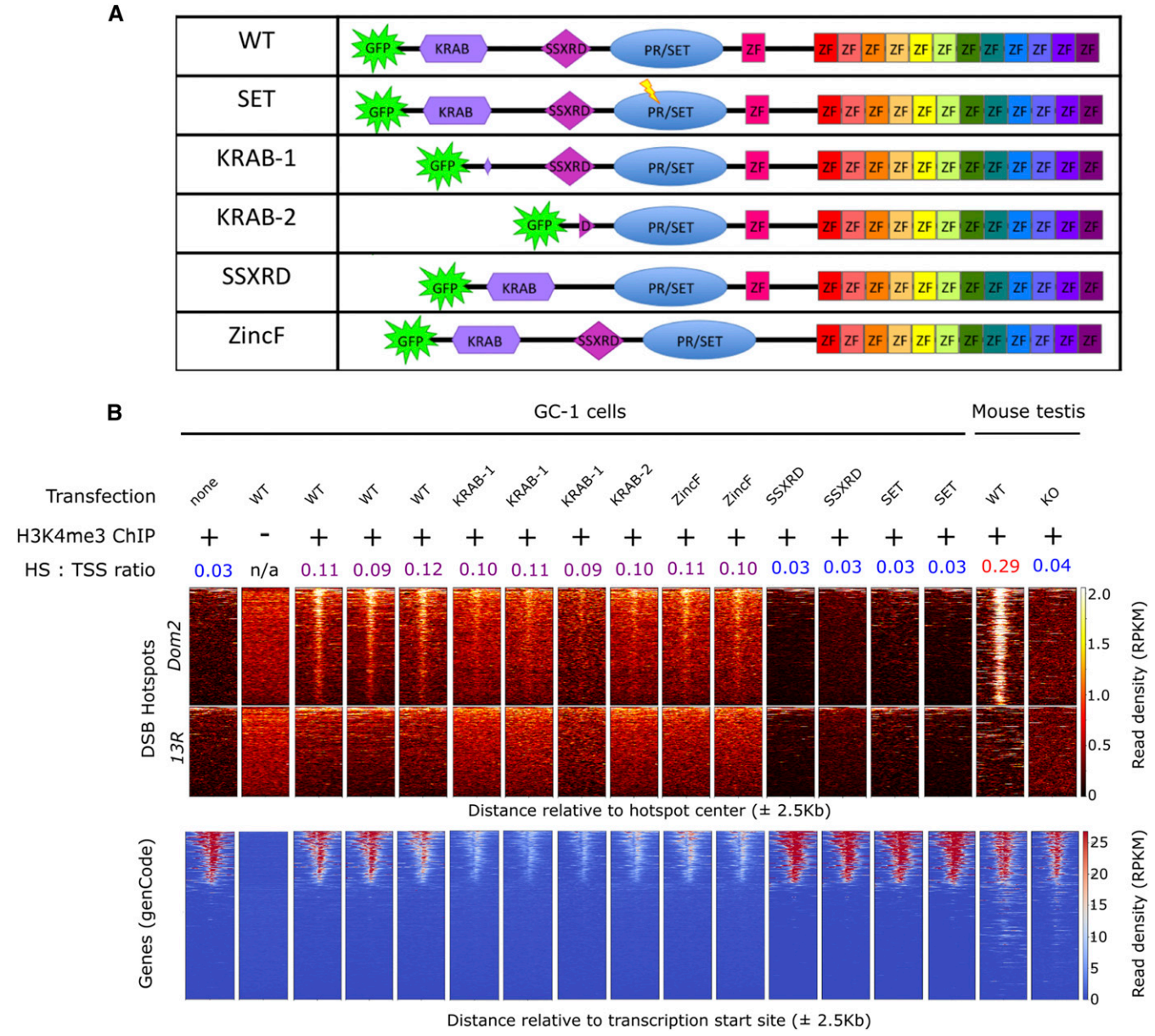

Figure 2 PRDM9 lacking the KRAB domain retains H3K4 methyltransferase activity in the GC-1 cell expression system. (A) Schematic of the constructs tested. (B) Indicated GFP-PRDM9 constructs were expressed in GC-1 cells followed by H3K4me3 ChIP-seq to assess methyltransferase activity of PRDM9 in reads per kilobase per million reads (RPKM). Upper rows of heatmaps (yellow-black color scheme) show enrichment of sequencing tags across 11,642 *Dom2* hotspots and 9834 *13R* hotspots. DSB hotspots overlapping transcription start sites (TSSs) are excluded. Lower row of heatmaps (red-blue color scheme) show enrichment of sequencing tags across transcription start sites (TSS) from GENCODE (v3.4). Input controls are indicated by "−" in "H3K4me3 ChIP" panel. "HS:TSS ratio" is the ratio of the H3K4me3 signal (FRiP; see *Materials and Methods*) at *Dom2*-defined DSB hotspots to the signal at GENCODE TSSs. This allows comparing the H3K4me3 signal at hotspots across samples with different background level.

regions at PRDM9 binding sites (Baker *et al.* 2014). To determine whether this was mediated by PRDM9 rather than downstream DSB formation and processing we carried out H3K4me3 profiling in the *Spo11*$^{-/-}$ mice that are deficient in DSB formation (Baudat *et al.* 2000; Romanienko and Camerini-Otero 2000). We found a clear nucleosome depleted region at hotspot locations, confirming that nucleosome reorganization at these sites occurs before the formation of DSBs, as a result of PRDM9 activity (Figure S4).

We next mapped DSB hotspots in the *Prdm9*$^{K/K}$ mice by ChIP with DMC1 antibodies followed by single-stranded DNA sequencing (SSDS; Khil *et al.* 2012). Not surprisingly, this analysis revealed that the vast majority of DSBs in the *Prdm9*$^{K/K}$ spermatocytes formed at promoters and enhancers, the default hotspots where DSBs form in the absence of PRDM9-dependent H3K4me3 marks (Brick *et al.* 2012) (Figure 4, B and C). Nevertheless, a small number of DSB hotspots formed at locations defined by PRDM9. With few exceptions, these were much weaker, but approximately half

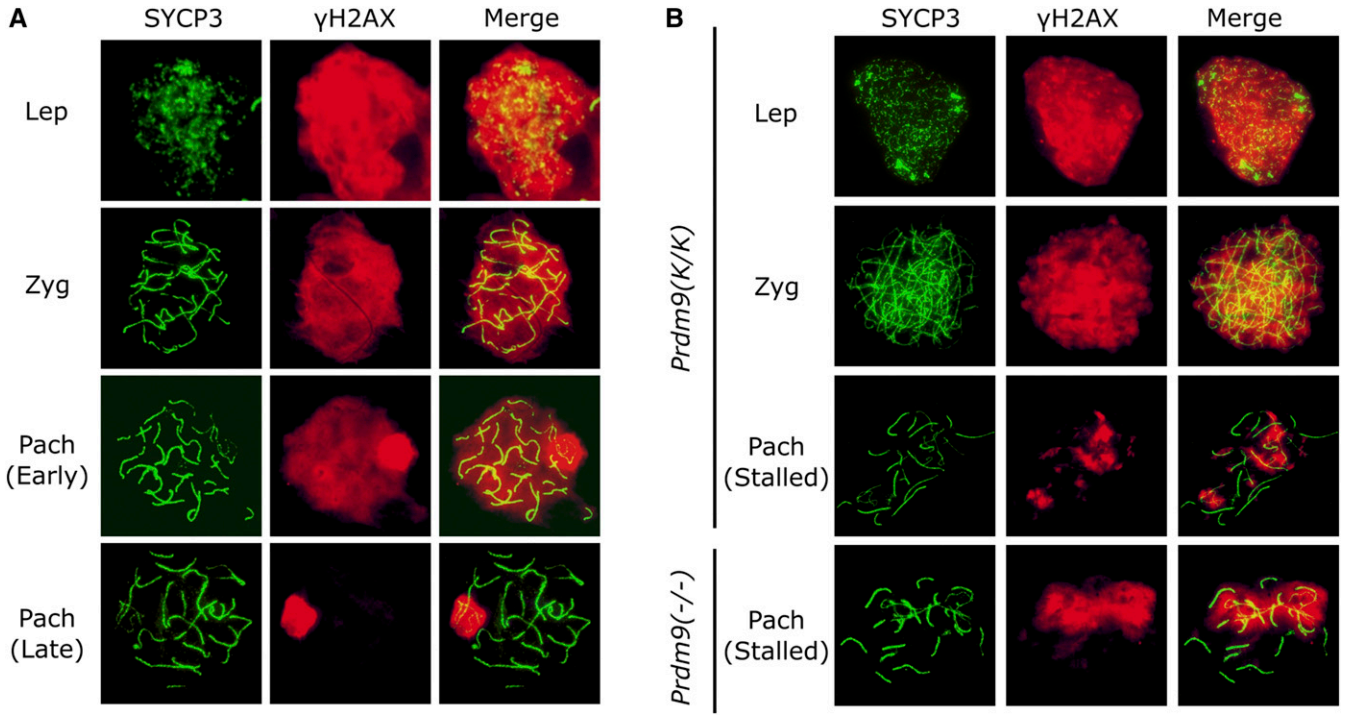

Figure 3 The KRAB domain of PRDM9 is essential for meiotic progression. Chromosome spreads from (A) *Prdm9*$^{+/+}$ or (B) *Prdm9*$^{K/K}$ and *Prdm9*$^{-/-}$ spermatocytes immunostained with SYCP3 and γH2AX antibodies as indicated. *Prdm9*$^{K/K}$ and *Prdm9*$^{-/-}$ arrest at pachytene-like stage with incomplete DSB repair and synapsis. Lep, leptonema; Zyg, zygonema; Pach, pachynema.

of them had a clear H3K4me3 signal, confirming that these were *bona fide* PRDM9-dependent hotspots. This indicates that the *Prdm9* mutant without the KRAB domain is largely incapable of trimethylating H3K4 at its binding sites *in vivo*, but DSB formation at the rare sites where such methylation occurs is not impaired.

Discussion

PRDM9 contains three different domains that have been implicated in regulation of transcription in other proteins: the KRAB domain, the SSXRD, and the PR/SET domain. Based on this, PRDM9 was assumed to be a transcription factor and early reports supported this assumption (Hayashi *et al.* 2005; Matsui and Hayashi 2007; Mihola *et al.* 2009; reviewed in Nowick *et al.* 2013; Capilla *et al.* 2016). Once the role of PRDM9 in defining recombination hotspots was identified (Baudat *et al.* 2010; Myers *et al.* 2010; Parvanov *et al.* 2010) it became clear that, rather than regulation of transcription, the methyltransferase activity of the PRDM9 PR/SET domain is needed to introduce H3K4me3 marks at the sites where recombination will be subsequently initiated. However, the role of the KRAB domain and the SSXRD remained uncertain. In this study, we found no change in global gene expression profile in *Prdm9*$^{-/-}$ mice, indicating that PRDM9 is not involved in transcription regulation; thus, the function of the KRAB and SSXRD domains of PRDM9 must be distinct from a role in transcriptional regulation.

The KRAB domain of PRDM9 has a noncanonical structure and shows ~25-fold weaker repression activity than the classic KRAB domains of the KRAB-ZFP family in *in vitro* assays (Lim *et al.* 1998). Therefore, involvement of the KRAB domain of PRDM9 in gene expression was somewhat unlikely. Indeed, two papers recently suggested that the role of the KRAB domain may be to anchor the PRDM9-bound DNA regions to the cores of meiotic chromosomes where the factors required for DSB formation are located. PRDM9 may exert this function through interaction with the CXXC1 protein (Imai *et al.* 2017; Parvanov *et al.* 2017), the mammalian homolog of yeast Spp1. Spp1 has a PHD domain that binds H3K4me3, but it also interacts with Mer2, the axis-associated protein required for DSB formation (Acquaviva *et al.* 2013; Sommermeyer *et al.* 2013). Therefore, Spp1 provides a link allowing H3K4 trimethylated chromatin to be physically anchored to the chromosome axis where much of the DSB formation machinery resides. Yeast lacks the PRDM9 protein, and recombination hotspots occur at H3K4me3 sites, usually promoters, regardless of the underlying DNA sequence (Wu and Lichten 1994; Borde *et al.* 2009; Pan *et al.* 2011). A similar situation occurs in *Prdm9* knockout mice, where DSBs are formed at promoter regions (Brick *et al.* 2012). In the presence of PRDM9, CXXC1 would interact with both PRDM9 (Imai *et al.* 2017; Parvanov *et al.* 2017) and H3K4me3 (Eberl *et al.* 2013) and then be tethered to the chromosome axis through interaction with IHO1 (Imai *et al.* 2017), the mammalian homolog of Mer2.

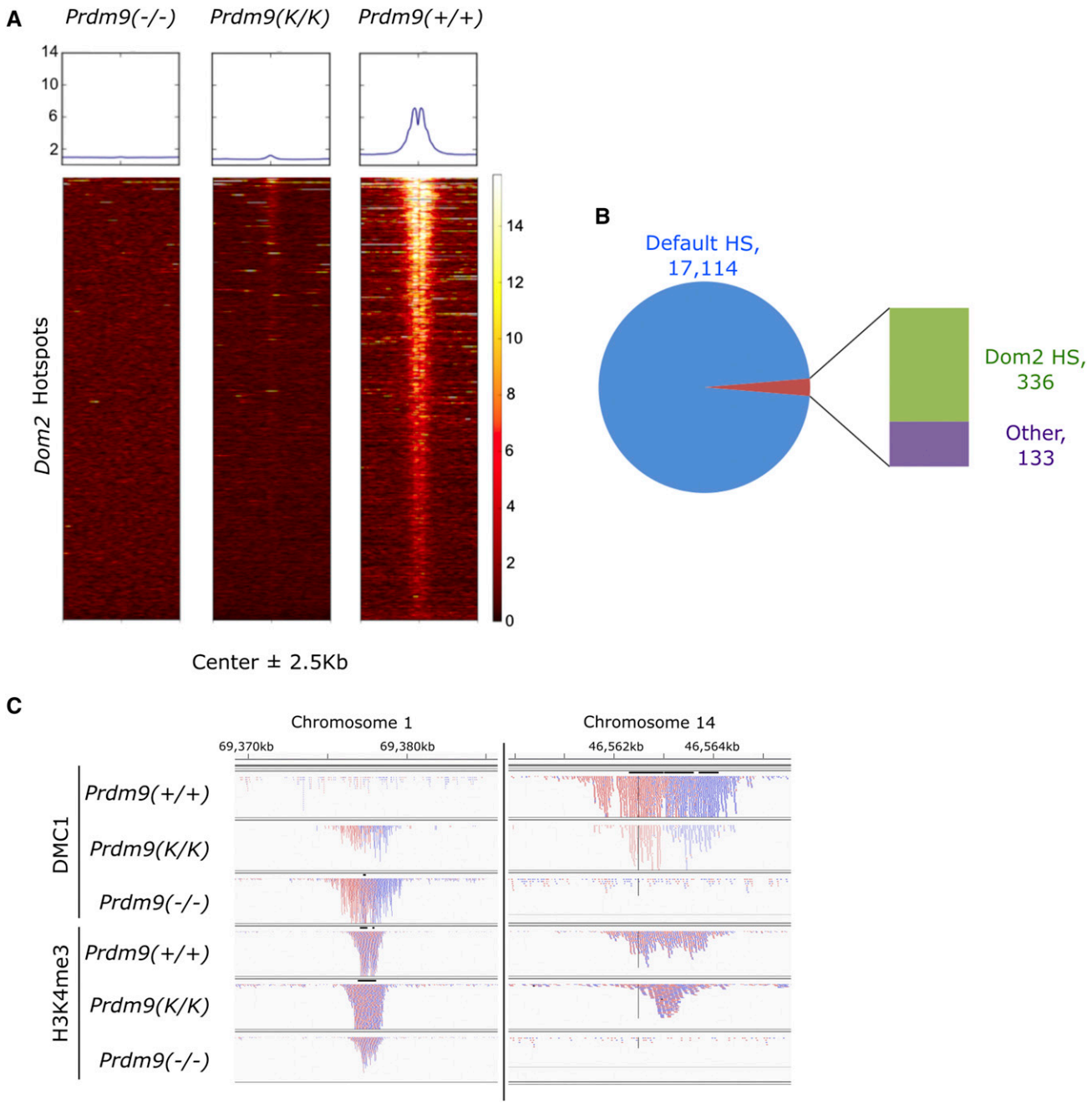

Figure 4 PRDM9 without the KRAB domain retains residual activity *in vivo*. (A) H3K4me3 ChIP-seq heat maps for *Prdm9*$^{-/-}$, *Prdm9*$^{K/K}$ and *Prdm9*$^{+/+}$ spermatocytes across 11,642 *Dom2* hotspots. (B) Overlap of DSB hotspots in *Prdm9*$^{K/K}$ with hotspots in *Prdm9*$^{-/-}$ (Default HS) and wild type (*Dom2* HS) mice. While the vast majority of DSBs occur at PRDM9-independent default hotspots there is a small number of PRDM9-directed DSBs. (C) Integrated Genome Viewer images of representative hotspots in *Prdm9*$^{K/K}$ mice that correspond to default hotspots (left) or to PRDM9-directed hotspots (right). SSDS tracks show sequencing reads from DMC1 ChIP-SSDS and H3K4me3 tracks show sequencing reads from H3K4me3 ChIP-seq.

Although the role of PRDM9 in bringing hotspot regions to the axis-associated DSB promoting factors is plausible, it does not explain why removal of the KRAB domain leads to the inhibition of the methyltransferase activity of PRDM9 during meiosis. The other six mammalian methyltransferases that methylate H3K4 (SET1A, SET1B and MLL1 through MLL4 in human) are catalytic cores in six nonredundant COMPASS complexes that have common, as well as complex-specific, subunits (Shilatifard 2012; Vedadi *et al.* 2017). Interestingly, CXXC1 that interacts with the KRAB domain of PRDM9 (Imai *et al.* 2017; Parvanov *et al.* 2017) is a subunit of the SET1A/B COMPASS, and removal of its yeast counterpart Spp1 leads to 80% reduction in the methyltransferase activity of the yeast COMPASS complex (Schneider *et al.* 2005). Therefore, it is conceivable that PRDM9 can also assemble into a COMPASS-like complex and recruit a CXXC1 subunit through its KRAB

domain. In this case, removal of the KRAB domain would result in dissociation of CXXC1 and loss of methyltransferase activity. Whether PRDM9 is indeed a part of a meiosis-specific canonical COMPASS complex or a part of an alternative complex, it looks like the KRAB domain has an earlier role in meiosis by facilitating H3K4 trimethylation and possibly other chromatin remodeling steps required for subsequent DSB formation and/or repair.

The deletion in one of our KRAB mutants (KRAB-2) included the entire KRAB domain and 24 out of 32 amino acids of the SSXRD. Since the resulting protein showed similar methyltransferase activity to the KRAB-1 mutant in cell culture, it appeared that the SSXRD did not have an additional role in the H3K4 trimethylation activity of PRDM9, at least in a nonmeiotic environment. It was therefore surprising that deletion of the entire SSXRD resulted in a complete loss of methyltransferase activity. Although it is possible that the last eight amino acids of the SSXRD are critical for its function, we cannot rule out the possibility that this deletion resulted in misfolding of the protein and indirect inhibition of the PR/SET domain.

In addition to the unknown roles of the KRAB domain and the SSXRD of PRDM9, the function of the single post-SET ZnF located right downstream of the SET domain was also unclear. While the other ZnFs in PRDM9 recognize and bind DNA, removal of this lone zinc finger does not affect the DNA-binding ability of PRDM9 in *in vitro* assays (Walker *et al.* 2015). At the same time, the post-SET ZnF was implicated in the autoregulation of the methyltransferase activity of PRDM9 through blocking the substrate binding site (Wu *et al.* 2013). If such regulation indeed occurred, removal of the post-SET ZnF would either increase the methyltransferase activity of PRDM9 (due to opening of the substrate binding site) or decrease it (if, for example, PRDM9 were automethylated and this inhibited the PR/SET activity and/or PRDM9 interactions with other proteins or DNA). As we found no significant changes in methyltransferase activity of PRDM9 with or without the post-SET ZnF in GC-1 cells, the regulatory role of the post-SET ZnF may be not that straightforward and needs to be evaluated *in vivo* using mouse models.

A recent study described the development of the *Prdm9* mutant mouse strain that lacked the N-terminal part of the protein including a portion of the KRAB domain (Imai *et al.* 2017). Although it was not clear whether the mutant phenotype was due to this truncation or because of substantially reduced expression level of the product, it matched the phenotype observed in our $Prdm9^{K/K}$ strain. We could not assess the level of PRDM9 protein in our mutant due to lack of antibodies of sufficient quality. However, the mRNA levels of $Prdm9^{K/K}$ and wild type *Prdm9* were similar (Figure S2), and, in cell culture, wild type *Prdm9* and the same KRAB mutant produced a comparable amount of protein product (Figure S1). Furthermore, we were able to detect residual H3K4me3 activity at the sites of expected hotspots and the evidence of DSB formation (Figure 4). If the KRAB domain of PRDM9 is required to communicate with the DSB formation machinery through interaction with CXXC1 (Imai *et al.* 2017) then deletion of this domain should prevent DSB formation at PRDM9 binding sites. The fact that these DSBs still form suggests that either (1) PRDM9 communicates with the DSB machinery through additional regions, outside the KRAB domain or (2) CXXC1 alone or within the SET1A/B COMPASS can interact with any H3K4me3 sites in the genome, regardless of whether they were introduced by PRDM9 or SET1A/B, and tether these sites to chromosome cores to enable DSB formation (Imai *et al.* 2017). Further studies will help to distinguish between these possibilities.

To summarize, we found that the PRDM9 protein is not involved in regulation of gene expression. Instead, the KRAB domain of PRDM9 appears to be required for the efficient H3K4 trimethylation activity of PRDM9 *in vivo*. The vast majority of meiotic DSBs in the $Prdm9^{K/K}$ mutants are formed at default hotspots, but those few H3K4me3 marks that are introduced by the KRAB-less PRDM9 are also recognized by the DSB formation machinery.

Acknowledgments

We thank Florencia Pratto for critical feedback on the manuscript, and members of the Camerini-Otero and Petukhova laboratories for helpful discussions and suggestions. We thank Jon Neumann and University of California (UC) Irvine Transgenic Mouse Facility for CRISPR targeting design and pronuclear injections. We thank both the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) genomics core and the American Genome Center at Uniformed Services University of Health Sciences (USUHS) for assistance with sequencing. This study used the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health (http://biowulf.nih.gov.lrc1.usuhs.edu). This study was supported by grants R01GM084104 and R01GM124402 from the National Institute of General Medical Sciences (G.V.P.) and the NIDDK Intramural Research Program (R.D.C.-O.).

Author contributions: S.T.-S. generated the mutant *Prdm9* constructs, performed the cell culture experiments and characterized the $Prdm9^{K}$ mouse model. F.S. cloned the PRDM9 cDNA and prepared RNA-seq libraries. J.C. constructed the miRNA libraries. S.T.-S. analyzed ChIP-seq data, Q.Y. analyzed the RNA-seq and microRNA-seq data. K.B. performed and supervised computational analyses. S.T.-S. and Q.Y. drafted the manuscript; S.T.-S., Q.Y., and K.B. designed and prepared the figures. All authors reviewed and commented on the manuscript. R.D.C.-O. and G.V.P. supervised the study.

Literature Cited

Acquaviva, L., L. Székvölgyi, B. Dichtl, B. S. Dichtl, C. de La Roche Saint André *et al.*, 2013 The COMPASS subunit Spp1 links histone methylation to initiation of meiotic recombination. Science 339: 215–218. https://doi.org/10.1126/science.1225739

Arnheim, N., P. Calabrese, and I. Tiemann-Boege, 2007 Mammalian meiotic recombination hot spots. Annu. Rev. Genet. 41: 369–399. https://doi.org/10.1146/annurev.genet.41.110306.130301

Baker, C. L., M. Walker, S. Kajita, P. M. Petkov, and K. Paigen, 2014 PRDM9 binding organizes hotspot nucleosomes and limits Holliday junction migration. Genome Res. 24: 724–732. https://doi.org/10.1101/gr.170167.113

Baker, C. L., P. Petkova, M. Walker, P. Flachs, O. Mihola *et al.*, 2015 Multimer formation explains allelic suppression of PRDM9 recombination hotspots. PLoS Genet. 11: e1005512. https://doi.org/10.1371/journal.pgen.1005512

Baker, Z., M. Schumer, Y. Haba, L. Bashkirova, C. Holland *et al.*, 2017 Repeated losses of PRDM9-directed recombination despite the conservation of PRDM9 across vertebrates. eLife 6: pii: e24133. https://doi.org/10.7554/eLife.24133

Baudat, F., K. Manova, J. P. Yuen, M. Jasin, and S. Keeney, 2000 Chromosome synapsis defects and sexually dimorphic meiotic progression in mice lacking Spo11. Mol. Cell 6: 989–998. https://doi.org/10.1016/S1097-2765(00)00098-8

Baudat, F., J. Buard, C. Grey, A. Fledel-Alon, C. Ober *et al.*, 2010 PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. Science 327: 836–840 (erratum: Science 328: 690). https://doi.org/10.1126/science.1183439

Billings, T., E. D. Parvanov, C. L. Baker, M. Walker, K. Paigen *et al.*, 2013 DNA binding specificities of the long zinc-finger recombination protein PRDM9. Genome Biol. 14: R35. https://doi.org/10.1186/gb-2013-14-4-r35

Borde, V., N. Robine, W. Lin, S. Bonfils, V. Géli *et al.*, 2009 Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites. EMBO J. 28: 99–111. https://doi.org/10.1038/emboj.2008.257

Bray, N. L., H. Pimentel, P. Melsted, and L. Pachter, 2016 Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. 34: 525–527 (erratum: Nat. Biotechnol. 34: 888). https://doi.org/10.1038/nbt.3519

Brick, K., F. Smagulova, P. Khil, R. D. Camerini-Otero, and G. V. Petukhova, 2012 Genetic recombination is directed away from functional genomic elements in mice. Nature 485: 642–645. https://doi.org/10.1038/nature11089

Buard, J., and B. de Massy, 2007 Playing hide and seek with mammalian meiotic crossover hotspots. Trends Genet. 23: 301–309. https://doi.org/10.1016/j.tig.2007.03.014

Capilla, L., R. A. Sánchez-Guillén, M. Farré, A. Paytuví-Gallart, R. Malinverni *et al.*, 2016 Mammalian comparative genomics reveals genetic and epigenetic features associated with genome reshuffling in Rodentia. Genome Biol. Evol. 8: 3703–3717. https://doi.org/10.1093/gbe/evw276

Clark, A. G., X. Wang, and T. Matise, 2010 Contrasting methods of quantifying fine structure of human recombination. Annu. Rev. Genomics Hum. Genet. 11: 45–64. https://doi.org/10.1146/annurev-genom-082908-150031

Davies, B., E. Hatton, N. Altemose, J. G. Hussin, F. Pratto *et al.*, 2016 Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice. Nature 530: 171–176. https://doi.org/10.1038/nature16931

Eberl, H. C., C. G. Spruijt, C. D. Kelstrup, M. Vermeulen, and M. Mann, 2013 A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics. Mol. Cell 49: 368–378. https://doi.org/10.1016/j.molcel.2012.10.026

Eram, M. S., S. P. Bustos, E. Lima-Fernandes, A. Siarheyeva, G. Senisterra *et al.*, 2014 Trimethylation of histone H3 lysine 36 by human methyltransferase PRDM9 protein. J. Biol. Chem. 289: 12177–12188. https://doi.org/10.1074/jbc.M113.523183

Everaert, C., M. Luypaert, J. L. V. Maag, Q. X. Cheng, M. E. Dinger *et al.*, 2017 Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data. Sci. Rep. 7: 1559. https://doi.org/10.1038/s41598-017-01617-3

Friedländer, M. R., S. D. Mackowiak, N. Li, W. Chen, and N. Rajewsky, 2012 miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res. 40: 37–52. https://doi.org/10.1093/nar/gkr688

Grey, C., P. Barthès, G. Chauveau-Le Friec, F. Langa, F. Baudat *et al.*, 2011 Mouse PRDM9 DNA-binding specificity determines sites of histone H3 lysine 4 trimethylation for initiation of meiotic recombination. PLoS Biol. 9: e1001176. https://doi.org/10.1371/journal.pbio.1001176

Grey, C., J. A. J. Clément, J. Buard, B. Leblanc, I. Gut *et al.*, 2017 In vivo binding of PRDM9 reveals interactions with noncanonical genomic sites. Genome Res. 27: 580–590. https://doi.org/10.1101/gr.217240.116

Handel, M. A., and J. C. Schimenti, 2010 Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. Nat. Rev. Genet. 11: 124–136. https://doi.org/10.1038/nrg2723

Hassold, T., H. Hall, and P. Hunt, 2007 The origin of human aneuploidy: where we have been, where we are going. Hum. Mol. Genet. 16: R203–R208. https://doi.org/10.1093/hmg/ddm243

Hayashi, K., K. Yoshida, and Y. Matsui, 2005 A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. Nature 438: 374–378. https://doi.org/10.1038/nature04112

Hofmann, M. C., S. Narisawa, R. A. Hess, and J. L. Millán, 1992 Immortalization of germ cells and somatic testicular cells using the SV40 large T antigen. Exp. Cell Res. 201: 417–435. https://doi.org/10.1016/0014-4827(92)90291-F

Hohenauer, T., and A. W. Moore, 2012 The Prdm family: expanding roles in stem cells and development. Development 139: 2267–2282. https://doi.org/10.1242/dev.070110

Imai, Y., F. Baudat, M. Taillepierre, M. Stanzione, A. Toth *et al.*, 2017 The PRDM9 KRAB domain is required for meiosis and involved in protein interactions. Chromosoma 126: 681–695. https://doi.org/10.1007/s00412–017–0631-z

Iyengar, S., and P. J. Farnham, 2011 KAP1 protein: an enigmatic master regulator of the genome. J. Biol. Chem. 286: 26267–26276. https://doi.org/10.1074/jbc.R111.252569

Keeney, S., and M. J. Neale, 2006 Initiation of meiotic recombination by formation of DNA double-strand breaks: mechanism and regulation. Biochem. Soc. Trans. 34: 523–525. https://doi.org/10.1042/BST0340523

Khil, P. P., F. Smagulova, K. M. Brick, R. D. Camerini-Otero, and G. V. Petukhova, 2012 Sensitive mapping of recombination hotspots using sequencing-based detection of ssDNA. Genome Res. 22: 957–965. https://doi.org/10.1101/gr.130583.111

Kozomara, A., and S. Griffiths-Jones, 2014 miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 42: D68–D73. https://doi.org/10.1093/nar/gkt1181

Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg, 2009 Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10: R25. https://doi.org/10.1186/gb-2009-10-3-r25

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lichten, M., 2008 Meiotic chromatin: the substrate for recombination initiation, pp. 165–193 in *Recombination and Meiosis*. Genome Dynamics and Stability. Springer, Berlin. https://doi.org/10.1007/7050_2008_040

Lim, F. L., M. Soulez, D. Koczan, H. J. Thiesen, and J. C. Knight, 1998 A KRAB-related domain and a novel transcription repression domain in proteins encoded by SSX genes that are disrupted in human sarcomas. Oncogene 17: 2013–2018. https://doi.org/10.1038/sj.onc.1202122

Liu, H., L. Chang, Y. Sun, X. Lu, and L. Stubbs, 2014 Deep vertebrate roots for mammalian zinc finger transcription factor subfamilies. Genome Biol. Evol. 6: 510–525. https://doi.org/10.1093/gbe/evu030

Love, M. I., W. Huber, and S. Anders, 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15: 550. https://doi.org/10.1186/s13059-014-0550-8

Lupo, A., E. Cesaro, G. Montano, D. Zurlo, P. Izzo *et al.*, 2013 KRAB-zinc finger proteins: a repressor family displaying multiple biological functions. Curr. Genomics 14: 268–278. https://doi.org/10.2174/13892029113149990002

Mack, K. L., and M. W. Nachman, 2017 Gene regulation and speciation. Trends Genet. 33: 68–80. https://doi.org/10.1016/j.tig.2016.11.003

Margolin, G., P. P. Khil, J. Kim, M. A. Bellani, and R. Daniel Camerini-Otero, 2014 Integrated transcriptome analysis of mouse spermatogenesis. BMC Genomics 15: 39. https://doi.org/10.1186/1471-2164-15-39

Matsui, Y., and K. Hayashi, 2007 Epigenetic regulation for the induction of meiosis. Cell. Mol. Life Sci. 64: 257–262. https://doi.org/10.1007/s00018-006-6281-6

Mihola, O., Z. Trachtulec, C. Vlcek, J. C. Schimenti, and J. Forejt, 2009 A mouse speciation gene encodes a meiotic histone H3 methyltransferase. Science 323: 373–375. https://doi.org/10.1126/science.1163601

Myers, S., R. Bowden, A. Tumian, R. E. Bontrop, C. Freeman *et al.*, 2010 Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. Science 327: 876–879. https://doi.org/10.1126/science.1182363

Narasimhan, V. M., K. A. Hunt, D. Mason, C. L. Baker, K. J. Karczewski *et al.*, 2016 Health and population effects of rare gene knockouts in adult humans with related parents. Science 352: 474–477. https://doi.org/10.1126/science.aac8624

Nowick, K., M. Carneiro, and R. Faria, 2013 A prominent role of KRAB-ZNF transcription factors in mammalian speciation? Trends Genet. 29: 130–139. https://doi.org/10.1016/j.tig.2012.11.007

Paigen, K., and P. Petkov, 2010 Mammalian recombination hot spots: properties, control and evolution. Nat. Rev. Genet. 11: 221–233. https://doi.org/10.1038/nrg2712

Pan, J., M. Sasaki, R. Kniewel, H. Murakami, H. G. Blitzblau *et al.*, 2011 A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. Cell 144: 719–731. https://doi.org/10.1016/j.cell.2011.02.009

Parvanov, E. D., P. M. Petkov, and K. Paigen, 2010 Prdm9 controls activation of mammalian recombination hotspots. Science 327: 835. https://doi.org/10.1126/science.1181495

Parvanov, E. D., H. Tian, T. Billings, R. L. Saxl, C. Spruce *et al.*, 2017 PRDM9 interactions with other proteins provide a link between recombination hotspots and the chromosomal axis in meiosis. Mol. Biol. Cell 28: 488–499. https://doi.org/10.1091/mbc.E16-09-0686

Patel, A., J. R. Horton, G. G. Wilson, X. Zhang, and X. Cheng, 2016 Structural basis for human PRDM9 action at recombination hot spots. Genes Dev. 30: 257–265. https://doi.org/10.1101/gad.274928.115

Persikov, A. V., R. Osada, and M. Singh, 2009 Predicting DNA recognition by Cys2His2 zinc finger proteins. Bioinformatics 25: 22–29. https://doi.org/10.1093/bioinformatics/btn580

Petukhova, G. V., P. J. Romanienko, and R. D. Camerini-Otero, 2003 The Hop2 protein has a direct role in promoting interhomolog interactions during mouse meiosis. Dev. Cell 5: 927–936. https://doi.org/10.1016/S1534-5807(03)00369-1

Pimentel, H., N. L. Bray, S. Puente, P. Melsted, and L. Pachter, 2017 Differential analysis of RNA-seq incorporating quantification uncertainty. Nat. Methods 14: 687–690. https://doi.org/10.1038/nmeth.4324

Ponting, C. P., 2011 What are the genomic drivers of the rapid evolution of PRDM9? Trends Genet. 27: 165–171. https://doi.org/10.1016/j.tig.2011.02.001

Powers, N. R., E. D. Parvanov, C. L. Baker, M. Walker, P. M. Petkov *et al.*, 2016 The meiotic recombination activator PRDM9 trimethylates both H3K36 and H3K4 at recombination hotspots in vivo. PLoS Genet. 12: e1006146. https://doi.org/10.1371/journal.pgen.1006146

Romanienko, P. J., and R. D. Camerini-Otero, 2000 The mouse Spo11 gene is required for meiotic chromosome synapsis. Mol. Cell 6: 975–987. https://doi.org/10.1016/S1097-2765(00)00097-6

Royo, H., H. Seitz, E. ElInati, A. H. F. M. Peters, M. B. Stadler *et al.*, 2015 Silencing of X–linked microRNAs by meiotic sex chromosome inactivation. PLoS Genet. 11: e1005461. https://doi.org/10.1371/journal.pgen.1005461

Schneider, J., A. Wood, J. Lee, R. Schuster, J. Dueker *et al.*, 2005 Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. Mol. Cell 19: 849–856. https://doi.org/10.1016/j.molcel.2005.07.024

Shilatifard, A., 2012 The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annu. Rev. Biochem. 81: 65–95. https://doi.org/10.1146/annurev-biochem-051710-134100

Smagulova, F., K. Brick, Y. Pu, U. Sengupta, R. D. Camerini-Otero *et al.*, 2013 Suppression of genetic recombination in the pseudoautosomal region and at subtelomeres in mice with a hypomorphic Spo11 allele. BMC Genomics 14: 493. https://doi.org/10.1186/1471-2164-14-493

Smagulova, F., K. Brick, Y. Pu, R. D. Camerini-Otero, and G. V. Petukhova, 2016 The evolutionary turnover of recombination hot spots contributes to speciation in mice. Genes Dev. 30: 266–280. https://doi.org/10.1101/gad.270009.115

Sommermeyer, V., and C. BéneutChaplais, E., M. E. Serrentino, and V. Borde, 2013 Spp1, a member of the Set1 complex, promotes meiotic DSB formation in promoters by tethering histone H3K4 methylation sites to chromosome axes. Mol. Cell 49: 43–54. https://doi.org/10.1016/j.molcel.2012.11.008

Song, R., S. Ro, J. D. Michaels, C. Park, J. R. McCarrey *et al.*, 2009 Many X–linked microRNAs escape meiotic sex chromosome inactivation. Nat. Genet. 41: 488–493. https://doi.org/10.1038/ng.338

Spandidos, A., X. Wang, H. Wang, and B. Seed, 2010 PrimerBank: a resource of human and mouse PCR primer pairs for gene expression detection and quantification. Nucleic Acids Res. 38: D792–D799. https://doi.org/10.1093/nar/gkp1005

Tam, S., M. Tsao, and J. D. McPherson, 2015 Optimization of miRNA-seq data preprocessing. Brief. Bioinform. 16: 950–963. https://doi.org/10.1093/bib/bbv019

Vedadi, M., L. Blazer, M. S. Eram, D. Barsyte-Lovejoy, C. H. Arrowsmith *et al.*, 2017 Targeting human SET1/MLL family of proteins. Protein Sci. 26: 662–676. https://doi.org/10.1002/pro.3129

Walker, M., T. Billings, C. L. Baker, N. Powers, H. Tian *et al.*, 2015 Affinity-seq detects genome-wide PRDM9 binding sites and reveals the impact of prior chromatin modifications on mammalian recombination hotspot usage. Epigenetics Chromatin 8: 31. https://doi.org/10.1186/s13072-015-0024-6

Wolfe, S. A., L. Nekludova, and C. O. Pabo, 2000 DNA recognition by Cys2His2 zinc finger proteins. Annu. Rev. Biophys. Biomol. Struct. 29: 183–212. https://doi.org/10.1146/annurev.biophys.29.1.183

Wu, H., N. Mathioudakis, B. Diagouraga, A. Dong, L. Dombrovski *et al.*, 2013 Molecular basis for the regulation of the H3K4

methyltransferase activity of PRDM9. Cell Rep. 5: 13–20. https://doi.org/10.1016/j.celrep.2013.08.035

Wu, T. C., and M. Lichten, 1994 Meiosis-induced double-strand break sites determined by yeast chromatin structure. Science 263: 515–518. https://doi.org/10.1126/science.8290959

Yu, Z., T. Raabe, and N. B. Hecht, 2005 MicroRNA Mirn122a reduces expression of the posttranscriptionally regulated germ cell transition protein 2 (Tnp2) messenger RNA (mRNA) by mRNA cleavage. Biol. Reprod. 73: 427–433. https://doi.org/10.1095/biolreprod.105.040998

Zhang, Y., T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson *et al.*, 2008 Model-based analysis of ChIP-seq (MACS). Genome Biol. 9: R137. https://doi.org/10.1186/gb-2008-9-9-r137

Communicating editor: D. Bishop