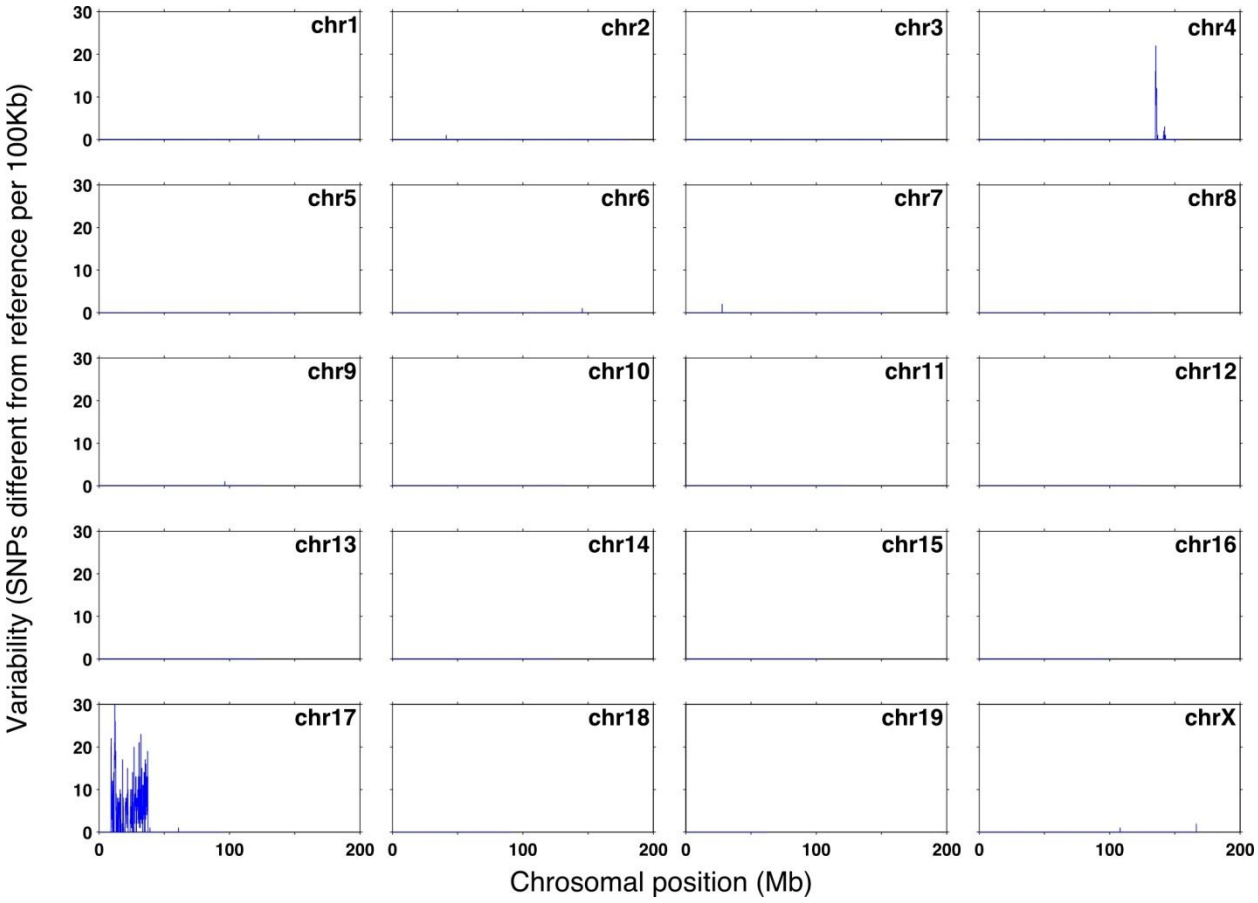
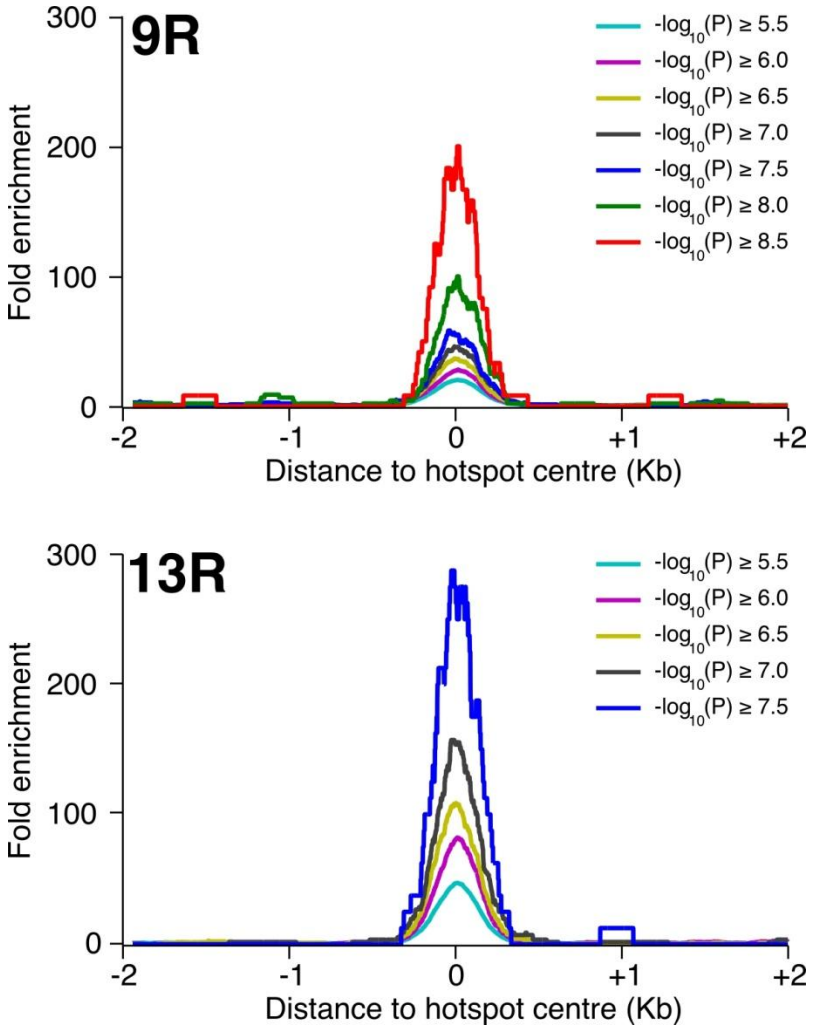
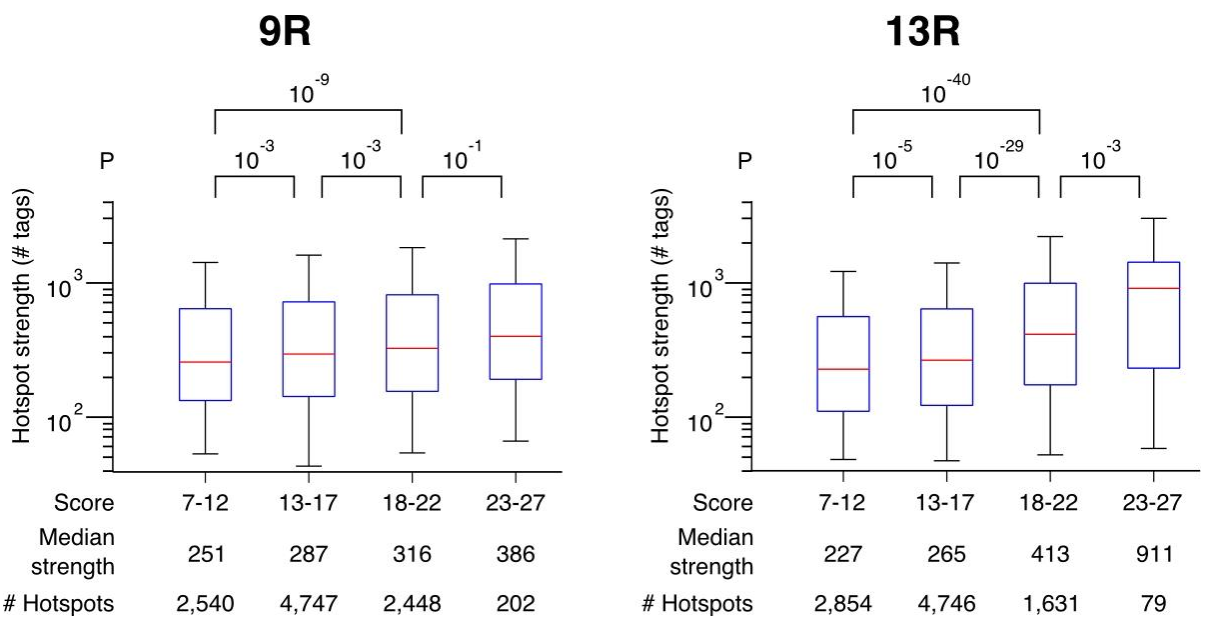
Supplementary Figures:



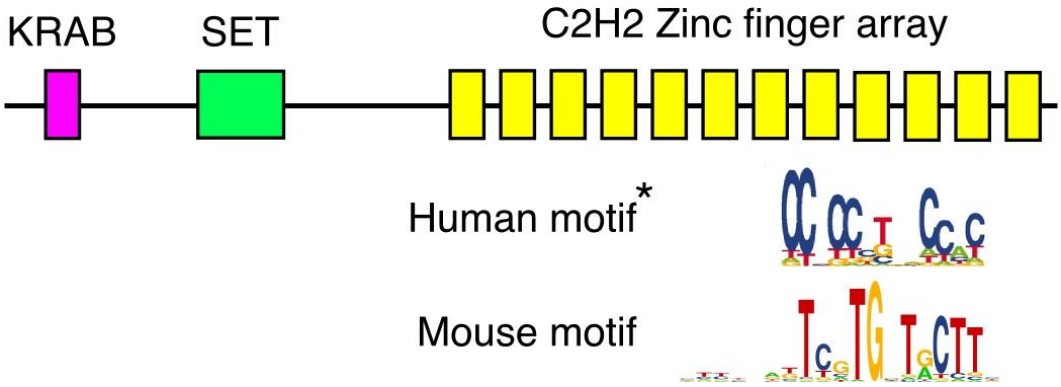
Supplementary Figure 1: 9R and 13R mice share an almost identical genetic background but different *Prdm9* alleles. Congenic 9R and 13R strains were produced by J.H. Stimpfling[31]. The variability between 9R and 13R mice was assessed using mouse diversity genotyping SNP arrays[32]. The variability is displayed as the number of SNPs which differ between two strains in each 100 Kb region across the genome. Most 9R and 13R chromosomes are identical at all SNP positions. The major exception is an approximately 30 Mb region on chromosome 17 that contains the *Prdm9* locus.



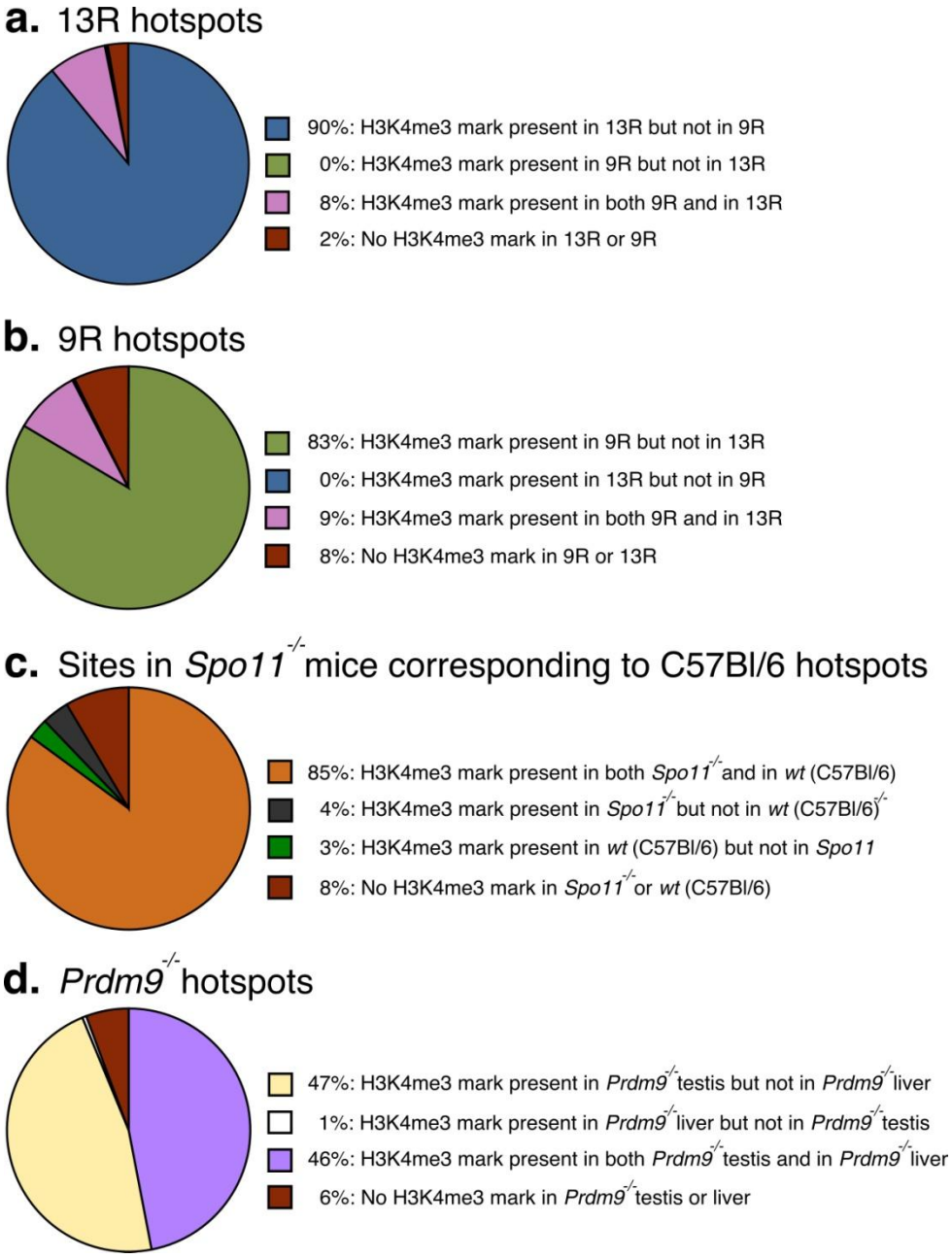
Supplementary Figure 2: Better scoring motifs are very highly enriched at hotspots. The spatial distribution around hotspot centres was generated for hits to the position specific scoring matrix. The MAST P-value (shown here as $-\log_{10}(P)$) reflects the quality of a motif hit. Different P-value thresholds were used to show that the motif enrichment at hotspots increases dramatically for better hits. This is true for both the 9R and the 13R motif. Distributions are generated with a 200 bp sliding window in single nucleotide steps. Enrichment is given relative to the mean motif frequency at each P-value cutoff in the 2.5 Kb regions flanking hotspots.



Supplementary Figure 3: Hotspot strength is correlated with the quality of the motif alignment. The quality is given as the best motif alignment score in the central 1 Kb of each hotspot (Score). Scores are rounded to the nearest integer. Hotspot strength is calculated as the number of ssDNA tags. Though a similar association is observed in both genetic backgrounds, this effect is more pronounced in 13R than for 9R. P-values are calculated using a Wilcoxon Rank Sum Test.

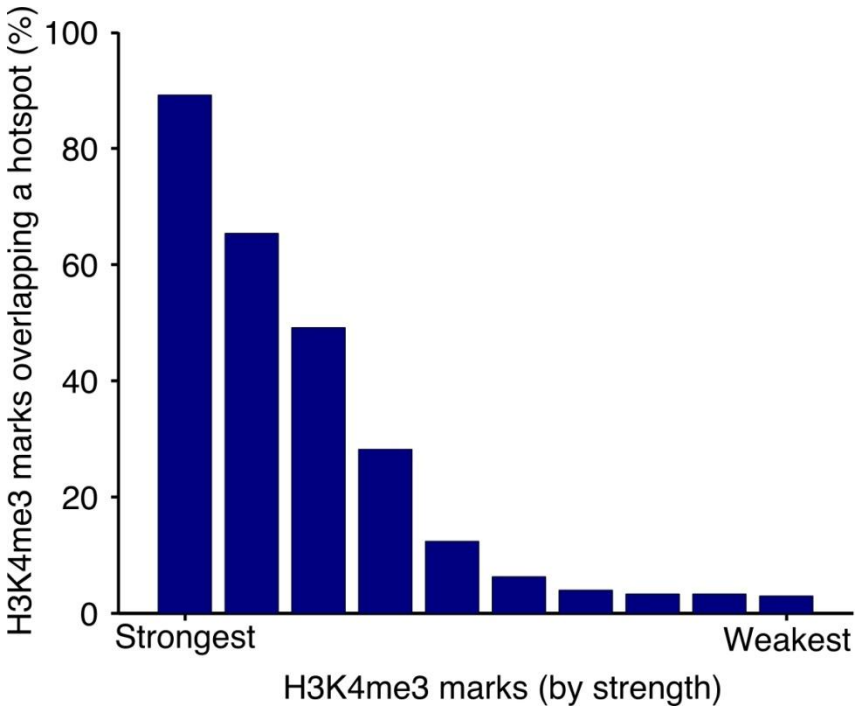


Supplementary Figure 4: Previously published DNA sequence motifs associated with mouse and human hotspots both align to the C-terminal end of the PRDM9 Zinc finger array. This schematic illustrates the domain structure of PRDM9. It is not drawn to scale. Furthermore, the PRDM9 alleles in mice and humans differ but are illustrated here by a single schematic. Alignments of human and mouse motifs are based on data from [2] and [12] respectively.

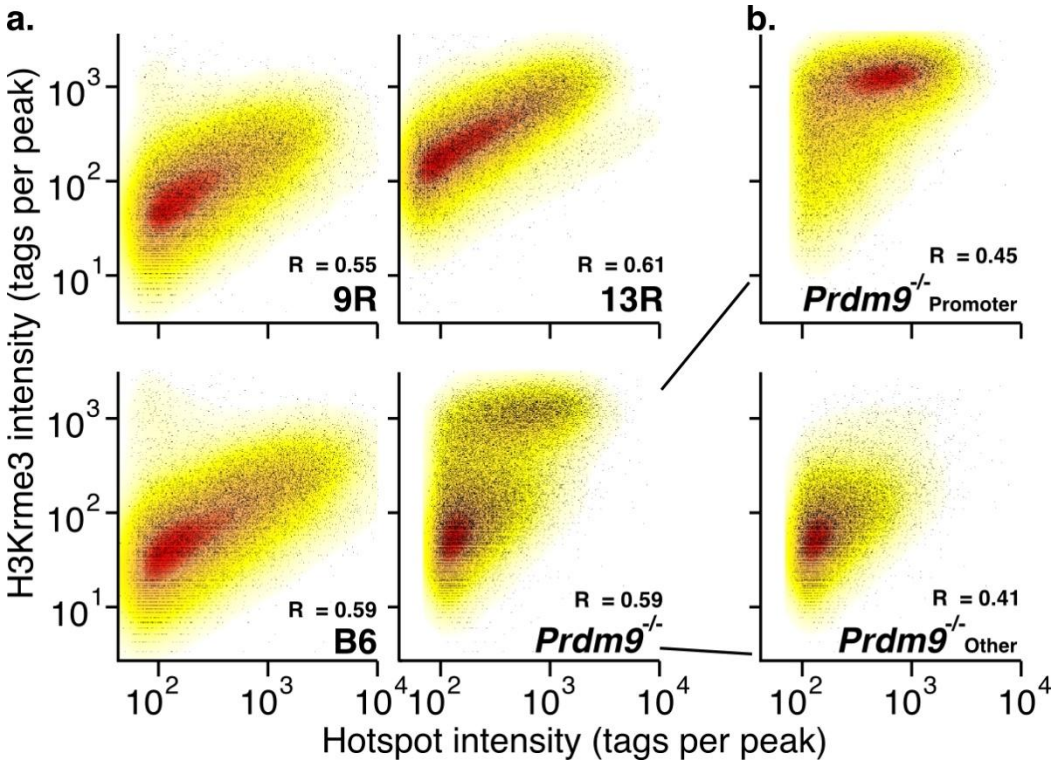


Supplementary Figure 5: H3K4me3 marks are present at sites of DSB hotspots. Pie charts show the proportion of DSB hotspots that overlap different subsets of H3K4me3 marks. Each panel represents hotspots from a different genetic background; a. 13R; b. 9R; c. Sites in *Spo11*$^{-/-}$ mice which correspond to C57Bl/6 hotspots; d: *Prdm9*$^{-/-}$. In a. 13R and b. 9R mice, most hotspots are associated with strain-specific H3K4me3 marks. The overlap with strain specific H3K4me3 sites is likely an under-estimate as the majority of shared H3K4me3 sites are actually PRDM9-

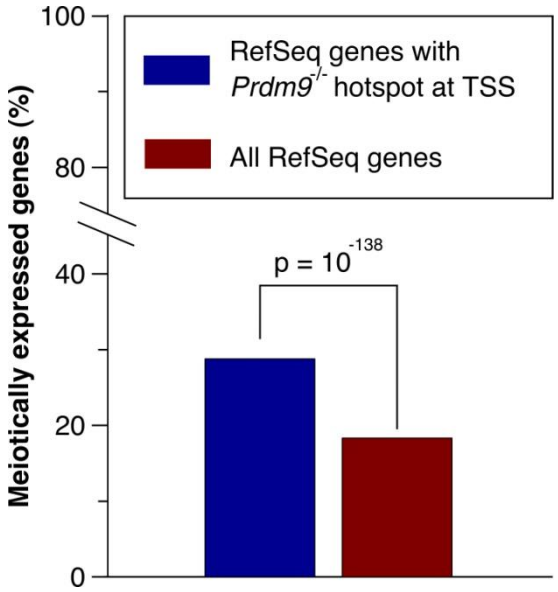
dependent H3K4me3 sites that cannot be resolved from a strong, adjacent PRDM9-independent H3K4me3 site. c. *Spo11* knockout mice are defective for DSB formation, however, the sites corresponding to potential DSB hotspots (from wild-type C57Bl/6 mice) are marked by H3K4me3. *Spo11* knockout mice have the C57Bl/6 *Prdm9* allele. Of the 89% of potential hotspots which overlap *Spo11*$^{-/-}$ H3K4me3 marks, 90% are found in B6 and not in 13R. d. DSB hotspots in *Prdm9*$^{-/-}$ mice are associated with H3K4me3. Many of these hotspot-associated H3K4me3 sites are also present in liver tissue. Hotspot associated H3K4me3 marks are not present in liver tissue in mice with a functional *Prdm9* allele ([12] and data not shown). All overlaps were counted between the central 1 Kb regions of hotspots and H3K4me3 marks.



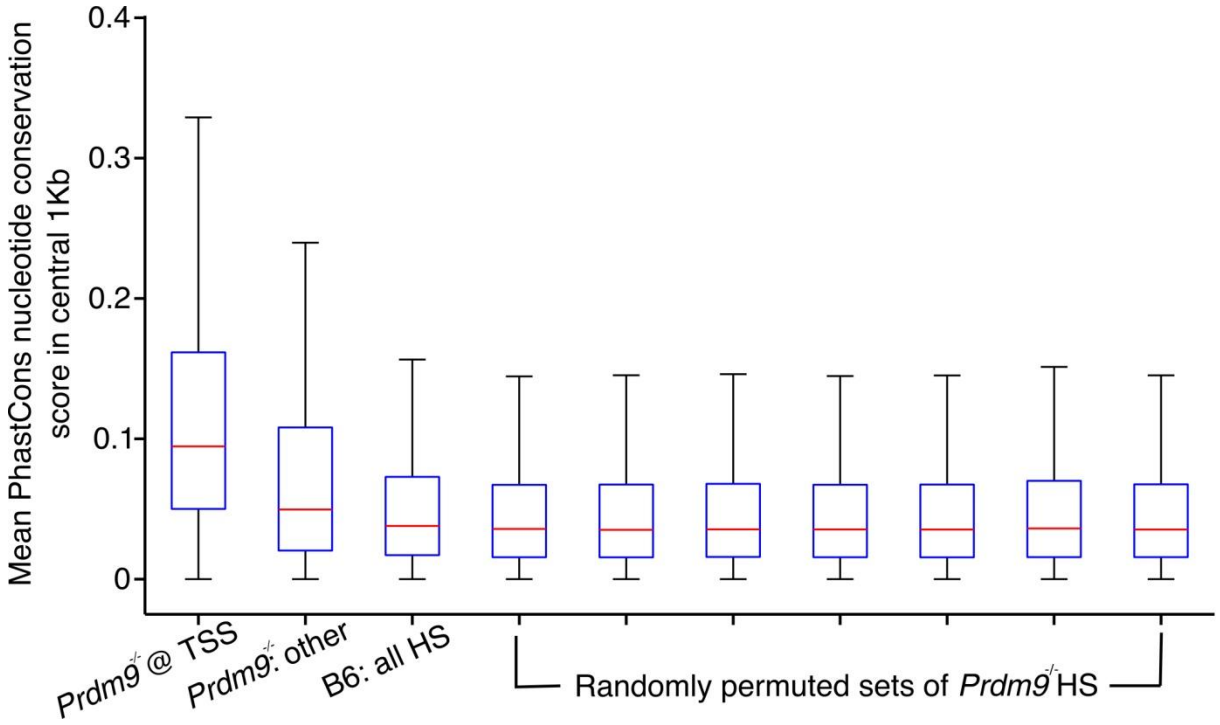
*Supplementary Figure 6***.** Many H3K4me3 marks in *Prdm9*$^{-/-}$ testis do not overlap hotspots. H3K4me3 marks in *Prdm9*$^{-/-}$ testis were binned by strength (x-axis). The percentage of H3K4me3 marks in each bin that overlapped a DSB hotspot is plotted. A significant proportion of H3K4me3 marks did not overlap a DSB hotspot, though the magnitude of this overlap is clearly strength dependent.



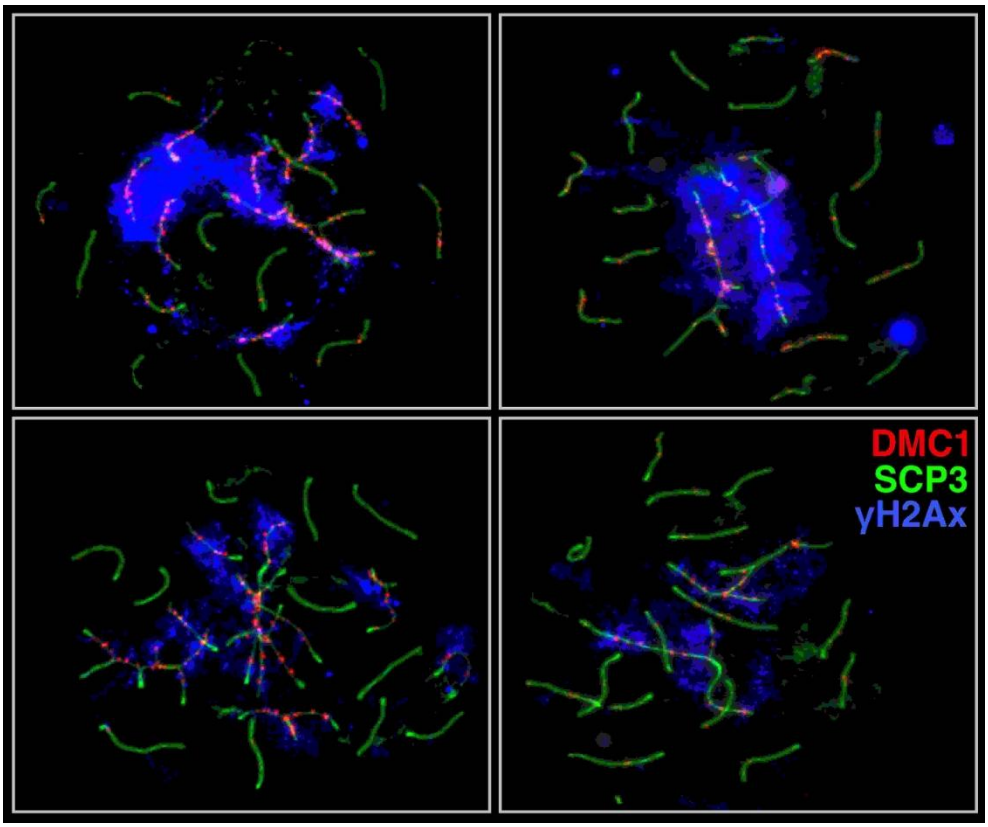
Supplementary Figure 7: Density scatter plots of hotspot strength versus H3K4me3 intensity inside hotspots. The DSB hotspot dataset is given in the bottom right corner of each panel. The spearman correlation coefficient (R) is given above. a. In 9R, 13R, and C57Bl/6 (B6) mice, the hotspot intensity correlates with the intensity of the H3K4me3 signal inside the hotspot. In *Prdm9*$^{-/-}$ mice, a more complex, bipartite correlation pattern is observed. This bipartite distribution gives an artificially high R (0.59). b. In *Prdm9*$^{-/-}$ mice, hotspots at promoters have a different relationship to H3K4me3 than those which are not at promoters. Most strong DSB hotspots in *Prdm9*$^{-/-}$ mice are at promoters, where the H3K4me3 signal is generally high (upper panel). Hotspots away from promoters are generally weaker and correlate with weaker H3K4-trimethylation (lower panel). It should be noted that the H3K4me3 strength at promoters may not be an accurate reflection of the H3K4me3 intensity in the cells where DSBs are formed. This is because these H3K4me3 marks are often persistent (present in tissue as diverse as liver), therefore the H3K4me3 signal at promoters is likely a cumulative signal from many cell types.



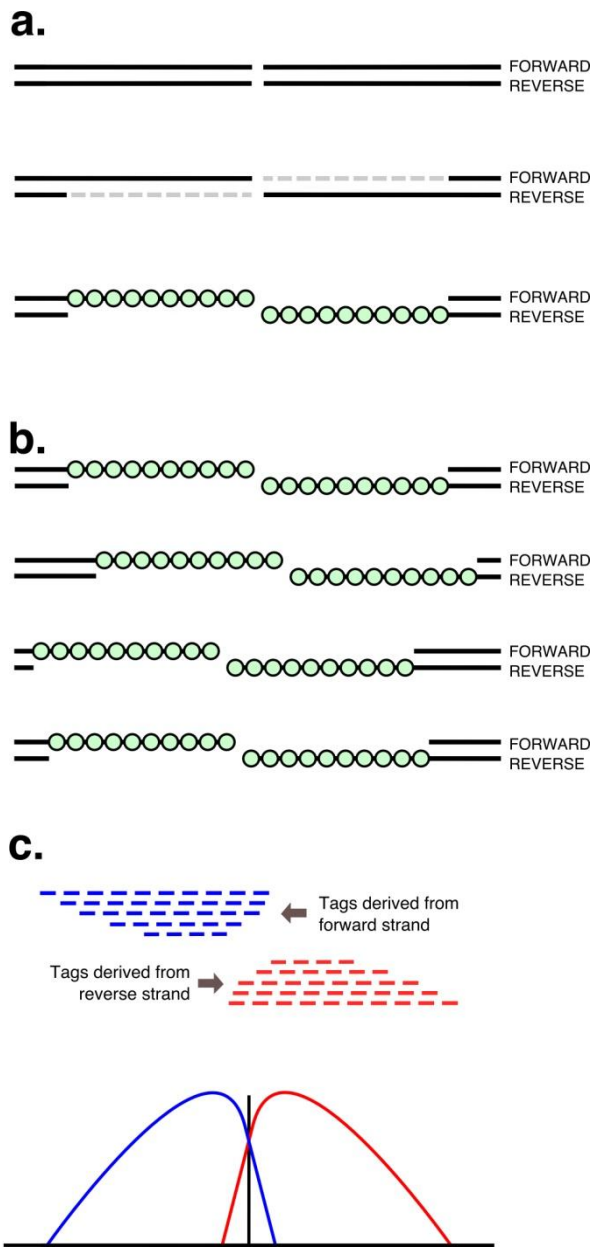
Supplementary Figure 8: Genes with a hotspot at the TSS in *Prdm9*$^{-/-}$ mice are expressed in meiosis more frequently than expected. Meiotic expression data were obtained from [33]. The expected fraction is defined as the fraction of all RefSeq genes expressed in this dataset. Genes were defined as meiotically expressed if they were assigned to expression cluster B, C or D in these data. Probability was defined using a one sided binomial test.



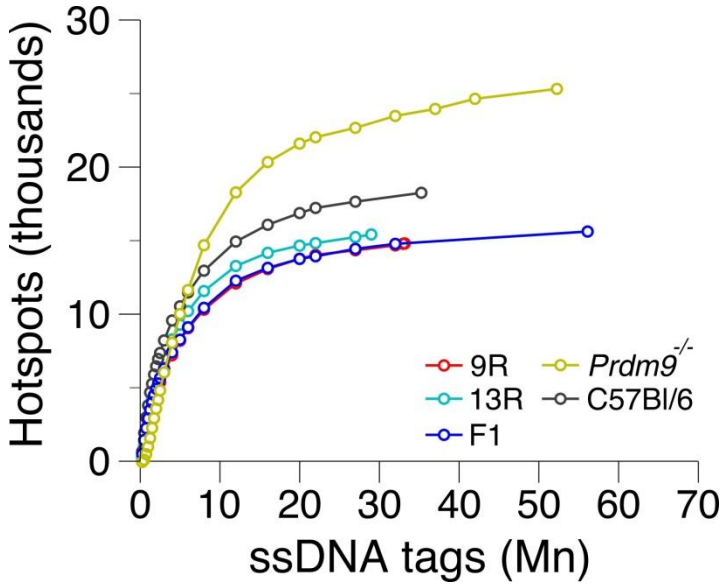
Supplementary Figure 9: Prdm9$^{-/-}$ hotspots are located in conserved genomic regions. The placental mammals PhastCons[34] conservation track was obtained from the UCSC genome browser [35]. This track was generated from the comparison of placental mammals and provides a conservation score from 0 (low) to 1 (high) for each nucleotide in the mouse genome. Here, we scored hotspots by calculating the mean PhastCons score across all nucleotide positions in the central 1 Kb. Boxplots illustrate the range of values for each set of hotspots (given on x-axis) and the median for each set is illustrated by a red bar. The highest conservation is observed in *Prdm9*$^{-/-}$ hotspots overlapping TSS. This is expected as the sequence at promoters is known to be well conserved. Sequence conservation at other *Prdm9*$^{-/-}$ hotspots (not at TSS) is also significantly elevated ($P < 10^{-100}$; Wilcoxon Rank Sum Test; compared to permuted set of hotspots). This suggests that at least some of these hotspots overlap putative functional elements. The permuted sets were generated as described in [12]. The conservation at B6 hotspots is similar to the genomic average as given by the permuted sets.



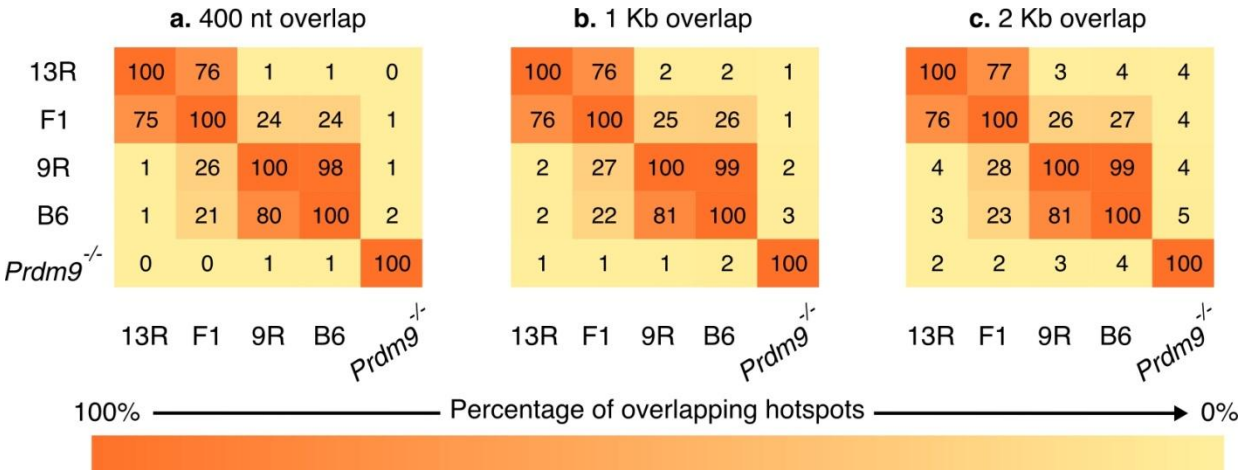
Supplementary Figure 10: Typical pachytene-like cells in *Prdm9*$^{-/-}$ mice showing DMC1 foci associated with γH2Ax on the asynapsed chromosomes. Though asynapsed chromosomes are apparent in almost all pachytene-like cells, the majority of chromosomes are consistently synapsed. DMC1 is stained red, SCP3 is stained green, γH2Ax is stained blue. Spermatocyte chromosome surface spread preparations were prepared using slight modifications[36] to the method of [37].



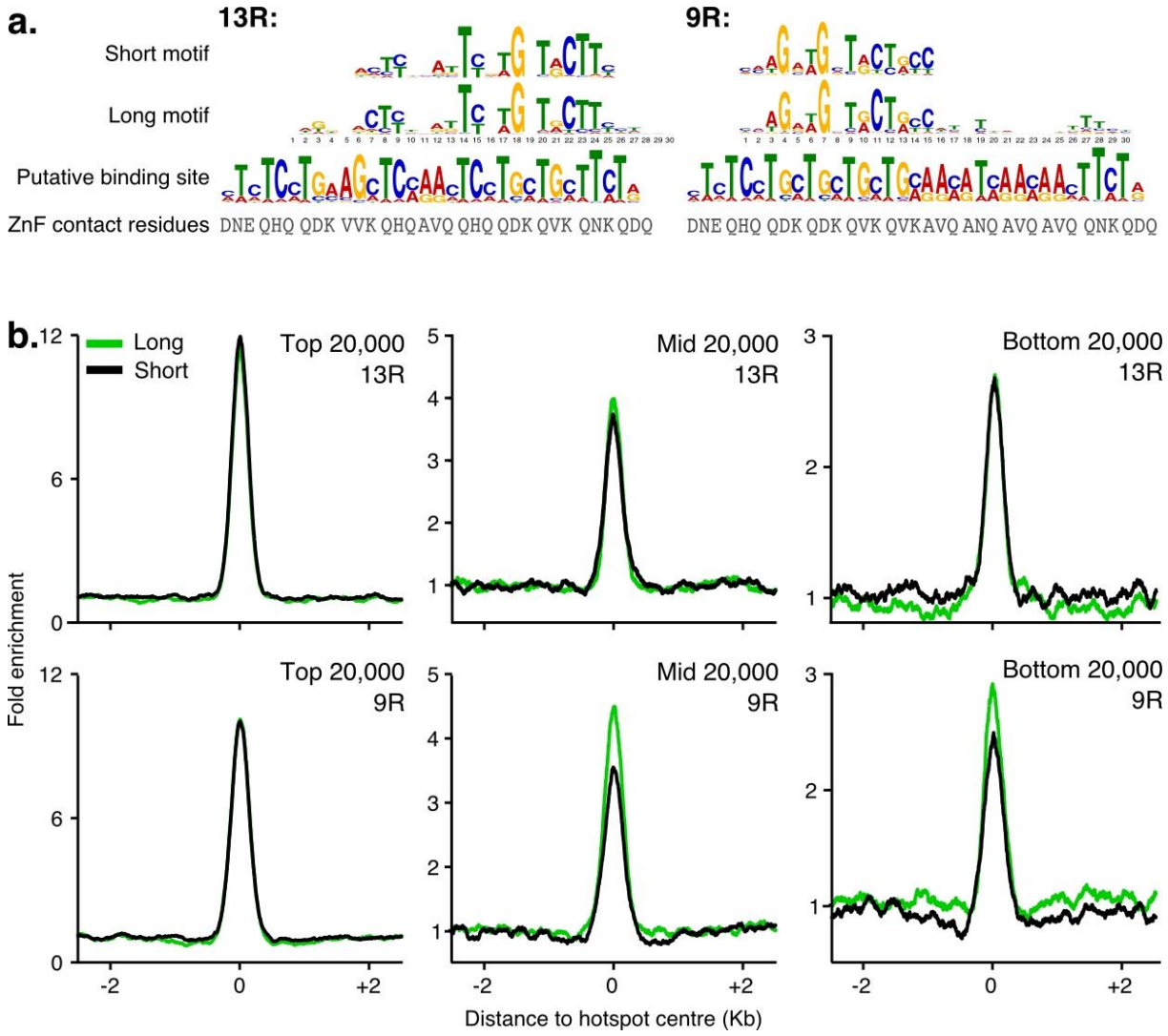
Supplementary Figure 11: SSDS generates a hotspot-specific signal. a. Meiotic recombination is initiated by the formation of a DNA (black lines) DSB by SPO11. Nuclease digestion results in the formation of 3' overhangs, which are subsequently bound by the recombinase DMC1 (green circles). This DMC1-bound ssDNA is the substrate for SSDS. b. DSBs form at various positions within a hotspot. c. SSDS allows sequencing of ssDNA fragments derived from the 3' overhangs while retaining the strand-of-origin (forward/blue or reverse/red) information for each fragment. This yields a very characteristic sequencing tag coverage pattern at hotspots (lower panel).



Supplementary Figure 12: All hotspot samples are sequenced almost to saturation for hotspot detection. To generate saturation curves, tag subsets of different sizes (x-axis) were randomly chosen from the full dataset and subsequently used to call peaks. The numbers of peaks identified for each subset are shown on the y-axis. From these plots it is clear that further sequencing will yield few new hotspots and that peak calling dynamics for *Prdm9*$^{-/-}$ hotspots differ from that for the other genetic backgrounds.



Supplementary Figure 13. Most shared DSB hotspots are detected using a 400 nt overlap window. Overlapping hotspots were counted between strains. Only overlaps between the central a. 400 nt b. 1 Kb or c. 2 Kb of hotspots were counted. The percentage of overlapping hotspots is given for each pair of strains. Comparisons were performed in both directions (i.e. B6 hotspots overlapping 9R AND 9R hotspots overlapping B6), as hotspot numbers were different for each strain. The calculation of overlaps using the 2 Kb overlap window only marginally increased the number of overlapping hotspots between strains (1% to 3%). Furthermore, the overlaps for strains sharing a *Prdm9* allele increased by a similar amount to the overlaps between strains with different *Prdm9* alleles. This suggests that the use of a larger overlap window resulted in more stochastic overlaps for all comparisons.



Supplementary Figure 14: Identification of 30-mer motifs at hotspots. MEME-ChIP[38] was constrained to identify motifs of 30 nt in each set of DSB hotspots. a. While overall, the residues outside of the core motif were information poor, residues at position 19, 27, 28 and 29 for 9R and at position 3 in the 13R 30-mer motifs agreed with the predicted PRDM9 binding sites for each allele and support the alignment positions of the short motifs. b. Use of the longer motif improves the specificity of weak motifs around DSB hotspots. The top 60,000 hits around hotspots (± 2.5 Kb) were considered for each motif. Hits were split into three equal bins based on the motif alignment score. The longer version of the 13R motif shows the same enrichment as the short version for each group of hits. For 9R, the longer version of the motif increases the specificity for weaker hits.

Methods:

Mouse strains

13R (alternative name B10.F-H2pb1/(13R)J, Jackson Labs stock number 001818) and 9R (B10.S-H2t4/(9R)/J, Jackson Labs stock number 001650) were received from Dr. N. Arnheim, University of Southern California. 9R and 13R mice were bred to produce a 9R/13R F1. *Prdm9*$^{-/-}$ mice were obtained from Jackson Labs (Jackson Labs stock number 010719). *Hop2*$^{-/-}$ mice on 9R and 13R background were previously described[12]. *Spo11*$^{-/-}$ mice (strain *deltaSpo.BC/B6*) have been produced by the deletion of exons 2-6 of the *Spo11* gene resulting in the absence of the SPO11 protein (P. J. Romanienko and Camerini-Otero, unpublished). The phenotype of this knockout strain is identical to the phenotype of the previously described[39] *Spo11*$^{-/-}$ strain *Spo11*^{tm1Rdco}. The *Spo11*$^{-/-}$ *(DeltaSpo.BC/B6*) mice used here were of a mixed 129S6/SvEvTac and C57Bl/6J background. 129S6/SvEvTac and C57Bl/6J share the same PRDM9 allele. All experiments were performed using adult (2-6 months old) mice. All animal procedures have been approved by the USUHS Animal Care and Use Committee.

Antibodies

The following antibodies were used for ChIP: anti-DMC1, Santa Cruz (C-20, sc 8973), anti-H3K4me3, Millipore (#07-473). The following antibodies were used for immunofluoresence microscopy: anti-SCP3, Santa Cruz (sc 74569), anti-DMC1, Santa Cruz (sc 22768), anti-γH2Ax, Upstate (16-193).

Sequencing library preparation:

We have previously described a method (Single-Stranded DNA Sequencing; SSDS) that allows highly specific and sensitive DSB hotspot detection in wild-type mice via identification of protein bound ssDNA after immunoprecipitation with DMC1. Here, we used this method (as described in [13]) to detect DSB hotspots. Briefly, sheared chromatin is immunoprecipitated with an αDMC1 antibody. Kinetic enrichment for ssDNAs is performed followed by paired-end high throughput sequencing.

The H3K4me3 ChIP-Seq libraries were prepared as described in [12].

Sequencing:

High-throughput sequencing was performed using the Illumina GaIIx and HiSeq 2000 platforms. Tags from SSDS experiments were aligned to the mouse mm9/NCBI37 genome using the ssDNA mapping pipeline (modified version of bwa[40]) described in[13]. Tags from H3K4me3 samples were aligned to the mouse mm9/NCBI37 genome using ELAND. Only uniquely mapping reads with less than three mismatches and which passed the quality filter were used.

Samples:

Mouse strain:	***Replicate***	***Antibody:***	***Method:***	***Sequencing type:***	***Accession #:***
C57BL/6J wt	1	Dmc1	SSDS	Paired-end	GSM869781
C57BL/6J wt	2	Dmc1	SSDS	Paired-end	GSM869782
13R wt	1	Dmc1	SSDS	Paired-end	GSM869783
13R wt	2	Dmc1	SSDS	Paired-end	GSM869784
9R wt	1	Dmc1	SSDS	Paired-end	GSM869785
9R wt	2	Dmc1	SSDS	Paired-end	GSM869786
9R / 13R F1 wt	1	Dmc1	SSDS	Paired-end	GSM869787
9R / 13R F1 wt	2	Dmc1	SSDS	Paired-end	GSM869788
9R / 13R F1 wt		n/a : Input	SSDS	Paired-end	GSM869789
Prdm9$^{-/-}$		Dmc1	SSDS	Paired-end	GSM869790
C57BL/6J wt	1	H3K4me3	ChIP-Seq	Single read	GSM869791
C57BL/6J wt	2	H3K4me3	ChIP-Seq	Single read	GSM869792
13R *Hop2*$^{-/-}$		H3K4me3	ChIP-Seq	Single read	GSM869793
9R *Hop2*$^{-/-}$		H3K4me3	ChIP-Seq	Single read	GSM869794
Prdm9$^{-/-}$: Testis		H3K4me3	ChIP-Seq	Single read	GSM869795
Prdm9$^{-/-}$: Liver		H3K4me3	ChIP-Seq	Single read	GSM869796
Spo11$^{-/-}$		H3K4me3	ChIP-Seq	Single read	GSM869797

Data processing and hotspot morphology:

SSDS post-processing entails computational identification of ssDNA-specific sequencing

signatures. This allowed us to specifically enrich for and detect DMC1-bound ssDNA. Fragments originating from dsDNA were excluded from our analyses. ssDNA-derived tags from different samples with the same genetic background were pooled. The strand of origin can also be determined for each sequencing read derived from ssDNA. This yields a very characteristic read coverage pattern around DSB hotspots; strand resection occurs on opposite strands either side of a DSB. Specifically, tags to the left of the DSB centre will map to the forward strand, while those to the right will map to the reverse strand (Supplementary Fig. 11).

Peak calling:

Peak calling algorithms for ChIP-Seq data have been generally designed to identify short (up to several hundred nucleotides) binding sites. The ssDNA signal around hotspots spans approximately 2 Kb on average. We used the SISSRS algorithm[29] for peak calling as it specifically uses the directionality of reads to define binding sites. We used the following parameters to run SISSRS: –a; -F 1000; -s 2,700,000,000; -E 10. SISSRs peaks within 100nt of each other were merged. The centre of each SISSRS peak was then redefined as the midpoint of the median of the forward and reverse strand tag distributions. If the median of the forward distribution is not to the left of that of the reverse fragment distribution, the peak was split in two and the distribution median re-examined in each new peak. Three iterations of this procedure were permitted and regions that did not meet these criteria were discarded. Peak deconvolution was particularly challenging when a weak peak overlapped a stronger one and this resulted in slight centring inaccuracies for a small fraction of the strongest peaks. SSDS has been shown to dramatically reduce the genomic background, therefore our peak calling strategy relies on a very low background level in the samples.

Genomic regions with high ssDNA background are problematic for such median-based peak calling. Problematic regions for peak calling were initially defined arbitrarily as the 0.1% of 25 Kb non-overlapping windows with the highest non-specific background (tags outside of hotspots) in each strain (9R, 13R, F1, *Prdm9*$^{-/-}$). Windows identified in each strain were pooled and this yielded a total of 106 25 Kb genomic intervals. These intervals were subsequently manually curated. Adjacent intervals were merged as were those spanning regions of low coverage. This yielded 91 contiguous regions. Most regions (84 of 91) had a comparable number of ssDNA tags in the control sample as in the other samples. 5 of 91 regions were included

because of the presence of strong apparent hotspots in one strain that were not called due to the absence of tags from one strand. The remaining two regions are found at the PAR (and internal to it) on chromosome X and at the non-centromeric end of chromosome 13. In each of these two regions, there is very high background in both the B6 and 13R strains but not in the control sample. Many apparent peaks in these two regions are present in the *Prdm9*$^{-/-}$ mice. The high background in these two regions may be indicative of DSB formation outside of strictly defined hotspots.

Given the broad and highly characteristic nature of hotspot signals, we expected very high specificity for peak calling. To estimate the false positive rate, we identified hotspots using 20 million ssDNA tags from the F1 αDMC1 ChIP-Seq sample and from the F1 control sample in which αDMC1 ChIP was not performed. 844 peaks were identified in the control, of which 414 occurred outside of the high background regions defined above. This compares to 13,819 in the F1 sample or an FDR of 3%. We performed peak calling using more permissive criteria (SISSRS –E 5), however this principally increased the false positive rate and only negligibly increased the true hotspot detection rate (not shown). We also performed a saturation analysis to evaluate the potential benefit of further sequencing (Supplementary Fig. 12). All samples were sequenced close to saturation – i.e. addition of further sequencing tags would yield few new hotspots. In fact, relaxing the peak calling criteria had the effect of shifting the saturation curve to the left; that is, hotspots were detected with fewer tags, however, few new hotspots were detected.

The number of hotspots detected is likely to depend on the ChIP and subsequent enrichment steps. The enrichment obtained experimentally can be estimated by calculating the sample specificity, defined as the fraction of sequencing reads which fall within hotspots. If this specificity is similar between the samples, it is reasonable to compare the number of determined hotspots. The specificity of the 9R, 13R, F1 and B6 samples are 29%, 30%, 35% and 45% respectively. The enrichment of the *Prdm9*$^{-/-}$ sample is 22%.

MACS[30] (v.1.3.7) was used to define peaks in the H3K4me3 ChIP-Seq data using a p-value of 0. 01. Local lambda was disabled to facilitate detection of weak, hotspot-associated peaks adjacent to larger peaks.

Overlapping intervals:

When comparing DSB hotspot positions, only the central 400 nt of each hotspot was used, in

order to minimize false positive overlaps. We note that one consequence of such stringent criteria is that the reported percentage overlaps were artificially low for some of the stronger hotspots, where multiple adjacent peaks could not be resolved. The reported overlap for very weak hotspots was also artificially low as centre definition is less robust and can be greatly influenced by few sequencing tags. Expanding the overlap window to 2 Kb allows us to estimate the maximum extent of these effects at about 1% (Supplementary Fig. 13). The detection of very weak hotspots, close to the sensitivity threshold of the method (10 tags in each orientation) is likely sensitive to experimental conditions. Therefore, though these hotspots frequently do not overlap those in other strains, the overlap estimates for these hotspots are not robust and are likely an underestimate. When counting overlaps between DSBs and H3K4me3 marks, the 1 Kb central region of both hotspots and H3K4me3 marks was used as we do not have an explicit model or estimate of the size of the peak centre for this histone modification. For overlap purposes, transcription start sites were defined as the 5' end of Ensembl transcripts downloaded from the UCSC genome browser[35] ± 1 Kb.

Data pooling and replicates:

To ensure that samples of the same genetic background were consistent, we called peaks for each sample and compared the peaks found in replicate samples. Since the number of sequencing tags for each replicate was not identical, we checked for the presence of the strongest 9,000 peaks of the smaller sample in the set of peaks from the larger sample. This number was chosen as the fewest peaks detected in any replicate sample was 9,242. All replicate samples were highly similar with more than 99% of peaks shared by this criterion. Thus, data from samples prepared from the same genetic background were pooled as described earlier.

PRDM9 dominance in F1 reciprocal crosses:

We wished to determine if the dominance of the 13R allele of PRDM9 in the F1 mice was dependent on the direction of the parental cross. We therefore examined this phenomenon in reciprocal crosses. SSDS data from a 13R female X 9R male (13Rfx9Rm) cross were generated in [13] (GEO accession number: GSM851663). RAD51 rather than DMC1 antibodies were used to generate this dataset, however we have previously shown that analogous results are obtained using either antibody[12]. We identified 6,472 hotspots in this dataset using the method described above. The F1 mice used in the present study were derived from the cross between a 9R female

and a 13R male (9Rfx13Rm). We used the strongest 6,472 hotspots for comparison with the 13Rfx9Rm dataset. The dominance of the 13R allele of PRDM9 is apparent in both crosses (Supplementary Table 1).

	13R derived hotspots	**9R derived hotspots**
13R female X 9R male	87 % (5,600)	11 % (745)
9R female X 13R male	86 % (5,589)	12 % (841)

Supplementary Table 1: The 13R allele of PRDM9 determines most hotspot locations in F1 mice of 13R X 9R reciprocal crosses.

Motif identification:

To define DSB-associated motifs in our datasets, we used MEME-ChIP[38]. We first indentified motifs using different combinations of command line arguments (below) to assess the sensitivity of the method to different parameters. The ultimate motifs varied little despite these permutations.

Sequences used: One argument passed to MEME-ChIP is a file of sequences. MEME-ChIP then randomly chooses 600 of these for motif finding. We wished to force the algorithm to use only stronger hotspots, therefore a file containing only the sequences of the 2,000 hotspots was used. Alternatively, the full set of hotspot sequences was used.

Required motif length: Default parameters OR length between 28 and 34 residues

Number of reported sequence motifs: N = 3,5 or 10.

The motifs reported in the main text were identified using the repeat-masked sequences of the top 2,000 DSB hotspots for each genetic background. The top 5 motifs were identified using the default parameters. For each motif, we used MAST with default parameters to identify significant hits around hotspots (centre ± 2.5 Kb). Only the top ranked motif in each background was co-centred with hotspots while the other motifs exhibited a flat distribution across this 5 Kb region. The enrichment at hotspot centres was calculated relative to the mean motif incidence in the flanking regions (± 2,501-5,001 nt). The mean motif incidence in these flanking regions is

comparable to the genomic average outside of hotspots. The motif enrichment in Fig. 1f, Supplementary Fig. 2 and Supplementary Fig. 14b were each calculated in the same way.

One interesting result arose when we searched only for long motifs (28 to 34 nt). As expected, the most informative region for each long motif coincided with the reported short motifs (Supplementary Fig. 14a). In addition, the highest information content was to the left of the long motif in 9R, while in 13R it was to the right. This again validates the alignments to the putative PRDM9 allele binding sites shown in Fig. 1g. We also compared the enrichment of the longer motifs around hotspots to that of the shorter versions (Supplementary Fig. 14b). The long version of the 13R motif actually contains less information than the shorter motif (Long: 17.1 bits; Short: 17.3 bits), however the long version of the 9R motif contains more information than the short version (Long: 16.0 bits; Short: 13.8 bits). This is reflected in the enrichment profiles around hotspots. The longer 13R motif does not increase specificity compared to the short version, however, the longer motif identified in 9R does increase the specificity of weaker matches to the motif.

PRDM9 binding site prediction:

We derived the consensus DNA binding sequence for each PRDM9 allele using the method defined in [41]. The first Zn-finger was excluded from predictions as it does not appear to be part of the tandem Zn-finger array. Linear SVM models were used for prediction.

Motifs were aligned to the predicted PRDM9 binding sites using the STAMP[42] web server with the following settings:

Column comparison metric: Pearson Correlation Coefficient

Alignment method: Ungapped Smith-Waterman

Multiple alignment strategy: Iterative Refinement

Tree-building algorithm: UPGMA

The 13R and 9R motifs aligned to their respective PRDM9 predicted binding sites with E-values of $4x10^{-4}$ and $2x10^{-4}$, respectively.

Histone lysine crotonylation analyses:

Histone lysine crotonylation (Kcr) is a recently discovered histone modification which has been

shown to preferentially associate with enhancers and promoters[19]. We downloaded the histone Kcr ChIP-Seq (GEO accession number: GSM810675) and Input DNA datasets (GEO accession number: GSM810673) for mouse spermatocytes from the GEO. We randomly subset the histone Kcr dataset to match the number of tags in the input, then used MACS[30] 1.3.7 (P-value = 0.0001) to identify sites of histone Kcr. 69,075 histone Kcr peaks were identified.

Analysis of dog recombination hotspots:

Dog recombination hotspots were downloaded from the supplementary information of [27]. The most precisely defined dog hotspots were 18 Kb in size. Of the 4,224 dog hotspots, 1,627 were defined to this resolution. We used this subset for subsequent analyses. We also generated matching hotspot sets from C57Bl/6 and *Prdm9*$^{-/-}$ by expanding these mouse hotspots to 18 Kb (centre ± 9 Kb).

The overlap between each set of hotspots and promoter regions (TSS ± 1 Kb) was then assessed. In mouse, 56% of *Prdm9*$^{-/-}{}_{18Kb}$ hotspots and 23% of C57Bl/6$_{18Kb}$ hotspots overlapped a promoter. 22% of dog$_{18Kb}$-hotspots overlapped promoters. This is similar to the proportion in C57Bl/6 ($P = 0.7$, two-sided binomial test) but significantly less than observed in *Prdm9*$^{-/-}$ mice ($P = 10^{-173}$, one-sided binomial test). To ensure that we did not greatly bias this analysis by selection of a subset of weak hotspots, we also checked the overlap for all 4,224 dog hotspots. This increased the overlap to 34%, still significantly less than that observed in *Prdm9*$^{-/-}$ mice ($P < 10^{-250}$, one-sided binomial test).

There are several caveats which must be considered here. First, the set of dog hotspots is far from complete. Second, there are 25% fewer annotated ensembl transcripts for dog than there are for mouse. Third, dog hotspots are not centred, therefore the 18 Kb region around hotspots is not directly analogous to the 18 Kb region which constitutes dog hotspots. Still, we think it unlikely that these factors will change the conclusion that dog hotspots do not exhibit the same distribution of hotspots at promoters as that observed in *Prdm9*$^{-/-}$ mice.

Supplementary References:

31. Hasenkrug, K.J. & Stimpfling, J.H. The antigenic properties of twenty-four new H-2 congenic recombinant mouse strains. *Immunogenetics* **24**, 423-7 (1986).
32. Yang, H. *et al.* A customized and versatile high-density genotyping array for the mouse. *Nat Methods* **6**, 663-6 (2009).
33. Shima, J.E., McLean, D.J., McCarrey, J.R. & Griswold, M.D. The murine testicular transcriptome: characterizing gene expression in the testis during the progression of spermatogenesis. *Biol Reprod* **71**, 319-30 (2004).
34. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* **15**, 1034-50 (2005).
35. Fujita, P.A. *et al.* The UCSC Genome Browser database: update 2011. *Nucleic Acids Res* **39**, D876-82 (2011).
36. Bellani, M.A., Romanienko, P.J., Cairatti, D.A. & Camerini-Otero, R.D. SPO11 is required for sex-body formation, and Spo11 heterozygosity rescues the prophase arrest of Atm-/- spermatocytes. *J Cell Sci* **118**, 3233-45 (2005).
37. Peters, A.H., Plug, A.W., van Vugt, M.J. & de Boer, P. A drying-down technique for the spreading of mammalian meiocytes from the male and female germline. *Chromosome Res* **5**, 66-8 (1997).
38. Machanick, P. & Bailey, T.L. MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics* **27**, 1696-7 (2011).
39. Romanienko, P.J. & Camerini-Otero, R.D. The mouse Spo11 gene is required for meiotic chromosome synapsis. *Mol Cell* **6**, 975-87 (2000).
40. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-60 (2009).
41. Persikov, A.V., Osada, R. & Singh, M. Predicting DNA recognition by Cys2His2 zinc finger proteins. *Bioinformatics* **25**, 22-9 (2009).
42. Mahony, S. & Benos, P.V. STAMP: a web tool for exploring DNA-binding motif similarities. *Nucleic Acids Res* **35**, W253-8 (2007).