The histone modification reader ZCWPW1 links histone methylation to PRDM9-induced double strand break repair

Tao Huang[1,2,3,4,*], Shenli Yuan[6,8,*], Lei Gao[6,8], Mengjing Li[1,2,3,4], Xiaochen Yu[1,2,3,4], Jianhong Zhang[6,8], Yingying Yin[1,2,3,4], Chao Liu[9], Chuanxin Zhang[1,2,3,4], Gang Lu[10], Wei Li[9], Jiang Liu[6,7,#], Zi-Jiang Chen[1,2,3,4,5,#], Hongbin Liu[1,2,3,4,10,#]

[1]Center for Reproductive Medicine, Cheeloo College of Medicine, Shandong University, Jinan, Shandong, 250012, China

[2]National Research Center for Assisted Reproductive Technology and Reproductive Genetics, Shandong University, Jinan, Shandong, 250012, China

[3]Key laboratory of Reproductive Endocrinology of Ministry of Education, Shandong University, Jinan, Shandong, 250012, China

[4]Shandong Provincial Clinical Medicine Research Center for Reproductive Health, Shandong University, Jinan, Shandong, 250012, China

[5]Shanghai Key Laboratory for Assisted Reproduction and Reproductive Genetics, Shanghai, 20000, China

[6]CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, 100101, China

[7]CAS Center for Excellence in Molecular Cell Science, University of Chinese Academy of Sciences, Beijing, 10010, China

[8]CAS Center for Excellence in Animal Evolution and Genetics, University of Chinese Academy of Sciences, Beijing, 100101, China

[9]State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing, 100101, China

[10]CUHK-SDU Joint Laboratory on Reproductive Genetics, School of Biomedical Sciences, the Chinese University of Hong Kong, Hong Kong, China

*These authors contributed equally to this work.

[#]Corresponding authors: liuj@big.ac.cn (J.L); chenzijiang@hotmail.com (Z-J.C); hongbin_sduivf@aliyun.com (HB.L)

ABSTRACT

The histone modification writer PRDM9 has been shown to deposit H3K4me3 and H3K36me3 at future double-strand break (DSB) sites during the very early stages of meiosis, but the reader of these marks remains unclear. Here, we demonstrate that ZCWPW1 is an H3K4me3 reader that is required for DSB repair and synapsis in mouse testes. We generated H3K4me3 reader-dead ZCWPW1 mutant mice and found that their spermatocytes were arrested at the pachytene-like stage, which phenocopies the *Zcwpw1* knock–out mice. Based on various ChIP-seq and immunofluorescence analyses using several mutants, we found that ZCWPW1's occupancy on chromatin is strongly promoted by the histone-modification activity of PRDM9. ZCWPW1 localizes to DMC1-labelled hotspots in a largely PRDM9-dependent manner, where it facilitates completion of synapsis by mediating the DSB repair process. In sum, our study demonstrates the function of ZCWPW1 that acts as part of the selection system for epigenetics-based recombination hotspots in mammals.

INTRODUCTION

Meiotic recombination ensures the faithful transmission of the genome through the pairing and segregation of homologous chromosomes, and it increases genetic diversity by disrupting linkage relationships (Bolcun-Filas and Schimenti, 2012; Handel and Schimenti, 2010). At the molecular level, meiotic recombination is initiated by the induction of programmed DSBs that are repaired by homologous recombination, leading to gene conversion and cross over formation (Gray and Cohen,

2016; Hunter, 2015; Zickler and Kleckner, 2015). DSB induction is a complex process, and DSB locations are known to be marked at the very earliest stages of meiosis by trimethylation of histone H3 on lysine 4 (H3K4me3) (Baudat et al., 2013; de Massy, 2013). In mammals, this is performed by the protein PRDM9, which is expressed in the leptotene and zygotene substages(Parvanov et al., 2017; Sun et al., 2015). PRDM9 is a DNA-binding zinc finger protein, with an exceptionally long and genetically variable zinc finger domain that determines its binding specificity (for defining recombination hotspots), while its SET domain possesses histone trimethyl transferase activity, and its KRAB domain is involved in protein-protein interactions(Grey et al., 2018; Paigen and Petkov, 2018). In yeast, the histone reader Spp1 links H3K4me3 sites at promoters with the DSB formation machinery, thus promoting DSB formation (Acquaviva et al., 2013; Sommermeyer et al., 2013). In mice, although multiple studies have shown that the H3K4me3 writer PRDM9 controls the locations of DSB formation (Baudat et al., 2010; Brick et al., 2012; Diagouraga et al., 2018; Grey et al., 2017; Myers et al., 2010; Parvanov et al., 2010; Powers et al., 2016), much less is known about the subsequent activities of any proteins that might read these epigenetic marks and thus participate in advancing the meiotic recombination process (Paigen and Petkov, 2018).

DSB formation at sites defined by PRDM9 is catalyzed by an evolutionarily conserved topoisomerase-like enzyme complex consisting of the SPO11 enzyme and its binding partner TOPOVIBL (Bergerat et al., 1997; Keeney et al., 1997; Panizza et al., 2011; Robert et al., 2016; Vrielynck et al., 2016). SPO11-mediated cleavage results in single-strand DNA overhangs that are subsequently coated by various proteins, including DMC1 and RAD51 (Dai et al., 2017; Pittman et al., 1998; Tarsounas et al., 1999). The DSBs enable homology searching and alignment to occur, which in turn promote homology synapsis and DSB repair (Inagaki et al., 2010). A basic feature of meiosis is that DSB-mediated interactions and repair processes occur differentially between homologous nonsister chromatids, rather than between sisters, as occurs in mitotic DSB repair (Garcia et al., 2015; Keeney et al., 2014; Lange et al., 2011). Some DSBs are repaired in a way that generates crossovers, wherein DNA is exchanged between homologous chromosomes (Baudat and de Massy, 2007). The ZMM proteins (*e.g.*, TEX11, MSH4/MSH5, and RNF212) are a group of functionally

related proteins known for their roles in promoting the formation of crossovers (Edelmann et al., 1999; Kneitz et al., 2000; Lynn et al., 2007; Reynolds et al., 2013; Yang et al., 2008).

We previously reported that the zinc finger CW-type and PWWP domain containing 1 (ZCWPW1) protein is required for meiosis prophase I in mice, and we found that *Zcwpw1* deficiency disrupted spermatogenesis in male mice but did not disrupt oogenesis in females to the same extent (Li et al., 2019a). ZCWPW1 is a member of the CW-domain containing protein family (Liu et al., 2016; Perry and Zhao, 2003), and its zinc finger CW (zf-CW) domain has three conserved tryptophan and four conserved cysteine residues. Structural analysis has shown that human ZCWPW1 zf-CW domain is a histone modification reader (He et al., 2010), while chromatin pulldown analysis has confirmed that ZCWPW1 zf-CW domain recognizes H3K4me3 marks (Hoppmann et al., 2011). A crystal structure of the human zf-CW domain of ZCWPW1 in complex with a peptide bearing an H3K4me3 mark revealed that four amino acids– W256, E301, T302, and W303 – are primarily responsible for the binding of ZCWPW1 zf-CW domain to H3K4me3 marks (He et al., 2010). However, whether the H3K4me3 reading function is required for ZCWPW1's physiological role in meiosis is still unknown.

To address the physiological role of ZCWPW1's H3K4me3 reading function, we generated an H3K4me3 reader-dead *Zcwpw1* knock-in mutant mouse line (*Zcwpw1*$^{KI/KI}$ mouse). We found that spermatocytes were arrested at the pachytene-like stage, which phenocopied the defect seen in *Zcwpw1* knock–out mice thus suggesting that H3K4me3 reader function of ZCWPW1 might facilitate meiotic recombination by facilitating the DSB repair process. Mechanistically, a series of chromatin immunoprecipitation sequencing (ChIP-seq) analyses of ZCWPW1, H3K4me3, and H3K36me3 in multiple knock–out and knock–in mouse lines established that ZCWPW1 is an H3K4me3 and H3K36me3 reader that exclusively binds at genomic loci bearing PRDM9-deposited histone modifications. ZCWPW1 localizes to DMC1-labelled DSB hotspots where it can read H3K4me3 and H3K36me3 marks. Thus, beyond demonstrating that the histone modification reader protein ZCWPW1 functions in an epigenetics-based recombination hotspot selection system, this study

advances our understanding of the sequence of recruitment events that are required for crossover formation during meiosis.

RESULTS

The H3K4me3 reader function of ZCWPW1 is essential for meiotic recombination

Previously, we developed *Zcwpw1* knockout mice in the C57BL/6 genetic background (Li et al., 2019a) and found that loss of *Zcwpw1* in male mice caused a complete failure of synapsis. This failure resulted in meiotic arrest at the zygotene to pachytene stage, and this was accompanied by incomplete DSB repair and lack of crossover formation, thus leading to male infertility. In light of the known capacity of ZCWPW1 to recognize epigenetic methylation modification marks, we designed a knock-in strategy to generate a H3K4me3 reader-dead ZCWPW1mutant mouse line (Figure 1–figure supplement 1A). Specifically, this knock-in mutant of ZCWPW1 had three mutations–W247I/E292R/W294P– and these mutations in mouse ZCWPW1 are equivalent to the previously reported W256I, E301R, and W303P mutations in the human ZCWPW1 protein (except for T302L in humans and S293 in mice, which are not conserved) (Figure 1–figure supplement 1B), and all of them are known to be essential for the H3K4me3 reader function of human ZCWPW1(He et al., 2010).

Western blot analysis confirmed the absence of the ZCWPW1 protein in *Zcwpw1*$^{-/-}$ testes, while the ZCWPW1$^{W247I/E292R/W294P}$ variant protein was expressed at a level similar to that of the wild type (WT) protein (Figure 1–figure supplement 2A and B). Consistent with the western blot data, immunofluorescence staining of frozen sections from 8-week-old WT, *Zcwpw1*$^{-/-}$, and *Zcwpw1*$^{KI/KI}$ mouse testes revealed that the ZCWPW1 protein was undetectable in *Zcwpw1*$^{-/-}$ spermatocytes but could still be found in ZCWPW1$^{W247I/E292R/W294P}$ mutant spermatocytes (Figure 1–figure supplement 2C). After confirming that the ZCWPW1$^{W247I/E292R/W294P}$ mutant protein could be expressed normally in *Zcwpw1*$^{KI/KI}$ mice, we prepared testis sections from 8-week-old WT, *Zcwpw1*$^{-/-}$, and the new *Zcwpw1*$^{KI/KI}$ mouse line. Hematoxylin staining showed that spermatogenesis was disrupted in both the *Zcwpw1*$^{-/-}$ and *Zcwpw1*$^{KI/KI}$ mice. Compared with the WT mice, the seminiferous tubules of the *Zcwpw1*$^{-/-}$ and *Zcwpw1*$^{KI/KI}$ mice lacked post-meiotic cell types, contained apoptotic cells, or were

nearly empty. Furthermore, the WT epididymides were full of sperm, but there were no obvious sperm detected in either the *Zcwpw1*$^{-/-}$ or *Zcwpw1*$^{KI/KI}$ samples, suggesting meiotic arrest in these mice (Figure 1A).

We then analyzed chromosome spreads of spermatocytes from the testes of adult mice by immunostaining for the synaptonemal complex (SC) markers SYCP1 and SYCP3 (Figure 1B). Immunostaining of SYCP1 and SYCP3 showed no differences among any of the genotypes with regard to leptotene-to-zygotene progression, which appeared normal in all mice. We quantified the synapsed chromosome pairs in the nuclei of WT, *Zcwpw1*$^{-/-}$, and *Zcwpw1*$^{KI/KI}$ testes from 8-week-old mice. We observed 169 spermatocytes in WT testes, and 153 spermatocytes (90.5%) had all chromosome pairs fully synapsed, with only 16 spermatocytes (9.5%) exhibiting synapsis between 4 and 18 pairs of chromosomes. In contrast, among 164 spermatocytes in *Zcwpw1*$^{-/-}$ testes and 158 spermatocytes in *Zcwpw1*$^{KI/KI}$ testes none had complete synapsis, and we only detected around of 8 synapsed chromosome pairs in *Zcwpw1*$^{-/-}$ and *Zcwpw1*$^{KI/KI}$ spermatocytes (Figure 1–figure supplement 2D)(Li et al., 2019a). Thus, spermatocytes lacking the H3K4me3-reader activity of the ZCWPW1 protein have severely disrupted synapsis.

Having established that ZCWPW1 facilitates the completion of synapsis during meiosis prophase I in male mice, we observed that the ZCWPW1$^{W247I/E292R/W294P}$ mutant mice exhibited the same synapsis defect as *Zcwpw1* knockout mice, suggesting that these residues are essential for the recombination-related functions of ZCWPW1. We then performed immunofluorescence staining of chromosome spreads to evaluate the recruitment of DMC1 and RAD51 to single-stranded overhang sequences in WT and *Zcwpw1*$^{KI/KI}$ mice (Figure 1C and E). There were no differences in the numbers of DMC1 or RAD51 foci in the leptotene or zygotene stages of the two genotypes. However, analysis of WT pachytene and *Zcwpw1*$^{KI/KI}$ pachytene-like spermatocytes revealed an obvious discrepancy. Decreased numbers of DMC1 and RAD51 foci were seen in the pachytene WT spermatocytes, indicating successful repair of DSBs, but the *Zcwpw1*$^{KI/KI}$ pachytene-like spermatocytes retained a large number of DMC1 and RAD51 foci (Figure 1D and F). These results suggest that the repair of DSBs is disrupted in the absence of a functional ZCWPW1 H3K4me3 reader protein and that ZCWPW1 might facilitate meiotic DSB repair downstream of strand invasion.

Seeking to further assess the functional contributions of ZCWPW1 in meiotic recombination, we analyzed chromosome spreads of spermatocytes from the testes of adult WT and *Zcwpw1*$^{KI/KI}$ mice by immunostaining for the recombination factors MSH4 and RNF212 and the Holliday junction resolution marker MLH1 (Figure 1–figure supplement 3A and C and E). Staining for MSH4 and RNF212 showed that the recombination machinery could assemble normally in both WT and *Zcwpw1*$^{KI/KI}$ spermatocytes at the zygotene stage. However, these MSH4 and RNF212 signals decreased as expected in WT pachytene spermatocytes, but persisted on the pachytene-like *Zcwpw1*$^{KI/KI}$ chromosomes (Figure 1–figure supplement 3B and D). Additionally, the MLH1 staining patterns indicated that Holliday junction resolution proceeded normally in mid- to late-pachytene WT spermatocytes but indicated that the recombination process was arrested in the pachytene-like spermatocytes lacking ZCWPW1 H3K4me3-reader function, which failed to progress to the pachytene stage and for which no crossover occurred, thus resulting in the absence of MLH1 foci (Figure 1–figure supplement 3E). These results suggest that both DSB repair and recombination are defective in the *Zcwpw1*$^{KI/KI}$ mice.

To determine the specific process that can mechanistically account for the observed failure to complete meiotic recombination, we stained the spreads of spermatocytes from the testes of adult WT, *Zcwpw1*$^{-/-}$, and *Zcwpw1*$^{KI/KI}$ mice for the DSB marker γH2AX. We found that DSBs could form normally in all of the genotypes (Figure 2A), but there were obvious differences between pachytene WT spermatocytes and pachytene-like *Zcwpw1*$^{-/-}$ and *Zcwpw1*$^{KI/KI}$ spermatocytes. The WT pachytene spermatocytes exhibited no obvious signal for γH2AX on autosomes, but retained a γH2AX signal on the sex chromosomes, which corresponds to the XY sex body and indicates silencing of the sex chromosomes. In contrast, both autosomes and sex chromosomes retained obvious γH2AX signals, and no XY bodies were observed in *Zcwpw1*$^{-/-}$ or *Zcwpw1*$^{KI/KI}$ pachytene-like spermatocytes. We next stained against the DSB-repair machinery component p-ATM and found that both autosomes and sex chromosomes retained obvious p-ATM signals in the pachytene-like *Zcwpw1*$^{-/-}$ and *Zcwpw1*$^{KI/KI}$ spermatocytes (Figure 2B). Through single−stranded DNA sequencing (SSDS) by ChIP-seq against DMC1, Wells *et al.* found that DSBs occur in the same hotspot regions in *Zcwpw1*$^{-/-}$ male mice (Mahgoub et al., 2019; Wells et al., 2019). Similarly, using quantitative END−seq, Mahgoub *et al.*

189 also confirmed that DSBs in both WT and *Zcwpw1*$^{-/-}$ mice completely overlapped with each other
190 and with previously identified hotpots (Mahgoub et al., 2019; Wells et al., 2019). These results
191 indicate that ZCWPW1 is dispensable for the induction and location of DSBs but is required for
192 proper interhomolog interactions, including synapsis and the repair of DSBs that occur during the
193 later steps of homologous recombination.

194

195 **ZCWPW1 is an H3K4me3/H3K36me3 reader**

196 Having thus established that ZCWPW1 promotes the completion of synapsis and that it functions
197 in meiotic recombination by facilitating DSB repair, we next investigated the mechanism by which
198 ZCWPW1 recognizes histone modification marks involved in male meiosis prophase I. To this end,
199 we conducted ChIP-seq using antibodies against the ZCWPW1 protein and against H3K4me3 marks.
200 The ZCWPW1 ChIP-seq data for C57BL/6 mouse testis chromatin revealed a total of 14,688
201 ZCWPW1 peaks, with 499 peaks localized within 2,000 bp upstream of a transcription start site
202 (TSS), 2,416 peaks localized in exons, 6,142 peaks localized in introns, and 5,873 peaks localized
203 within intergenic regions (Figure 3–figure supplement 1A).

204 Among all ZCWPW1 binding sites detected in mouse testes, 11.5% of the ZCWPW1 binding sites
205 (peaks) overlapped with promoters, while 6.1% of ZCWPW1 binding sites overlapped with CpG
206 islands (Figure 3–figure supplement 1B). Compared with the random binding sites, ZCWPW1
207 binding sites were not significantly enriched in the transposable element regions (Figure 3–figure
208 supplement 1C D). In HEK293T cells, Wells *et al*. found that a large proportion of the weakly-
209 binding ZCWPW1 sites overlapped with Alu repeats. Notably, the weakest ZCWPW1 peaks
210 overlapped most frequently with Alus repeats, while the strongest peaks were depleted of Alus repeats
211 relative to random overlap. They also found that ZCWPW1 appears to have a greater affinity for
212 methylated CpG pairs but retains some affinity even for non-methylated regions (Mahgoub et al.,
213 2019; Wells et al., 2019). Because we found that 1,766 of the ZCWPW1 binding sites overlapped with
214 promoters, we sought to examine the transcriptome in *Zcwpw1*$^{-/-}$ testes by RNA-seq to investigate
215 whether ZCWPW1 affected the expression of those genes whose promoters overlapped with

ZCWPW1 binding sites. Analysis of RNA-seq data of postnatal day 14 (PD14) WT and *Zcwpw1*$^{-/-}$ mice identified 567 differentially expressed genes (DEGs), including 464 downregulated and 103 upregulated DEGs in *Zcwpw1*$^{-/-}$ testes compared with WT testes (Figure 3–figure supplement 2A). Gene ontology analysis showed that the down-regulated genes were enriched in axoneme assembly, male gamete generation and flagellated sperm motility (Figure 3–figure supplement 2B). However, most of the DEGs were not the genes whose promoters overlapped with ZCWPW1 binding sites (Figure 3–figure supplement 2C). These results strongly suggest that ZCWPW1 may not affect the transcription level of many genes sharing promoter overlap with ZCWPW1, even though *Zcwpw1*$^{-/-}$ (vs WT) spermatocytes have 2 up-regulated genes and 45 down-regulated genes with such promoter overlap (Supplemental File 3).

The H3K4me3 ChIP-seq data in C57BL/6 mice revealed a total of 55,801 H3K4me3 peaks, consistent with a previous report of 55,497 H3K4me3 peaks in whole testes(Smagulova et al., 2011). Lam *et al.* described a method for isolating pure sub-populations of meiotic substage nuclei, and they detected a total of 75,771 H3K4me3 peaks among isolated SCP3^{+}H1T^{-} spermatocytes; this signature defines pre-leptotene to early-pachytene substage spermatocytes(Lam et al., 2019). In our work, we obtained a weaker average H3K4me3 signal in ZCWPW1 peaks in whole testes than that reported by Lam *et al.* in isolated, stage-specific spermatocyte nuclei (Figure 3C). The ChIP-seq data from sorted meiotic cells thus allowed the elimination of H3K4me3 peaks originating from cells that did not express ZCWPW1. In light of the known capacity of ZCWPW1 to recognize epigenetic methylation modification marks, we compared the ZCWPW1 peaks with these two sets of H3K4me3 peaks, and we found that 97.8% (14,369 of 14,688 peaks) of the ZCWPW1 peaks overlapped with the H3K4me3 peaks reported by Lam *et al*, while 39.4% (5,792 of 14,688 peaks) of the ZCWPW1 peaks overlapped with the H3K4me3 peaks in our data (Figure 3A and B), therefore supporting the hypothesis that this specific overlap with H3K4me3 peaks serves as a means for ZCWPW1 recognition of histone modification marks.

To determine whether the H3K4me3 binding ability of ZCWPW1's CW-domain is necessary for its recruitment to chromatin *in vivo*, we conducted an additional ZCWPW1 ChIP-seq analysis of testes samples from PD14 WT, *Zcwpw1*$^{-/-}$, and *Zcwpw1*$^{KI/KI}$ mice. The analysis indicated that no

ZCWPW1 peaks were detected in the *Zcwpw1*$^{-/-}$ or *Zcwpw1*$^{KI/KI}$ mice (Figure 3A and D). These *in vivo* results, viewed alongside the previous reports of ZCWPW1 function in the meiotic process demonstrating that these specific mutations in the ZCWPW1 zf-CW domain affect the protein's ability to read histone modifications (including H3K4me3), together indicate that the ZCWPW1$^{W247I/E292R/W294P}$ mutant is an H3K4me3 reader-dead variant of ZCWPW1. Furthermore, these results suggest that the H3K4me3 reader function of this protein is essential for its ability to bind to chromatin and to function in meiosis prophase I in male mice.

ZCWPW1 also has a PWWP domain which was found in multiple other proteins to specifically bind to histone H3 containing an H3K36me3 mark (Qin and Min, 2014), so we next sought to better understand the overlap between ZCWPW1 peaks and H3K36me3 peaks in our ChIP-seq dataset. We found that 90.1% of the ZCWPW1 peaks overlapped with the H3K36me3 peaks reported by Lam *et al.* in isolated stage-specific spermatocyte nuclei, while 24.8% of the ZCWPW1 peaks overlapped with the H3K36me3 peaks identified by Grey *et al*. in whole testes (Figure 3–figure supplement 3A). As with our data from whole testes, the average H3K36me3 signal of ZCWPW1 peaks obtained by Grey *et al.* was considerably weaker than that from isolated stage-specific spermatocyte nuclei obtained by Lam *et al*. (Figure 3–figure supplement 3B).

In analyzing the correlation between ZCWPW1 binding sites and these two histone modification marks, we found that 88.8% of the ZCWPW1 peaks overlapped with regions containing both H3K4me3 and H3K36me3 marks, while only 9.1% and 1.3% of ZCWPW1 peaks overlapped with H3K4me3 and H3K36me3 peaks individually (Figure 3E). Furthermore, the ZCWPW1 peak intensity was significantly higher for the dual overlapping regions than for regions containing either H3K4me3 or H3K36me3 alone (Figure 3–figure supplement 3C). We also found that ZCWPW1 bound H3K4me3 regions had higher H3K36me3 levels than H3K4me3 regions were not bound by ZCWPW1 (Figure 3–figure supplement 3D). We conducted immunofluorescence analysis of chromosome spreads of spermatocytes from adult mice using rat anti-ZCWPW1 and rabbit anti-H3K4me3/H3K36me3 antibodies and found, consistent with the ChIP-seq data, that the pattern for both H3K4me3 and H3K36me3 signals were similar to the ZCWPW1 expression pattern we observed in the leptotene and zygotene stages (Figure 3–figure supplement 4A-B). Moreover, Mahgoub *et al.*

(2019) confirmed that recombinant ZCWPW1 (1–440aa) binds with the highest affinity to H3K4me3/K36me3 peptides *in vitro*. Taken together, these results demonstrate that ZCWPW1 preferentially binds to sites with both H3K4me3 and H3K36me3 marks.

ZCWPW1 binding is strongly promoted by the histone modification activity of PRDM9

To identify the factors responsible for ZCWPW1 recruitment to chromatin *in vivo*, we searched for enriched motifs within the ZCWPW1 binding sites in our ChIP-seq data (Figure 4–figure supplement 1A). This analysis identified a *de novo* motif that is highly correlated with a known PRDM9 binding motif in mice (Figure 4-figure supplement 1B) (Billings et al., 2013; Segurel, 2013; Walker et al., 2015), and this suggested that ZCWPW1 binding to chromatin might occur in a PRDM9-dependent manner. To pursue this possibility, we compared our ZCWPW1 ChIP-seq data with previously published *in vivo* ChIP-seq data generated using an anti-PRDM9 antibody and with *in vitro* data from an affinity-seq analysis of genome-wide PRDM9 binding sites (Grey et al., 2017; Walker et al., 2015). At the genome-wide level, 13% of the ZCWPW1 peaks obtained in our study overlapped with Grey *et al.*'s *in vivo* PRDM9 peaks, while 74% of the ZCWPW1 peaks overlapped with Walker *et al.*'s *in vitro* PRDM9 peaks. Moreover, it should be noted that 99.5% of the ZCWPW1 peaks overlapped with PRDM9 in Grey *et al* are covered by that ZCWPW1 peaks overlapped with PRDM9 peaks in Walker *et al* (Figure 4–figure supplement 3B). Conversely, we found that 1,934 of 2,601 PRDM9 peaks (74%) from Grey *et al.* and 10,975 of 36,898 PRDM9 peaks (29.7%) from Walker *et al.* overlapped with our ZCWPW1 peaks (Figure 4A and B). The high overlap between ZCWPW1 and PRDM9 peaks further suggested that ZCWPW1 occupancy occurs in a PRDM9-dependent manner.

To further explore this finding of high overlap between ZCWPW1 and PRDM9 peaks in our ChIP-seq data, and in light of the well-known overlap of PRDM9 peaks with H3K4me3 and H3K36me3 marks (Grey et al., 2017), we compared the ZCWPW1/PRDM9 overlap with the ZCWPW1/histone mark overlap. We found that the majority of ZCWPW1 peaks overlapped with PRDM9 binding sites containing both H3K4me3 and H3K36me3 marks (Figure 4A and C, Figure 4–figure supplement 1C and F). Our further analysis of H3K4me3 peak intensity in whole testes showed that among the PRDM9-occupied regions from Grey *et al.*, the intensity of H3K4me3 peaks

overlapping with ZCWPW1 was significantly weaker than that of ZCWPW1-non-overlapping regions (Figure 4–figure supplement 1D left panel), which was consistent with previous reports (Smagulova et al., 2011). In contrast, the H3K4me3 and H3K36me3 peak intensities of isolated stage-specific spermatocyte nuclei (Lam *et al.*) showed that among the PRDM9-occupied regions from Grey *et al.* and Walker *et al.*, the intensities of H3K4me3 and H3K36me3 peaks overlapping with ZCWPW1 were significantly greater than the intensities of ZCWPW1-non-overlapping regions (Figure 4–figure supplement 1D and E). Allowing for differences in the binding performance of different antibodies in different ChIP-seq analyses, the fact that some but certainly not all of the ZCWPW1 peaks overlapped with PRDM9 peaks suggests that it is the H3K4me3 and perhaps H3K36me3 epigenetic marks deposited by PRDM9, rather than the PRDM9 protein *per se*, that can explain the observed overlap of the ZCWPW1 and PRDM9 peaks.

To determine whether the activity of PRDM9 is necessary for ZCWPW1 recruitment to chromatin *in vivo*, we conducted an additional ZCWPW1 and H3K4me3 ChIP-seq analysis of testes samples from PD14 WT and *Prdm9*$^{-/-}$ mice (Figure 1–figure supplement 1C). Consistent with a previous report (Brick et al., 2012), the majority of PRDM9-dependent H3K4me3 peaks disappeared in *Prdm9*$^{-/-}$ mice (Figure 4D and E, Figure 4–figure supplement 2A-C). In our ChIP-seq data, we found that the H3K4me3 peaks overlapped with PRDM9 binding sites (Grey et al., 2017); with the notable exception of a 94.7% loss in ZCWPW1 binding sites, we found no obvious discrepancies between WT and *Prdm9*$^{-/-}$ testes (Figure 4E right panel, Figure 4–figure supplement 2A). This suggested that ZCWPW1 binding is strongly promoted by the specific activity of PRDM9 (Figure 4D and E, Figure 4–figure supplement 2A-C).

Having established that ZCWPW1 binding to chromatin is strongly promoted by the histone modification activity of PRDM9, we next examined changes in ZCWPW1 binding sites between WT and *Prdm9*$^{-/-}$ mutant testes. We found that although 94.7% of the ZCWPW1 peaks were apparently lost in *Prdm9*$^{-/-}$ mutant testes, 781 ZCWPW1 peaks were maintained and were accompanied by 652 newly generated ZCWPW1 peaks in *Prdm9*$^{-/-}$ mice (Figure 4–figure supplement 3A). Furthermore, examination of peak intensities showed that the new ZCWPW1 peaks were significantly weaker than those of both the maintained and the lost ZCWPW1 peaks in *Prdm9*$^{-/-}$mice (Figure 4–figure

supplement 3C). The majority of these gained (67.8%) and maintained (83.2%) ZCWPW1 peaks overlapped with promoter regions, while only 7.4% of the lost ZCWPW1 peaks overlapped with promoter regions in *Prdm9*$^{-/-}$mice (Figure 4–figure supplement 3D). Further analysis showed that nearly 80% of the lost ZCWPW1 peaks overlapped with PRDM9 binding sites, while the majority of the maintained and gained ZCWPW1 peaks did not overlap with PRDM9 peaks (Figure 4–figure supplement 3E). Surprisingly, a motif analysis showed that 3,028 ZCWPW1 peaks, *i.e.,* those that were lost and did not overlap with PRDM9 binding sites, were significantly enriched at PRDM9 binding sites (Figure 4–figure supplement 3F), suggesting that ZCWPW1 binding to these sites is highly PRDM9 dependent.

ZCWPW1 localizes to DMC1-labelled DSB hotspots in a PRDM9-dependent manner

A previous study developed a SSDS analysis using an antibody against DMC1 in mouse testes—that specifically detects protein-bound single-stranded DNA at DSB ends (Khil et al., 2012). SSDS provides insights into the shape and evolution of the mammalian DSB landscape (Davies et al., 2016). Lange *et al*. sequenced mouse SPO11 oligos and provided nucleotide-resolution DSB maps with low background and high dynamic range and found that SPO11 oligo counts correlated well with SSDS coverage(Lange et al., 2016). A previously study found that 94% of future DMC1 binding sites are enriched at H3K4me3 labelled hotspots, and asserted that such enrichment can be considered a global feature of DSB sites in multicellular organisms (Smagulova et al., 2011).Because we found that ZCWPW1 recognized dual histone modifications via PRDM9, we compared the distribution of the ZCWPW1 peaks with the DMC1 peaks and SPO11 oligos in the publicly available datasets(Grey et al., 2017; Lange et al., 2016). For the WT mice, 11,124 of the 14,688 total ZCWPW1 peaks overlapped with DMC1-defined DSB hotspots, while 10,340 of ZCWPW1 peaks overlapped with SPO11 oligo-defined DSB hotspots (Figure 5–figure supplement 1A). Both a heatmap and a scatter plot emphasized strong enrichment for ZCWPW1 signals at DMC1 binding sites (Figure 5–figure supplement 1B and C). Specifically, the greater the ZCWPW1 peak intensity, the better the overlap between the ZCWPW1 peaks and the DMC1 peaks (Figure 5–figure supplement 1D). These results strongly suggest that ZCWPW1 localizes to DMC1-labelled DSB hotspots.

Further analysis showed that 65.1% of the DMC1 peaks overlapped with both ZCWPW1 binding sites and merged PRDM9 peaks (Figure 5–figure supplement 1E). Our ZCWPW1 and H3K4me3 ChIP-seq data in WT and *Prdm9*$^{-/-}$ mice indicated an apparent lack of ZCWPW1 peaks and H3K4me3 signals at DMC1-labelled DSB hotspots (Figure 5A and B). These results reinforce the idea that occupancy of ZCWPW1 at DMC1-labelled DSB hotspots in *Prdm9*$^{-/-}$ spermatocytes is largely dependent on PRDM9-mediated histone modifications.

However, it bears mentioning that we also detected 781 ZCWPW1 peaks in WT testes that did not obviously overlap with DSB hotspots and we detected 652 ZCWPW1 peaks that did not obviously overlap with DSB hotspots in *Prdm9*$^{-/-}$ mice (Figure 5B, Figure 4–figure supplement 3A). We analyzed these 781 ZCWPW1 binding sites in detail and we found that 83.2% of these maintained peaks occurred within TSS ± 2kb, a substantially larger proportion than for the average position among all lost ZCWPW1 peaks (Figure 5–figure supplement 1F). We also found that the distribution pattern of H3K4me3 and H3K36me3 peaks, which overlapped with those 1,433 ZCWPW1 peaks, was significantly different compared to the distribution pattern of H3K4me3 and H3K36me3 peaks that overlapped with DMC1-labelled DSB hotspots (Figure 5B). Thus, although it is clear that the majority of the ZCWPW1 peaks resulted from PRDM9 activity, it is possible that ZCWPW1 might have an additional transcription regulation function that is not obviously related to the PRDM9-mediated hotspot selection system.

DISCUSSION

Our data support a working model wherein PRDM9 binds to specific DNA motifs in the genome and writes histone modifications (H3K4me3 and H3K36me3) via the methyltransferase activity of its PR/SET domain (Diagouraga et al., 2018; Powers et al., 2016). This leads to the recruitment of proteins required for the formation of DSBs in the vicinity of its binding site (e.g., SPO11, etc.) (Kumar et al., 2018; Panizza et al., 2011; Stanzione et al., 2016; Tesse et al., 2017). After these PRDM9-catalyzed epigenetic modifications are deposited, ZCWPW1 can specifically read these H3K4me3 and H3K36me3 marks in the vicinity of DSB sites, where ZCWPW1 functions to promote

DSB repair; ZCWPW1 is not however required for DSB activity and localization (Mahgoub et al., 2019; Wells et al., 2019). This DSB-repair-promoting function greatly increases the overall completion rates of synapsis, crossover formation, and ultimately meiotic progression (Figure 5C).

The identification of recombination hotspots was first made in genetically-tractable experimental organisms such as bacteriophages and fungi, but it is now apparent that such hotspots are ubiquitous and active in all organisms (Wahls, 1998). Higher-order chromosome architecture, which can be described using the terminology of the "tethered-loop/axis complex" model, contributes to DSB hotspot localization (Blat et al., 2002), and different strategies and mechanisms for the spatial regulation of DSB formation have evolved in different species, although these have many common features (Baudat et al., 2013; de Massy, 2013). In considering the evolution of hotspot selection systems, we are interested in whether other meiotic factors might have evolved in vertebrates to link PRDM9 to the meiotic recombination machinery and/or the synaptonemal complex, which would permit direct interactions with the histone marks deposited by PRDM9 (Paigen and Petkov, 2018; Tock and Henderson, 2018).

In *Saccharomyces cerevisiae*, Spp1–whose PHD finger domain is known to read H3K4me3 marks– promotes meiotic DSB formation by interacting with the axis-bound Spo11 accessory protein Mer2 (Acquaviva et al., 2013; Sommermeyer et al., 2013). Our study in mammals supports that one or more other as-yet unknown proteins might function in a similar role during DSB formation. It is noteworthy that there is structural similarity between the zf-CW domain and the PHD finger of Spp1 that helps recognize histone H3 tails (Adams-Cioaba and Min, 2009). Moreover, structural analysis has indicated that the zf-CW domain of human ZCWPW1 is a histone modification reader (He et al., 2010), and chromatin pulldown analysis has confirmed that this domain recognizes H3K4me3 marks (Hoppmann et al., 2011). In the present study, we showed that ZCWPW1 can specifically read H3K4me3 and H3K36me3 marks in the vicinity of DSB sites. However, somewhat surprisingly, our subsequent experiments indicated that deficiency of ZCWPW1 did not affect the recruitment of recombination-related factors like DMC1, MSH4, or RNF212, thereby implying that there might be other unknown proteins that function to link PRDM9 to meiotic recombination machinery.

The zf-CW domain of ZCWPW1 has previously been shown to bind to H3K4me3 peptides (He et al., 2010), and the PWWP domain, another type of "reader" module, has been shown to recognize H3K36me3 in both peptide and nucleosome contexts (Eidahl et al., 2013; Rondelet et al., 2016; Vezzoli et al., 2010). Consistent with a recently deposited pre-print at bioRxiv showing that ZCWPW1 can bind to histone H3 peptides with double H3K4me3 and H3K36me3 marks with high affinity at a 1:1 ratio *in vitro*(Mahgoub et al., 2019), we also found that ZCWPW1 localized to H3K4me3 and H3K36me3 enrichment regions in ChIP-seq analysis. Notably, most of the ZCWPW1 peaks overlapping H3K4me3 peaks disappeared in *Prdm9* knockout mice. One functional implication of our study is that it is PRDM9's histone modification activity, rather than the chromatin residence of the PRDM9 protein *per se*, that might account for the functional interactions of the apparently co-involved ZCWPW1 and PRDM9 proteins.

Our H3K4me3 reader-dead mutant mice showed that, upon loss of the binding affinity of the ZCWPW1 zf-CW domain for H3K4me3 marks, the ZCWPW1 protein completely lost its ability to bind chromatin, and spermatocytes in mice expressing this knock-in H3K4me3 reader-dead mutant ZCWPW1 exhibited a nearly complete failure of meiosis prophase I. We also hypothesize that this protein is unable to bind unmodified histones, that is, the mutant protein might be a histone binding-dead variant in addition to being a reader-dead variant. However, it remains unclear whether the ZCWPW1 PWWP domain, which likely functions in reading H3K36me3 marks, and/or other regions of the ZCWPW1 protein confer similarly important functions. Indeed, in future work we plan to pursue the selective disruption of the function of particular ZCWPW1 domains in our attempts to elucidate this protein's functions in male meiosis I.

While we clearly show that ZCWPW1 greatly facilitates PRDM9-dependent DSB repair, we do not yet have strong evidence for the precise nature of its functional role. One possibility is that ZCWPW1, upon binding to PRDM9-dependent histone modification hotspots, might serve as a DSB mark, which can perhaps subsequently recruit other factors involved in DSB repair. Recent studies have reported that PRDM9 binds on both the cut and uncut template chromosomes to promote meiotic recombination (Hinch et al., 2019; Li et al., 2019b). It is also possible that ZCWPW1 might directly

interact with the SC machinery by using its SCP1-like domain to tether PRDM9-bound loops to the SC in order to promote homologous DSB repair.

In summary, our study identifies ZCWPW1 as an H3K4me3 and H3K36me3 reader that promotes the repair of DNA DSBs during meiotic recombination. Complementary analyses and similar conclusions were obtained by Maghoub *et al* and by Wells *et al*, establishing that ZCWPW1 is not required for DSB activity and localization(Mahgoub et al., 2019; Wells et al., 2019), findings which do not support the previous speculation that ZCWPW1 directs the location or the formation of DSBs (Li et al., 2019a). In future studies, we plan to focus on additional proteins (*e.g.*, ZCWPW2, MORC3/4, etc.) that have similar functional domains as ZCWPW1 (Liu et al., 2016), with the aim of identifying any unknown biomolecules that might act to link PRDM9 to the DSB machinery specifically or to meiotic recombination more generally.

MATERIALS AND METHODS

Key Resources Table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
gene (Mus musculus)	*zinc finger, CW type with PWWP domain 1(Zcwpw1)*	Mouse Genome Informatics	MGI:2685899	
gene (Mus musculus)	*PR domain containing 9 (Prdm9)*	Mouse Genome Informatics	MGI:2384854	
Strain, strain background (Mus musculus, C57BL6/J)	Mouse: C57BL/6J	Jackson laboratory	RRID: IMSR_JAX:000664	All sexes used

genetic reagent (Mus musculus, C57BL6/J)	*Zcwpw1*$^{-/-}$	PMID: 31453335		Materials and methods section Generation of Zcwpw1 knockout mice
genetic reagent (Mus musculus, C57BL6/J)	*Zcwpw1*$^{KI/KI}$	This study		Materials and methods section Mice
genetic reagent (Mus musculus, C57BL6/J)	*Prdm9*$^{-/-}$	This study		Materials and methods section Mice
antibody	Rabbit polyclonal anti-ZCWPW1	PMID: 31453335		IF 1:1000; WB 1:5000; ChIP 35ug
antibody	Rabbit polyclonal anti-alpha Tubulin	Proteintech Group	11224-1-AP RRID: AB_2210206	WB 1 : 10000
antibody	Rat polyclonal anti-ZCWPW1	This study		Materials and methods section IF 1:200
antibody	Mouse polyclonal anti-SCP3	Abcam	#ab97672 RRID: AB_10678841	IF 1:500
antibody	Rabbit polyclonal anti-SCP1	Abcam	# ab15090 RRID: AB_301636	IF 1:2000
antibody	Rabbit polyclonal anti-RAD51	Thermo Fisher Scientific	#PA5-27195 RRID: AB_2544671	IF 1:200

antibody	Rabbit polyclonal anti-DMC1	Santa Cruz Biotechnology	#sc-22768 RRID: AB_2277191	IF 1:100
antibody	Mouse polyclonal anti-γH2AX	Millipore	#05-636 RRID: AB_309864	IF 1:300
antibody	Mouse polyclonal anti-pATM	Sigma-Aldrich	#05-740 RRID: AB_309954	IF 1:500
antibody	Rabbit polyclonal anti-MSH4	Abcam	#ab58666 RRID: AB_881394	IF 1:500
antibody	Rabbit polyclonal anti-RNF212	this study	a gift from Mengcheng Luo, Wuhan University	Materials and methods section IF 1:500
antibody	mouse polyclonal anti-MLH1	BD Biosciences	#550838 RRID: AB_2297859	IF 1:50
antibody	Rabbit polyclonal anti-H3K4me3	Abcam	#ab8580 RRID: AB_306649	IF 1:500 ChIP 5ug
antibody	Rabbit polyclonal anti-H3K36me3	Abcam	#ab9050 RRID: AB_30696	IF 1:500 ChIP 5ug
antibody	Goat anti-rabbit IgG (H+L) cross-adsorbed secondary Antibody, Alexa Fluor 488	Thermo Fisher Scientific	#A-11070 RRID: AB_142134	IF 1:500

antibody	Goat anti-rabbit IgG H&L (Alexa Fluor 594) preadsorbed	Abcam	ab150084 RRID: AB_2734147	IF 1:500
antibody	Goat anti-mouse IgG H&L (Alexa Fluor488)	Abcam	ab150113 RRID: AB_2576208	IF 1:500
antibody	Goat anti-mouse IgG H&L (Alexa Fluor 594) preadsorbed	Abcam	ab150120 RRID: AB_2631447	IF 1:500
antibody	Goat anti-rat IgG H&L (Alexa Fluor 488) preadsorbed	Abcam	ab150165 RRID: AB_2650997	IF 1:500
sequence-based reagent	Primers used in this study	This study		All primers used in this study are shown in the methods
commercial assay or kit	NEBNext Ultra II DNA Library Prep Kit	NEB	E7645S	
software, algorithm	Trimmomatic v0.32	doi:10.1093/bioinformatics/btu170	http://www.usadellab.org/cms/index.php?page=trimmomatic	
software, algorithm	Bowtie2 v2.3.4.2	doi:10.1038/nmeth.1923	RRID:SCR_005476	
software, algorithm	Samtools and Picard	doi:10.1038/ng.806	http://samtools.sourceforge.net	
software, algorithm	MACS2 v2.1.0	doi:10.1186/gb-2008-9-9-r137	RRID:SCR_013291	
software, algorithm	Integrative Genomics Viewer	doi:10.1038/nbt.1754	RRID:SCR_011793	

software, algorithm	Deeptools2	doi:10.1093/nar/gkw257	http://deeptools.ie-freiburg.mpg.de	
software, algorithm	HOMER software	doi: 10.1016/j.molcel.2010.05.004	RRID:SCR_010881	
software, algorithm	GREATER software	doi:10.1038/nbt.1630	http://great.stanford.edu	
software, algorithm	regioneR package version1.18.1	doi: 10.1093/bioinformatics/btv562	http://www.bioconductor.org/packages/regioneR	
software, algorithm	hisat2 v2.1.0	doi: 10.1038/nprot.2016.095	http://github.com/infphilo/hisat2	
software, algorithm	DESeq2 v 1.18.0	doi: 10.1186/s13059-014-0550-8	RRID:SCR_015687	
software, algorithm	Metacape	doi: 10.1038/s41467-019-09234-6	http://metascape.org	

Mice

The *Zcwpw1* gene (NCBI Reference Sequence: NM_001005426.2) is located on mouse chromosome 5 and comprises 17 exons, with its ATG start codon in exon 2 and a TAG stop codon in exon 17. The *Zcwpw1* knockout mice were generated in our previous study (Li et al., 2019a). The *Zcwpw1* knock-in H3K4me3-reader-dead mutant mice were generated by mutating three sites. The W247I (TGG to ATT) point mutation was introduced into exon 8 in the 5' homology arm, and the E292R (GAG to CGG) and W294P (TGG to CCG) point mutations were introduced into exon 9 in the 3' homology arm. The W247I, E292R, and W294P mutations created in the mouse *Zcwpw1* gene are positionally equivalent to the W256I, E301R, and W303P mutations previously reported in the human *ZCWPW1* gene. To engineer the targeting vector, homology arms were generated by PCR using BAC clones RP24-387B18 and RP24-344E7 from the C57BL/6 library as templates. In the targeting vector, the Neo cassette was flanked by SDA (self-deletion anchor) sites. DTA was used for negative selection.

C57BL/6 ES cells were used for gene targeting, and genotyping was performed by PCR amplification of genomic DNA extracted from mouse tails. PCR primers for the *Zcwpw1* Neo deletion were Forward: 5'-CAC TGA GTT AAT CCC ACC TAC GTC-3' and Reverse: 5'CTC TCC CAA ACC ATC TCA AAC ATT-3', with targeted point mutants yielding a 318 bp fragment and WT mice yielding a 174 bp fragment (Cyagen Biosciences Inc, Guangzhou, China).

The mouse *Prdm9* gene (GenBank accession number: NM_144809.3) is located on mouse chromosome 17. Ten exons have been identified, with the ATG start codon in exon 1 and the TAA stop codon in exon 10. The *Prdm9* knockout mice in the C57BL/6 genetic background were generated by deleting the genomic DNA fragment covering exon 1 to exon 9 using the CRISPR/Cas9-mediated genome editing system (Cyagen Biosciences Inc, Guangzhou, China). The founders were genotyped by PCR followed by DNA sequencing analysis. Genotyping was performed by PCR amplification of genomic DNA extracted from mouse tails. PCR primers for the *Prdm9* mutant allele were Forward: 5'-GCT TAG GTA GCA GAA TTG AAG GGA AAG TC-3' and Reverse: 5'- GTT TGT GTC TTT CTA ACT CAA ACT TCT GCA-3', yielding a 580 bp fragment. PCR primers for the *Prdm9* WT allele were Forward: 5'- GCT TAG GTA GCA GAA TTG AAG GGA AAG TC-3' and Reverse: 5'-TCG TGG CGT AAT AAT AGA GTG CCT TG-3', yielding a 401 bp fragment.

All mice were housed under controlled environmental conditions with free access to water and food, and illumination was on between 6 a.m. and 6 p.m. All experimental protocols were approved by the Animal Ethics Committee of the School of Medicine of Shandong University.

Production of the rat ZCWPW1 antibody

Antibodies to mouse ZCWPW1 were produced by Dia-an Biological Technology Incorporation (Wuhan, China). Briefly, a complementary DNA (cDNA) fragment encoding amino acids 448 to 622 of mouse *Zcwpw1* was inserted into the p-ET-32a + vector (EMD Millipore) and transfected into BL21-CodonPlus (DE3) Escherichia colicells. The cells were cultured at 37°C overnight and induced by addition of 0.2 mM isopropyl-1-thio-β-d-galactoside (Sigma-Aldrich) for 4 hours at 28°C. Cells were harvested by centrifugation and disrupted by sonication, and the soluble homogenates were purified by Ni-nitrilotriacetic acid (NI-NTA) Agarose (Qiagen) according to the manufacturer's

instructions. The protein was dialyzed in phosphate-buffered saline (PBS) and used to immunize rats, and the antiserum was affinity-purified on antigen-coupled CNBr-activated agarose (GE Healthcare).

Tissue collection and histological analysis

Testes from at least three mice for each genotype were dissected immediately after euthanasia, fixed in 4% (mass/vol) paraformaldehyde (Solarbio) for up to 24 h, stored in 70% ethanol, and embedded in paraffin after dehydration, and 5 μm sections were prepared and mounted on glass slides. After deparaffinization, slides were stained with hematoxylin for histological analysis using an epifluorescence microscope (BX52, Olympus), and images were processed using Photoshop (Adobe).

Chromosome spread immunofluorescence analysis

Spermatocyte spreads were prepared as previously described (Peters et al., 1997). Primary antibodies used for immunofluorescence were as follows: rabbit anti-ZCWPW1 (1:1,000 dilution; Dia-an Biological Technology Incorporation (Li et al., 2019a)), rat anti-ZCWPW1 (1:200 dilution; Dia-an Biological Technology Incorporation), mouse anti-SCP3 (1:500 dilution; Abcam #ab97672), rabbit anti-SCP1 (1:2,000 dilution; Abcam # ab15090), rabbit anti-RAD51 (1:200 dilution; Thermo Fisher Scientific #PA5-27195), rabbit anti-DMC1 (1:100 dilution; Santa Cruz Biotechnology #sc-22768), mouse anti-γH2AX (1:300 dilution; Millipore #05-636), mouse anti-pATM (1:500 dilution; Sigma-Aldrich #05-740), rabbit anti-MSH4 (1:500 dilution; Abcam #ab58666), rabbit anti-RNF212 (1:500 dilution; a gift from Mengcheng Luo, Wuhan University), mouse anti-MLH1 (1:50 dilution; BD Biosciences #550838), rabbit anti-H3K4me3 (1:500 dilution; Abcam #ab8580), and rabbit anti-H3K36me3 (1:500 dilution; Abcam #ab9050). Primary antibodies were detected with Alexa Fluor 488-, 594-, or 647-conjugated secondary antibodies (1:500 dilution, Thermo Fisher Scientific #A-11070, Abcam #ab150084, #ab150067, #ab150113, #ab150120, #ab150119, #ab150165, #ab150168, and #ab150167) for 1 h at room temperature. The slides were washed with PBS several times and mounted using VECTASHIELD medium with DAPI (Vector Laboratories, #H-1200). Immunolabeled chromosome spreads were imaged by confocal microscopy using a Leica TCS SP5 resonant-scanning

confocal microscope. Projection images were then prepared using ImageJ Software (NIH, v. 1.6.0-65) or Bitplane Imaris (v8.1) software.

Immunoblotting

To prepare protein extracts, tissues were collected from male C57BL/6 mice and lysed in TAP lysis buffer (50 mM HEPES-KOH, pH 7.5, 100 mM KCl, 2 mM EDTA, 10% glycerol, 0.1% NP-40, 10 mM NaF, 0.25 mM Na3VO4 and 50 mM β-glycerolphosphate) plus protease inhibitors (Roche, 04693132001) for 30 min on ice, followed by centrifugation at 4°C at 13,000 × *g* for 15 min. The supernatants were used for Western blotting. Equal amounts of protein were electrophoresed on 10% Bis-Tris protein gels (Invitrogen, NP0315), and the bands were transferred to polyvinylidene fluoride membranes (Millipore). The primary antibodies for immunoblotting included anti-tubulin (1:10,000 dilution; Proteintech Group, #11224-1-AP) and anti-ZCWPW1 (1:5,000 dilution; homemade). Immunoreactive bands were detected and analyzed with a Bio-Rad ChemiDoc MP Imaging System and Image Lab Software (Bio-Rad).

ChIP-seq experiments

PD 14 male mice were used to prepare cells for ChIP-seq (Chen et al., 2018). After removal of the tunica albuginea, testes were incubated in 5 ml PBS with collagenase type I (120 U/ml) at 35°C with gentle agitation for 10 min. The dispersed seminiferous tubules were further digested with 5 ml 0.25% trypsin, plus 0.1 ml DNase I (5 mg/ml) at 35°C for 8 min, and then terminated by adding 0.5 ml fetal bovine serum (FBS). The resulting suspension was passed through a 70 μm cellular filter. After centrifugation at 500 g for 5 min, the cells were resuspended in PBS and separated into 3×10^5 per tube. The collected cells from testes were cross-linked in 100 μL of 1% formaldehyde in PBS at room temperature for 10 min and this was followed by 25 μL of 1.25M glycine solution and mixing via gentle tapping and incubation at room temperature for 5 min. After centrifugation, the cell pellet was washed in PBS three times. Dynabeads Protein A beads (Life Technologies, 10001D) in a total volume of 25 μL were washed twice with 200 μL ice-cold 140 mM RIPA buffer (10 mM Tris-HCl pH 7.5, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% SDS, 0.1% Na-deoxycholate, 1% Triton X-

100, 1mM PMSF, 1×proteinase inhibitor Cocktail, and 20 mM Na-butyrate), followed by resuspension in RIPA buffer to a final volume of 200 μL in a 1.5ml tube. A total volume of 5 μl H3K4me3 antibody (Abcam, ab8580) or 7μl ZCWPW1 antibody (homemade, 5ug/μl) or 5μl H3K36me3 antibody (Abcam, ab9050) was added to the beads suspension, and this was followed by incubation on a tube rotator for at least 2.5 h at 4°C. The antibody-coated beads were then washed twice in 140mM RIPA buffer, followed by resuspension with 200 μL 140mM RIPA buffer.

The cross-linked cells were incubated in 150 μL lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA pH8.0, 0.5% SDS, 1mM PMSF, 1× proteinase inhibitor cocktail, and 20 mM Na-butyrate) for 20 min on ice, then sonicated using a Diagenode Bioruptor sonication device for 23 cycles (30 s on and then 30s off). A total volume of 150 μl 300 mM SDS-free RIPA buffer (10 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% Na-deoxycholate, 1mM PMSF, 1× Cocktail proteinase inhibitor, and 20 mM Na-butyrate) and 200μl 140mM SDS-free RIPA buffer were added to the samples. After centrifugation at 13,000 × *g* for 10 min at 4°C, 40 μL supernatant was removed and used as the sample input. The remaining supernatant was transferred to a 1 ml tube containing suspended antibody-coated Protein A beads, and this was followed by incubation on a tube rotator overnight at 4°C.

For the H3K4me3 and H3K36me3 antibodies, the incubated Protein A beads were washed once with RIPA buffer containing 250 mM NaCl, three times with RIPA buffer containing 500 mM NaCl, and once with TE buffer (10 mM Tris-HCl pH 8.0, 1mM EDTA). For the ZCWPW1 antibody, the incubated Protein A beads were washed twice with RIPA buffer containing 250 mM NaCl, once with RIPA buffer containing 500mM NaCl, and once with TE buffer. Next, the beads were transferred to a new 0.5ml tube and incubated in 100 μL ChIP elution buffer (10mM Tris-HCl pH8.0, 5mM EDTA, 300mM NaCl, 0.5% SDS) containing 5 μL proteinase K (Qiagen, 20mg/ml stock) at 55°C for 2 h and then at 65°C for 4 h. The eluate was transferred to a 0.5 mL tube, and the enriched DNA was purified by phenol–chloroform, followed by dissolution in 50 μL TE buffer.

An NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, E7645S) was used for library construction according to the product's instructions. DNA was first end repaired and A-tailed by adding 7 μL NEBNext Ultra II End Prep Reaction Buffer and 3 μL NEBNext Ultra II End Prep

Enzyme Mix. Samples were incubated at 20°C for 30min and then at 65°C for 30min, and finally cooled to 4°C in a thermal cycler. Adaptor ligation was performed by adding 30 μL NEBNext Ultra II Ligation Master Mix, 1 μL NEBNext Ligation Enhancer, 0.8 μL 200mM ATP, and 2.5 μL 15 μM Illumina Multiplexing Adaptors. Samples were thoroughly mixed and incubated at 20°C for 40 min. Following adaptor ligation, 1.2 volume SPRIselect beads (Beckman Coulter, B23318) were used to purify the DNA. PCR amplification was performed with NEBNext Ultra II Q5 Master Mix. The PCR cycle number was evaluated using a FlashGelTM System (Lonza, 57063). The volume of the PCR product was adjusted to 100 μL by adding 50 μl TE buffer. The 300–700 bp DNA fragments were selected with 0.5 volumes plus 0.5 volumes SPRIselect beads and then eluted in 20 μL water. The libraries were sequenced on a Hiseq X-ten instrument set for paired-end 150 bp sequencing (Illumina).

ChIP-seq Bioinformatics Analysis

The ChIP-seq raw reads were cropped to 100 bp, and the low quality reads were removed using Trimmomatic v0.32 (Bolger et al., 2014). Paired reads were mapped to the mouse genome (version mm10) by Bowtie2 v2.3.4.2 with the parameters "-X 2000 -no-discordant –no-contain" (Langmead and Salzberg, 2012). Reads with low mapping quality (MAPQ < 10) and PCR duplicated reads were removed by Samtools and Picard (DePristo et al., 2011; Li et al., 2009).

Reads of two replicates were merged to call the necessary peaks, while only one replicate of H3K4me3 and H3K36me3 in Lam et al was used to call peaks in $SCP3^{+}$ /$H1T^{-}$ spermatocytes using relatively-stringent conditions. The H3K4me3 peaks in this work and in Lam *et al.* were called by MACS2 v2.1.0 (Zhang et al., 2008) with the parameters " --SPMR -p 0.01 -nomodel" and "--SPMR -nomodel -q 0.05 ", the ZCWPW1 peaks were called with the parameters " -SPMR -p 0.001 –nomodel", the H3K36me3 peaks in Grey *et al.* and Lam *et al.* were called with the parameters " --SPMR --broad -nomodel" and "--SPMR --broad -nomodel -p 0.001", and the affinity-seq PRDM9 peaks in Walker *et al.* were called with the parameters "--SPMR -nomodel -p 0.05". The peaks for DMC1 and PRDM9 in Grey *et al*. were directly obtained from their published work (Khil et al.,

2012) and transformed to mm10 using the LiftOver application from UCSC. The SPO11 hotspots in Lange *et al* were directly obtained from their published work(Lange et al., 2016). The peaks intensities were denoted as the fold changes over input lambdas, which were obtained from the results produced by MACS2 callpeak. ZCWPW1 and H3K4me3 (Lam et al., 2019) peaks were further selected based on peak intensities greater than a 3-fold enrichment over the input lambda. Affinity-seq PRDM9 peaks in Walker *et al.* and H3K36me3 peaks in Lam *et al.* were further selected based on peak intensity greater than a 2-fold enrichment over the input lambda. The normalized signals of H3K4me3, H3K36me3, ZCWPW1, PRDM9, and DMC1 were generated using MACS2 bdgcmp, following the output produced by MACS2 callpeak with SPMR (reads per million for each covered position). The fold change over input lambda worked as the signal enrichment and was transformed into Bigwig using bedGraphToBigWig. ChIP-seq signal tracks were visualized by Integrative Genomics Viewer (Robinson et al., 2011). The computeMatrix algorithm in Deeptools2 (Ramirez et al., 2016) was used to calculate the normalized signal of each 40bp-size bins in the regions of peak center ± 2k bp. Deeptools plotHeatmap, plotProfile and R (3.4.4) were used to generate the profile plot and heatmap. The script findMotifsGenome.pl function in the HOMER software (Heinz et al., 2010) was used to examine the enrichment for transcription factor binding motifs. The gene-region association was determined using the GREATER software (McLean et al., 2010). The genomic regions including promoters (TSS ± 2k bp), exons, introns, intergenic regions, transposon elements, CpG islands, and distal enhancers were downloaded from the UCSC Table Browser under the mm10 version.

Analysis of the spatial overlap of ZCWPW1 peaks with genomic regions and transposon elements

The peak distribution over genome elements and the overlap between two types of peaks were calculated using bedtools intersect (v2.25.0) with the parameters -u, and minimum overlap.was 1bp (as default). The random binding sites (peaks) used as a control were created with the same number and size distribution as the observed peaks by using the regioneR package version1.18.1 (Gel et al.,

2016) implemented in R. Using regioneR, a Monte Carlo permutation test with 1,000 iterations was performed. In each iteration, the random binding sites were obtained through arbitrarily shuffled in the mouse genome. From this shuffling, the average overlap and standard deviation of the random binding site set was determined, as well as the statistical significance of the association between ZCWPW1 binding sites. Taking the overlap of ZCWPW1 peaks and transposable elements (from RepeatMasker) as an example, pair-end reads were aligned to mouse genome, and then uniquely mapped reads were used to call peaks in the ChIP-seq Bioinformatics Analysis. The random ZCWPW1 peaks were generated using the aforementioned regioneR package. The number of random ZCWPW1 peaks sharing overlap with different types of transposable elements (TE) was obtained in regioneR (whose overlap calculation is same as in bedtoool intersect). The final number of random peaks sharing overlap with transposon elements was indicated as the average of the results of 1000 iterations. To assess the significance of the overlap difference with TEs between observed ZCWPW1 peaks and random peaks, a permutation test was perfomed in regioneR.

RNA-seq experiments and bioinformatics analysis

The RNA was extracted from the testis with the Direct-zo RNA MiniPrep kit (Zymo). A total amount of 1.5 μg RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using the NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, USA) following the manufacturer's recommendations. QC-passed libraries were sequenced on the Hiseq X-ten instrument with the paired-end 150 bp.

The low-quality reads were removed using Trimmomatic v0.32 (Bolger et al., 2014). Paired reads were mapped to the mouse genome (version mm10) by hisat2 v2.1.0 (Pertea et al., 2016) and to the transcriptome by Salmon v 0.8.2. The DESeq2 v 1.18.0 software (Love et al., 2014) was used to identify DEGs from the raw counts produced by Salmon with the two conditions: P adjust < 0.05, and fold change ≥2. Metascape (Zhou et al., 2019) was used to perform Gene Ontology analysis of DEGs.

Statistical analysis

Two-tailed Wilcoxon rank sum tests were performed to obtain inferential statistical significance (p values) in related analyses by using the R function wilcox.test. No statistical methods were used to predetermine sample size.

Data availability

All data generated or analyzed during this study are included in the manuscript and supporting files. The raw sequencing data produced in this study (ChIP-seq data listed in Supplemental file1) and the RNA-seq data have been deposited with the Genome Sequence Archive (https://bigd.big.ac.cn/gsa/s/CjjpbIjf) under accession number PRJCA001901.

Acknowledgments

We are grateful for the interesting discussion with K. Liu from the University of Hong Kong, China, in the very initial phase of the study. We thank Translational Medicine Core Facility of Shandong University for consultation and instrument availability that supported this work and Jing Xin for assistance with animal work (Shandong University). This work was supported by the Major Program of the National Natural Science Foundation of China [31890780], the National Key Research and Development Programs of China [2018YFC1003400] and the Young Scholars Program of Shandong University (2016WLJH50).

Declaration of interests

The authors declare no competing interests.

References

Acquaviva, L., Szekvolgyi, L., Dichtl, B., Dichtl, B.S., de La Roche Saint Andre, C., Nicolas, A., and Geli, V. (2013). The COMPASS subunit Spp1 links histone methylation to initiation of meiotic recombination. Science *339*, 215-218.

682 Adams-Cioaba, M.A., and Min, J. (2009). Structure and function of histone
683 methylation binding proteins. Biochem Cell Biol *87*, 93-105.
684 Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G.,
685 and de Massy, B. (2010). PRDM9 is a major determinant of meiotic recombination
686 hotspots in humans and mice. Science *327*, 836-840.
687 Baudat, F., and de Massy, B. (2007). Regulating double-stranded DNA break repair
688 towards crossover or non-crossover during mammalian meiosis. Chromosome Res
689 *15*, 565-577.
690 Baudat, F., Imai, Y., and de Massy, B. (2013). Meiotic recombination in mammals:
691 localization and regulation. Nat Rev Genet *14*, 794-806.
692 Bergerat, A., de Massy, B., Gadelle, D., Varoutas, P.C., Nicolas, A., and Forterre, P.
693 (1997). An atypical topoisomerase II from Archaea with implications for meiotic
694 recombination. Nature *386*, 414-417.
695 Billings, T., Parvanov, E.D., Baker, C.L., Walker, M., Paigen, K., and Petkov, P.M.
696 (2013). DNA binding specificities of the long zinc-finger recombination protein
697 PRDM9. Genome Biol *14*, R35.
698 Blat, Y., Protacio, R.U., Hunter, N., and Kleckner, N. (2002). Physical and functional
699 interactions among basic chromosome organizational features govern early steps of
700 meiotic chiasma formation. Cell *111*, 791-802.
701 Bolcun-Filas, E., and Schimenti, J.C. (2012). Genetics of meiosis and recombination
702 in mice. Int Rev Cell Mol Biol *298*, 179-227.
703 Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for
704 Illumina sequence data. Bioinformatics *30*, 2114-2120.
705 Brick, K., Smagulova, F., Khil, P., Camerini-Otero, R.D., and Petukhova, G.V.
706 (2012). Genetic recombination is directed away from functional genomic elements in
707 mice. Nature *485*, 642-645.
708 Dai, J., Voloshin, O., Potapova, S., and Camerini-Otero, R.D. (2017). Meiotic
709 Knockdown and Complementation Reveals Essential Role of RAD51 in Mouse
710 Spermatogenesis. Cell Rep *18*, 1383-1394.
711 Davies, B., Hatton, E., Altemose, N., Hussin, J.G., Pratto, F., Zhang, G., Hinch, A.G.,
712 Moralli, D., Biggs, D., Diaz, R.*, et al.* (2016). Re-engineering the zinc fingers of
713 PRDM9 reverses hybrid sterility in mice. Nature *530*, 171-176.

714 de Massy, B. (2013). Initiation of meiotic recombination: how and where?
715 Conservation and specificities among eukaryotes. Annu Rev Genet *47*, 563-599.
716 DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C.,
717 Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M.*, et al.* (2011). A framework
718 for variation discovery and genotyping using next-generation DNA sequencing data.
719 Nat Genet *43*, 491-498.
720 Diagouraga, B., Clement, J.A.J., Duret, L., Kadlec, J., de Massy, B., and Baudat, F.
721 (2018). PRDM9 Methyltransferase Activity Is Essential for Meiotic DNA Double-
722 Strand Break Formation at Its Binding Sites. Mol Cell *69*, 853-865 e856.
723 Edelmann, W., Cohen, P.E., Kneitz, B., Winand, N., Lia, M., Heyer, J., Kolodner, R.,
724 Pollard, J.W., and Kucherlapati, R. (1999). Mammalian MutS homologue 5 is
725 required for chromosome pairing in meiosis. Nat Genet *21*, 123-127.
726 Eidahl, J.O., Crowe, B.L., North, J.A., McKee, C.J., Shkriabai, N., Feng, L., Plumb,
727 M., Graham, R.L., Gorelick, R.J., Hess, S.*, et al.* (2013). Structural basis for high-
728 affinity binding of LEDGF PWWP to mononucleosomes. Nucleic Acids Res *41*, 3924-
729 3936.
730 Garcia, V., Gray, S., Allison, R.M., Cooper, T.J., and Neale, M.J. (2015). Tel1(ATM)-
731 mediated interference suppresses clustered meiotic double-strand-break formation.
732 Nature *520*, 114-118.
733 Gel, B., Diez-Villanueva, A., Serra, E., Buschbeck, M., Peinado, M.A., and
734 Malinverni, R. (2016). regioneR: an R/Bioconductor package for the association
735 analysis of genomic regions based on permutation tests. Bioinformatics *32*, 289-291.
736 Gray, S., and Cohen, P.E. (2016). Control of Meiotic Crossovers: From Double-
737 Strand Break Formation to Designation. Annu Rev Genet *50*, 175-210.
738 Grey, C., Baudat, F., and de Massy, B. (2018). PRDM9, a driver of the genetic map.
739 PLoS Genet *14*, e1007479.
740 Grey, C., Clement, J.A., Buard, J., Leblanc, B., Gut, I., Gut, M., Duret, L., and de
741 Massy, B. (2017). In vivo binding of PRDM9 reveals interactions with noncanonical
742 genomic sites. Genome Res *27*, 580-590.
743 Handel, M.A., and Schimenti, J.C. (2010). Genetics of mammalian meiosis:
744 regulation, dynamics and impact on fertility. Nat Rev Genet *11*, 124-136.

He, F., Umehara, T., Saito, K., Harada, T., Watanabe, S., Yabuki, T., Kigawa, T., Takahashi, M., Kuwasako, K., Tsuda, K.*, et al.* (2010). Structural insight into the zinc finger CW domain as a histone modification reader. Structure *18*, 1127-1139.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Hinch, A.G., Zhang, G., Becker, P.W., Moralli, D., Hinch, R., Davies, B., Bowden, R., and Donnelly, P. (2019). Factors influencing meiotic recombination revealed by whole-genome sequencing of single sperm. Science *363*.

Hoppmann, V., Thorstensen, T., Kristiansen, P.E., Veiseth, S.V., Rahman, M.A., Finne, K., Aalen, R.B., and Aasland, R. (2011). The CW domain, a new histone recognition module in chromatin proteins. EMBO J *30*, 1939-1952.

Hunter, N. (2015). Meiotic Recombination: The Essence of Heredity. Cold Spring Harb Perspect Biol *7*.

Inagaki, A., Schoenmakers, S., and Baarends, W.M. (2010). DNA double strand break repair, chromosome synapsis and transcriptional silencing in meiosis. Epigenetics *5*, 255-266.

Keeney, S., Giroux, C.N., and Kleckner, N. (1997). Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell *88*, 375-384.

Keeney, S., Lange, J., and Mohibullah, N. (2014). Self-organization of meiotic recombination initiation: general principles and molecular pathways. Annu Rev Genet *48*, 187-214.

Khil, P.P., Smagulova, F., Brick, K.M., Camerini-Otero, R.D., and Petukhova, G.V. (2012). Sensitive mapping of recombination hotspots using sequencing-based detection of ssDNA. Genome Res *22*, 957-965.

Kneitz, B., Cohen, P.E., Avdievich, E., Zhu, L., Kane, M.F., Hou, H., Jr., Kolodner, R.D., Kucherlapati, R., Pollard, J.W., and Edelmann, W. (2000). MutS homolog 4 localization to meiotic chromosomes is required for chromosome pairing during meiosis in male and female mice. Genes Dev *14*, 1085-1097.

776 Kumar, R., Oliver, C., Brun, C., Juarez-Martinez, A.B., Tarabay, Y., Kadlec, J., and
777 de Massy, B. (2018). Mouse REC114 is essential for meiotic DNA double-strand
778 break formation and forms a complex with MEI4. Life Sci Alliance *1*, e201800259.
779 Lam, K.G., Brick, K., Cheng, G., Pratto, F., and Camerini-Otero, R.D. (2019). Cell-
780 type-specific genomics reveals histone modification dynamics in mammalian
781 meiosis. Nat Commun *10*, 3821.
782 Lange, J., Pan, J., Cole, F., Thelen, M.P., Jasin, M., and Keeney, S. (2011). ATM
783 controls meiotic double-strand-break formation. Nature *479*, 237-240.
784 Lange, J., Yamada, S., Tischfield, S.E., Pan, J., Kim, S., Zhu, X., Socci, N.D., Jasin,
785 M., and Keeney, S. (2016). The Landscape of Mouse Meiotic Double-Strand Break
786 Formation, Processing, and Repair. Cell *167*, 695-708 e616.
787 Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie
788 2. Nat Methods *9*, 357-359.
789 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G.,
790 Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The
791 Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.
792 Li, M., Huang, T., Li, M.J., Zhang, C.X., Yu, X.C., Yin, Y.Y., Liu, C., Wang, X., Feng,
793 H.W., Zhang, T.*, et al.* (2019a). The histone modification reader ZCWPW1 is
794 required for meiosis prophase I in male but not in female mice. Sci Adv *5*, eaax1101.
795 Li, R., Bitoun, E., Altemose, N., Davies, R.W., Davies, B., and Myers, S.R. (2019b).
796 A high-resolution map of non-crossover events reveals impacts of genetic diversity
797 on mammalian meiotic recombination. Nat Commun *10*, 3900.
798 Liu, Y., Tempel, W., Zhang, Q., Liang, X., Loppnau, P., Qin, S., and Min, J. (2016).
799 Family-wide Characterization of Histone Binding Abilities of Human CW Domain-
800 containing Proteins. J Biol Chem *291*, 9000-9013.
801 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change
802 and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.
803 Lynn, A., Soucek, R., and Borner, G.V. (2007). ZMM proteins during meiosis:
804 crossover artists at work. Chromosome Res *15*, 591-605.
805 Mahgoub, M., Paiano, J., Bruno, M., Wu, W., Pathuri, S., Zhang, X., Ralls, S.,
806 Cheng, X., Nussenzweig, A., and Macfarlan, T. (2019). Dual Histone Methyl Reader
807 ZCWPW1 Facilitates Repair of Meiotic Double Strand Breaks. BioRxiv.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol *28*, 495-501.

Myers, S., Bowden, R., Tumian, A., Bontrop, R.E., Freeman, C., MacFie, T.S., McVean, G., and Donnelly, P. (2010). Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. Science *327*, 876-879.

Paigen, K., and Petkov, P.M. (2018). PRDM9 and Its Role in Genetic Recombination. Trends Genet *34*, 291-300.

Panizza, S., Mendoza, M.A., Berlinger, M., Huang, L., Nicolas, A., Shirahige, K., and Klein, F. (2011). Spo11-accessory proteins link double-strand break sites to the chromosome axis in early meiotic recombination. Cell *146*, 372-383.

Parvanov, E.D., Petkov, P.M., and Paigen, K. (2010). Prdm9 controls activation of mammalian recombination hotspots. Science *327*, 835.

Parvanov, E.D., Tian, H., Billings, T., Saxl, R.L., Spruce, C., Aithal, R., Krejci, L., Paigen, K., and Petkov, P.M. (2017). PRDM9 interactions with other proteins provide a link between recombination hotspots and the chromosomal axis in meiosis. Mol Biol Cell *28*, 488-499.

Perry, J., and Zhao, Y. (2003). The CW domain, a structural module shared amongst vertebrates, vertebrate-infecting parasites and higher plants. Trends Biochem Sci *28*, 576-580.

Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc *11*, 1650-1667.

Peters, A.H., Plug, A.W., van Vugt, M.J., and de Boer, P. (1997). A drying-down technique for the spreading of mammalian meiocytes from the male and female germline. Chromosome Res *5*, 66-68.

Pittman, D.L., Cobb, J., Schimenti, K.J., Wilson, L.A., Cooper, D.M., Brignull, E., Handel, M.A., and Schimenti, J.C. (1998). Meiotic prophase arrest with failure of chromosome synapsis in mice deficient for Dmc1, a germline-specific RecA homolog. Mol Cell *1*, 697-705.

Powers, N.R., Parvanov, E.D., Baker, C.L., Walker, M., Petkov, P.M., and Paigen, K. (2016). The Meiotic Recombination Activator PRDM9 Trimethylates Both H3K36 and H3K4 at Recombination Hotspots In Vivo. PLoS Genet *12*, e1006146.

841 Qin, S., and Min, J. (2014). Structure and function of the nucleosome-binding PWWP
842 domain. Trends Biochem Sci *39*, 536-547.
843 Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne,
844 S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for
845 deep-sequencing data analysis. Nucleic Acids Res *44*, W160-165.
846 Reynolds, A., Qiao, H., Yang, Y., Chen, J.K., Jackson, N., Biswas, K., Holloway,
847 J.K., Baudat, F., de Massy, B., Wang, J.*, et al.* (2013). RNF212 is a dosage-sensitive
848 regulator of crossing-over during mammalian meiosis. Nat Genet *45*, 269-278.
849 Robert, T., Nore, A., Brun, C., Maffre, C., Crimi, B., Bourbon, H.M., and de Massy, B.
850 (2016). The TopoVIB-Like protein family is required for meiotic DNA double-strand
851 break formation. Science *351*, 943-949.
852 Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz,
853 G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26.
854 Rondelet, G., Dal Maso, T., Willems, L., and Wouters, J. (2016). Structural basis for
855 recognition of histone H3K36me3 nucleosome by human de novo DNA
856 methyltransferases 3A and 3B. J Struct Biol *194*, 357-367.
857 Segurel, L. (2013). The complex binding of PRDM9. Genome Biol *14*, 112.
858 Smagulova, F., Gregoretti, I.V., Brick, K., Khil, P., Camerini-Otero, R.D., and
859 Petukhova, G.V. (2011). Genome-wide analysis reveals novel molecular features of
860 mouse recombination hotspots. Nature *472*, 375-378.
861 Sommermeyer, V., Beneut, C., Chaplais, E., Serrentino, M.E., and Borde, V. (2013).
862 Spp1, a member of the Set1 Complex, promotes meiotic DSB formation in promoters
863 by tethering histone H3K4 methylation sites to chromosome axes. Mol Cell *49*, 43-
864 54.
865 Stanzione, M., Baumann, M., Papanikos, F., Dereli, I., Lange, J., Ramlal, A.,
866 Trankner, D., Shibuya, H., de Massy, B., Watanabe, Y.*, et al.* (2016). Meiotic DNA
867 break formation requires the unsynapsed chromosome axis-binding protein IHO1
868 (CCDC36) in mice. Nat Cell Biol *18*, 1208-1220.
869 Sun, F., Fujiwara, Y., Reinholdt, L.G., Hu, J., Saxl, R.L., Baker, C.L., Petkov, P.M.,
870 Paigen, K., and Handel, M.A. (2015). Nuclear localization of PRDM9 and its role in
871 meiotic chromatin modifications and homologous synapsis. Chromosoma *124*, 397-
872 415.

873 Tarsounas, M., Morita, T., Pearlman, R.E., and Moens, P.B. (1999). RAD51 and
874 DMC1 form mixed complexes associated with mouse meiotic chromosome cores
875 and synaptonemal complexes. J Cell Biol *147*, 207-220.
876 Tesse, S., Bourbon, H.M., Debuchy, R., Budin, K., Dubois, E., Liangran, Z., Antoine,
877 R., Piolot, T., Kleckner, N., Zickler, D*., et al.* (2017). Asy2/Mer2: an evolutionarily
878 conserved mediator of meiotic recombination, pairing, and global chromosome
879 compaction. Genes Dev *31*, 1880-1893.
880 Tock, A.J., and Henderson, I.R. (2018). Hotspots for Initiation of Meiotic
881 Recombination. Front Genet *9*, 521.
882 Vezzoli, A., Bonadies, N., Allen, M.D., Freund, S.M., Santiveri, C.M., Kvinlaug, B.T.,
883 Huntly, B.J., Gottgens, B., and Bycroft, M. (2010). Molecular basis of histone
884 H3K36me3 recognition by the PWWP domain of Brpf1. Nat Struct Mol Biol *17*, 617-
885 619.
886 Vrielynck, N., Chambon, A., Vezon, D., Pereira, L., Chelysheva, L., De Muyt, A.,
887 Mezard, C., Mayer, C., and Grelon, M. (2016). A DNA topoisomerase VI-like
888 complex initiates meiotic recombination. Science *351*, 939-943.
889 Wahls, W.P. (1998). Meiotic recombination hotspots: shaping the genome and
890 insights into hypervariable minisatellite DNA change. Curr Top Dev Biol *37*, 37-75.
891 Walker, M., Billings, T., Baker, C.L., Powers, N., Tian, H., Saxl, R.L., Choi, K., Hibbs,
892 M.A., Carter, G.W., Handel, M.A*., et al.* (2015). Affinity-seq detects genome-wide
893 PRDM9 binding sites and reveals the impact of prior chromatin modifications on
894 mammalian recombination hotspot usage. Epigenetics Chromatin *8*, 31.
895 Wells, D., Bitoun, E., Moralli, D., Zhang, G., Hinch, A.G., Donnelly, P., Green, C.,
896 and Myers, S.R. (2019). ZCWPW1 is recruited to recombination hotspots by
897 PRDM9, and is essential for meiotic double strand break repair. BioRxiv.
898 Yang, F., Gell, K., van der Heijden, G.W., Eckardt, S., Leu, N.A., Page, D.C.,
899 Benavente, R., Her, C., Hoog, C., McLaughlin, K.J*., et al.* (2008). Meiotic failure in
900 male mice lacking an X-linked factor. Genes Dev *22*, 682-691.
901 Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E.,
902 Nusbaum, C., Myers, R.M., Brown, M., Li, W*., et al.* (2008). Model-based analysis of
903 ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun *10*, 1523.

Zickler, D., and Kleckner, N. (2015). Recombination, Pairing, and Synapsis of Homologs during Meiosis. Cold Spring Harb Perspect Biol *7*.

Figure legends

Figure 1 The H3K4me3 reader function of ZCWPW1 is required for synapsis and meiotic recombination

(A) Hematoxylin staining of adult C57BL/6 wild type, *Zcwpw1*$^{-/-}$, and *Zcwpw1*$^{KI/KI}$ testes (left panel) and epididymides (right panel). Adult *Zcwpw1*$^{-/-}$ and *Zcwpw1*$^{KI/KI}$ testis sections showed near complete arrest of spermatogenesis. Arrows, apoptotic spermatocytes; arrowheads, empty seminiferous tubules; asterisks, seminiferous tubules lacking post-meiotic spermatocytes. The spermatogenic arrest led to empty epididymides in adult *Zcwpw1*$^{-/-}$ and *Zcwpw1*$^{KI/KI}$ mice. (St) Seminiferous tubules, (Ep) Epididymides. Adult mice (6-8weeks) with n = 5 for each genotype. **(B)** Chromosome spreads of spermatocytes from the testes of adult WT (upper panel), *Zcwpw1*$^{-/-}$ (middle panel), and *Zcwpw1*$^{KI/KI}$ (lower panel) males were immunostained for the SC marker proteins SYCP1 (green) and SYCP3 (red). The arrow indicates a pachytene spermatocyte in WT mice, with completely synapsed chromosomes, and the arrowheads indicate the pachytene-like spermatocytes in adult *Zcwpw1*$^{-/-}$ and *Zcwpw1*$^{KI/KI}$ mice with incompletely synapsed chromosomes. **(C)** Chromosome spreads of spermatocytes from the testes of adult WT and *Zcwpw1*$^{KI/KI}$ males were immunostained for DMC1 (green) and SYCP3 (red). Representative images of spermatocytes at zygotene and pachytene in WT and at zygotene and pachytene-like stages in *Zcwpw1*$^{KI/KI}$ are shown. **(D)** Each dot represents the number of DMC1 foci per cell, with black dots indicating WT spermatocytes and red dots indicating *Zcwpw1*$^{KI/KI}$ spermatocytes. Solid lines show the mean and SD of foci number in each group of spermatocytes. P values were calculated by Student's t-test. N represents the number of cells counted, with black indicating WT spermatocytes and red indicating *Zcwpw1*$^{KI/KI}$ spermatocytes. **(E)** Chromosome spreads of spermatocytes from the testes of adult WT and *Zcwpw1*$^{KI/KI}$ males immunostained for RAD51 (green) and SYCP3 (red). **(F)** Each dot represents the number of RAD51 foci per cell, with black dots indicating WT spermatocytes and red dots indicating *Zcwpw1*$^{KI/KI}$ spermatocytes. Solid lines show the mean and SD of the foci number for each group of spermatocytes. P values were calculated by Student's t test. Adult mice (6-8weeks) with n = 3 for each genotype.

Figure 2 The H3K4me3 reader function of ZCWPW1 is required for DSB repair

(A) Chromosome spreads of spermatocytes from the testes of adult WT, *Zcwpw1*$^{-/-}$, and *Zcwpw1*$^{KI/KI}$ males immunostained for the DSB marker γH2AX (green) and SYCP3 (red). **(B)** Chromosome spreads of spermatocytes from the testes of adult WT, *Zcwpw1*$^{-/-}$, and *Zcwpw1*$^{KI/KI}$ males immunostained for the DSB repair protein p-ATM (green) and SYCP3 (red). Representative images are shown for spermatocytes at the leptotene, zygotene, pachytene (arrow indicates the XY body), and pachytene-like (arrowheads indicate the p-ATM signal) stages of the three genotypes. All experiments were performed on adult mice (6-8weeks) with n = 3 for each genotype.

Figure 3 ZCWPW1 is an H3K4me3 reader

(A) ChIP-seq genome snapshot of the distribution of H3K4me3 and ZCWPW1 peaks in C57BL/6 WT, *Zcwpw1*$^{-/-}$, and *Zcwpw1*$^{KI/KI}$ mice along a 554-kb region of chromosome 14. H3K4me3 and ZCWPW1 signals were normalized (See Methods). Overlapping regions are indicated by grey or blue shaded areas, while non-overlapping regions of interest are indicated by orange shaded areas. R1 and R2 represent two independent replicates. The H3K4me3 tract (red)was generated with isolated stage-specific (SCP3^{+}H1T^{-}) spermatocyte nuclei (Lam et al., 2019) The unit of Y axis is fold change as described in the method. **(B)** Venn diagram showing the overlap between ZCWPW1 peaks and H3K4me3 peaks. H3K4me3 data generated in whole testes (left, this study) compared with H3K4me3 data generated with isolated stage-specific (SCP3^{+}H1T^{-}) spermatocyte nuclei (right, Lam *et al.*). Italics (*14,387*) indicates the number of H3K4me3 peaks overlapping ZCWPW1, while standard font (14,369) indicates the number of ZCWPW1 peaks overlapping with H3K4me3 marks. Sometimes, one broad ZCWPW1 peak perhaps overlap with two narrow H3K4me3 peaks, which will cause these two numbers differ. P-values were calculated by using the permTest (see Methods, ntimes=1000). **(C)** Profile plot of averaged normalized H3K4me3 signals (see Methods) in ZCWPW1 peaks obtained in this work and in Lam *et al*. The profile shows the average values over 4-kb intervals for all 14,688 detected peaks (binding sites). The unit of Y axis is average fold change as described in the method. **(D)** Profile plot of the averaged ZCWPW1 signal in 14,688 ZCWPW1 peaks, in C57BL/6 WT,

Zcwpw1$^{-/-}$ and *Zcwpw1*$^{KI/KI}$ mice. The unit of Y axis is average fold change as described in the method. **(E)** Pie chart showing the ratio of four ZCWPW1 peak groups determined by their overlap with histone modification peaks generated with isolated stage-specific spermatocyte nuclei (Lam *et al.*). The "+" indicates overlap, while "–" indicates no overlap. All ChIP-seq experiments were performed in PD14 mice with n = 4 for each genotype.

Figure 4 ZCWPW1 binding is strongly promoted by the histone modification activity of PRDM9

(A) ChIP-seq genome snapshot of the distribution of H3K4me3, H3K36me3, ZCWPW1, and PRDM9 peaks in C57BL/6 mice along a 372-kb region of chromosome 12. Overlapping peaks in samples from all four analyses are indicated by grey or blue shaded areas, while non-overlapping regions of interest are indicated by orange shaded areas. The unit of Y axis is fold change as described in the method. **(B)** Venn diagram showing the overlap between PRDM9 peaks and ZCWPW1 binding sites. On the left are *in vivo* PRDM9 data generated by Grey *et al.* (2018), while on the right are *in vitro* affinity-seq PRDM9 data generated by Walker *et al.* (2015). Italics indicate PRDM9 peak overlap with ZCWPW1, while standard font indicates ZCWPW1 peak overlap with PRDM9 peaks. **(C)** Heatmap showing the correlation among H3K4me3, H3K36me3, and ZCWPW1 with PRDM9 peaks (Grey et al. 2018). Each row represents a PRDM9 binding site of ±2 kb around the center and ranked by ZCWPW1 signal from the highest to the lowest. Color indicates normalized ChIP-seq signal (See Methods). **(D)** ChIP-seq genome snapshot showing changes in H3K4me3 and ZCWPW1 binding distributions following *Prdm9* knockout (*Prdm9*$^{-/-}$) along a 27/12-kb region of chromosome 16 and a 76-kb region of chromosome 7. The PRDM9 data were obtained from Grey *et al* (2018). The unit of Y axis is fold change as described in the method. **(E)** Profile plot of averaged H3K4me3 and ZCWPW1 signals obtained in this work with two types of PRDM9 peaks (Grey et al., 2018) following *Prdm9* knockout. The Y-axis shows log base-2 transformation of the normalized signal. The unit of Y axis is average fold change as described in the method. All ChIP-seq experiments were performed using PD14 mice with n = 4 for each genotype.

Figure 5 ZCWPW1 localizes to DMC1-labelled DSB hotspots in a PRDM9-dependent manner

(A) ChIP-seq genome snapshot of the distribution of H3K4me3, H3K36me3, ZCWPW1, PRDM9, and DMC1 peaks in C57BL/6 mice with changes in H3K4me3 and ZCWPW1 binding sites in *Prdm9* knockout mice. ZCWPW1 and DMC1 overlapping regions are indicated by grey shaded areas, while non-overlapping regions of interest are indicated by orange or blue shaded areas. The unit of Y axis is fold change as described in the method. (**B)** Heatmap showing DMC1 (Grey *et al*), H3K4me3 (Lam *et al*), H3K36me3 (Lam *et al*), H3K4me3 (WT and *Prdm9*$^{-/-}$), and ZCWPW1 (WT and *Prdm9*$^{-/-}$) corresponding with Gained, Lost, or Maintained groups of ZCWPW1 peaks (also see Figure 4–figure supplement 3A). We used k-means clustering to define the “gain” and “maintain” group. We performed k-means clustering analysis to define the "gain" and "maintain" groups, and found two different subtypes within each group: a subtype with H3K4me3 signal similar along the ZCWPW1 peaks ± 2kb, and a subtype with H3K4me3 mostly enriched on the ZCWPW1 peak center. All ChIP-seq experiments were performed using PD14 mice with n = 4 for each genotype. (**C)** Working model of ZCWPW1 in recognizing H3K4me3 and H3K36me3 deposited by PRDM9.

Figure 1–figure supplement 1 Generation of *Zcwpw1* reader-dead-mutant mice and *Prdm9* knockout mice

(A) Schematic representation of the genome editing strategy to generate knock-in *Zcwpw1* reader-dead-mutant mice. **(B)** Sequence alignment of human ZCWPW1 (NP_060454.3) and mouse ZCWPW1 (NP_001005426.2). Conserved regions of ZCWPW1 are indicated by green shaded areas for the zf-CW domain and grey shaded areas for the PWWP domain. Red asterisks indicate the point mutation sites. **(C)** Schematic representation of the CRISP-cas9 genome editing strategy used to generate the *Prdm9* knockout mice showing the gRNAs (arrows), the corresponding coding exons (black and red thick lines), and the non-coding exons (gray thick lines). Red thick lines (coding exons) represent approximately12 kb deleted from the wild-type *Prdm9* allele.

Figure 1–figure supplement 2 Distribution pattern of ZCWPW1 in WT, *Zcwpw1*$^{KI/KI}$ and *Zcwpw1*$^{-/-}$ mice

(A) Western blots showing that ZCWPW1 was not detected in *Zcwpw1*$^{-/-}$ testes but was present in WT testes, and with a similarly intense signal for *Zcwpw1*$^{KI/KI}$. Tubulin was used as the loading control. The experiments were performed using PD20 mice with n = 3 for each genotype. **(B)** Western blots using an antibody against ZCWPW1 showing that ZCWPW1 was not detected in *Zcwpw1*$^{-/-}$ testes but was present in WT testes. Asterisks indicate candidate band for ZCWPW1. **(C)** Immunofluorescence staining of ZCWPW1 and SYCP3 in WT, *Zcwpw1*$^{KI/KI}$ and *Zcwpw1*$^{-/-}$ histological cross-sections. DNA was stained with DAPI. Abbreviations: L/Z, leptotene/zygotene spermatocytes; Pa, pachytene spermatocytes; Spc, spermatocytes; ES, elongated spermatids. **(D)** The numbers of synapsed chromosome pairs in WT, *Zcwpw1*$^{-/-}$ and *Zcwpw1*$^{KI/KI}$ spermatocytes. In *Zcwpw1*$^{-/-}$ and *Zcwpw1*$^{KI/KI}$ spermatocytes, the average number of synapsed chromosome pairs was 8. All experiments were performed on adult mice (6−8weeks old) with n = 3 for each genotype.

Figure 1–figure supplement 3 Meiotic recombination defects in *Zcwpw1* knock-in mice

(A) Chromosome spreads of spermatocytes from the testes of adult WT and *Zcwpw1*$^{KI/KI}$ males immunostained for the recombination factor MSH4 (green) and SYCP3 (red). **(B)** Each dot represents the number of MSH4 foci per cell. Black dots indicate WT spermatocytes, and red dots indicate *Zcwpw1*$^{KI/KI}$ spermatocytes. Solid lines show the mean and SD of foci in each group of spermatocytes. P -values were calculated by Student's t-test. **(C)** Chromosome spreads of spermatocytes from WT and *Zcwpw1*$^{KI/KI}$ male mice immunostained for the recombination factor RNF212 (green) and SYCP3 (red). Representative images are shown for spermatocytes at the zygotene, pachytene, and pachytene-like stages of the three genotypes. **(D)** Each dot represents the number of RNF212 foci per cell, with black dots indicating WT spermatocytes and red dots indicating *Zcwpw1*$^{KI/KI}$ spermatocytes. Solid lines show the mean and SD of foci number for each group of spermatocytes. P-values were calculated by Student's t-test. **(E)** Chromosome spreads of spermatocytes from the testes of adult WT

and *Zcwpw1*$^{KI/KI}$ males immunostained for MLH1 (green) and SYCP3 (red). Representative images are shown for spermatocytes at the pachytene stage in WT and the pachytene-like stage in *Zcwpw1*$^{KI/KI}$. All experiments were performed on adult mice (6−8weeks old) with n = 3 for each genotype.

Figure 3–figure supplement 1 Genome-wide properties of ZCWPW1-associated binding sites

(A) The distribution of ZCWPW1 peak intensity and peak length by location in genomic elements. (***P < 0.001 by two-tailed Wilcoxon rank sum test). The red boxplots and left Y-axis indicate peak intensity, while the blue boxplots and right Y-axis indicate peak length. **(B)** Percentage of ZCWPW1 binding sites (peaks) overlapping with DNA elements compared with the random binding sites obtained by random shuffling of the identified ZCWPW1 binding sites. The X-axis indicates the percentage of binding sites overlapping with DNA elements. The random groups and P-values were generated using permTest, an R function in the regioneR package (see Methods, ntimes=1000). The barplot height of the random group represents the means of 1000 tests. Bars represent ± SD. **(C)** Percentage of ZCWPW1 binding sites (peaks) overlapping with transposable elements (TEs) compared with the random binding sites obtained by random shuffling of the ZCWPW1 binding sites. The X-axis indicates the percentage of binding sites overlapping with TEs. The random groups and P-values were generated using the same method as Figure 3–figure supplement 1B. **(D)** Percentage of ZCWPW1 binding sites (peaks) overlapping with Alu repeats compared with the binding sites obtained by random shuffling of the identified ZCWPW1 binding sites. The X-axis indicates the percentage of binding sites overlapping with Alu repeats. The random groups and P-value were generated as above.

Figure 3–figure supplement 2 Transcriptional profiling analysis of WT and *Zcwpw1*$^{-/-}$ testes

(A) Scatterplot of DEGs between WT and *Zcwpw1*$^{-/-}$ at PD14. Downregulated in *Zcwpw1*$^{-/-}$, red; upregulated in *Zcwpw1*$^{-/-}$, blue; no significant change (NC), green. P-adjust < 0.05 and FC≥ 2 (Fold change). **(B)** Functional enrichment analysis of DEGs by Metascape. The upper bar chart shows

DEGs with lower expression in *Zcwpw1*$^{-/-}$, and the lower bar chart shows DEGs with higher expression in *Zcwpw1*$^{-/-}$. **(C)** Venn diagram of the overlap between the genes with promoter regions overlapping ZCWPW1 binding sites and DEGs showing lower expression in *Zcwpw1*$^{-/-}$. All RNA-seq experiments were performed on PD14 mice with n = 3 for each genotype.

Figure 3–figure supplement 3 Correlations between ZCWPW1 peaks and H3K4me3 and H3K36me3 peaks

(A) Percentage of ZCWPW1 binding sites (peaks) overlapping with H3K36me3 peaks identified by Lam *et al*. (2019) or Grey *et al.* (2018), compared with the random binding sites obtained by random shuffling of the identified ZCWPW1 peaks. The X-axis indicates the percentage of binding sites overlapping with H3K36me3 peaks. The random groups and P-values were generated using the method described above. **(B)** Profile plot of the averaged normalized H3K36me3 signal (See Methods) obtained from Grey *et al.* (whole testes) and Lam *et al*. (isolated stage-specific spermatocyte nuclei) in ZCWPW1 peaks. The profile shows the average values over 4-kb intervals for all 14688 peaks (binding sites). **(C)** Boxplots showing the peak intensity of four ZCWPW1 groups determined by their overlap with histone modification peaks (Lam *et al*.). + indicates overlap, - indicates no overlap. The P-values was calculated using the two-tailed Wilcoxon rank sum test.
(D) Heatmap and averaged profile plot of the H3K36me3 signal (Grey *et al.*) on two types of H3K4me3 peaks (Grey *et al.*), including H3K4me3 peaks overlapping ZCWPW1 and H3K4me3 peaks not overlapping ZCWPW1. The 14,387 H3K4me3 peaks were randomly sampled from the total H3K4me3 peaks not overlapping with ZCWPW1 (61,384). Each row of the heatmap shows the H3K36me3 distribution on an H3K4me3 peak center ± 2k bp. The color change from white to purple indicates a change in the normalized H3K4me3 signal from weak to strong.

Figure 3–figure supplement 4 ZCWPW1, H3K4me3 and H3K36me3 localize to the nucleus in leptotene and zygotene spermatocytes

(A) Chromosome spreads of spermatocytes from the testes of adult WT were immunostained for H3K4me3 (red), ZCWPW1 (green), and SYCP3 (blue). Representative images of spermatocytes at leptotene and zygotene stages are shown. **(B)** Chromosome spreads of spermatocytes from the testes of adult WT were immunostained for H3K36me3 (red), ZCWPW1 (green), and SYCP3 (blue). Representative images of spermatocytes at leptotene (upper panels) and zygotene (lower panels) stages are shown. All experiments were performed on adult mice (6-8weeks old) with n = 3 for each genotype.

Figure 4–igure supplement 1 Correlation between ZCWPW1 binding sites and PRDM9-induced dual histone methylation

(A) The rank-first *de novo* binding motif of ZCWPW1. **(B)** The similarity of this ZCWPW1 binding motif with those of other transcription factors. **(C)** Barplots showing the constituent ratio in four types of ZCWPW1 peaks, representing the ratio of ZCWPW1 peaks sharing overlap with PRDM9 peaks, and those not sharing overlap with PRDM9 peaks. Black indicates overlap, white indicates no overlap. The four ZCWPW1 groups classified by their overlap with histone modification peaks (Lam *et al.*). + indicates overlap, - indicates no overlap. **(D)** Profile plot of averaged H3K4me3 and H3K36me3 signals in different types of PRDM9 peaks generated by Grey *et al*. The profile shows the average values over 4-kb intervals for all 2,601 detected PRDM9 binding sites. **(E)** Profile plot of averaged H3K4me3 and H3K36me3 in different types of PRDM9 peaks generated by Walker et al.; the profile shows the average values over 4-kb intervals for all 36898 of the detected PRDM9 binding regions. **(F)** Heatmap showing the correlation among H3K4me3, H3K36me3, and ZCWPW1 in the PRDM9 peaks (Walker *et al.*). Each row in the heatmap represents a PRDM9 binding site of ±2 kb around the center and ranked from the highest to the lowest according to ZCWPW1 signal intensity. Color indicates normalized ChIP-seq signal (see methods).

Figure 4–figure supplement 2 Correlation between ZCWPW1 chromatin occupancy and PRDM9-induced H3K4me3

(A) Heatmap showing H3K4me3 and ZCWPW1 signals obtained in this work, in two types of PRDM9 peaks (Grey *et al*.) following *Prdm9* knockout. Each row in the heatmaps represents a PRDM9 binding site of ±2 kb around the center and ranked from the highest to the lowest according to PRDM9 density. Color indicates ChIP-seq density. **(B)** Heatmap showing H3K4me3 and ZCWPW1 signals obtained from this work, on two types of PRDM9 peaks identified by Walker *et al*. following *Prdm9* knockout. Each row in the heatmaps represents a PRDM9 binding site of ±2 kb around the center and ranked from the highest to the lowest according to PRDM9 density. Color indicates ChIP-seq density. **(C)** Profile plot of averaged H3K4me3 and ZCWPW1 signals obtained in this work and in two types of PRDM9 peaks identified by Walker *et al*. following *Prdm9* knockout. The Y-axis indicates the log base-2 transformation of the normalized signals. All ChIP-seq experiments were performed using PD14 mice with n = 4 for each genotype.

Figure 4–figure supplement 3 Change in ZCWPW1 chromatin occupancy following *Prdm9* knockout

(A) Venn diagram showing the overlap of the ZCWPW1 peaks between WT and *Prdm9* knockout (*Prdm9*$^{-/-}$). ZCWPW1 peaks were grouped by their detection in either WT, *Prdm9*$^{-/-}$, or both as either gained (652) – indicating peaks only found after *Prdm9* knockout, maintained (781) – indicating peaks found in both backgrounds, or lost (13,907) – indicating peaks found only in WT. **(B)** Pie chart showing the ratio of four ZCWPW1 peak groups determined by their overlap with two types PRDM9 peaks (Grey e*t al.*; Walker *et al.*). + indicates overlap, - indicates no overlap. **(C)** Boxplots showing the peak intensity of the three ZCWPW1 groups. The Y-axis indicates the log base-2 transformation of normalized signals. P-values were calculated by two-tailed Wilcoxon rank sum test. **(D)** Barplots showing the percentage of ZCWPW1 peaks overlapping with promoter regions for the three groups of ZCWPW1 peaks. (**E)** Barplots showing the constituent ratio of each ZCWPW1 peak group according to whether they overlapped with PRDM9, which includes the merged peaks of Grey *et al* and Walker *et al*. Black indicates overlap, white indicates no overlap. **(F)** Motif analysis of the lost ZCWPW1 peaks that did not overlap with merged PRDM9 peaks. The value of background

had been transformed to the condition that the ratio was stable and the number of regions was scaled to 3028.

Figure 5–figure supplement 1 Correlation between the chromatin occupancy of ZCWPW1 and DMC1

(A) Venn diagram showing the overlap of the DMC1 (Grey *et al*) and Spo11 (Lange *et al.*) peaks with ZCWPW1 binding sites. P-values were calculated using the permTest (see Methods, ntimes=1000). **(B)** Heatmap and profile plot of the ZCWPW1 signal in all DMC1 peaks. Each row in the heatmap represents a DMC1 binding site of ±2 kb around the center and ranked according to ZCWPW1 signal from the highest to the lowest. Color indicates the normalized ChIP-seq signal (See methods). (C) Scatterplot showing the correlation between the normalized ZCWPW1 signal and DMC1 signal on all DMC1 peaks, with spearman correlation coefficient. The x and y axis indicated 2-based log transformation of normalized signal the ZCWPW1 and DMC1 after adding 1. **(D)** Line plots showing the overlap of ZCWPW1 peaks with DMC1 peaks based on different minimum cutoff values for ZCWPW1 peak intensity. The Y-axis indicates the overlap number, and the X-axis indicates the minimum cutoff of ZCWPW1 peak intensity. The red line represents the total number of ZCWPW1 peaks with a specific minimum cutoff for ZCWPW1 peak intensity. The black line represents the overlap number of ZCWPW1 peaks with DMC1 peaks with a specific minimum cutoff for ZCWPW1 peak intensity. **(E)** Pie chart showing the ratio of four DMC1 peak groups (Grey *et al.*) determined by their overlap with merged PRDM9 peaks (Grey e*t al.*; Walker *et al.*) and ZCWPW1 peaks. + indicates overlap, - indicates no overlap. **(F)** The number of region (peak)-gene associations among distance stratifications of ZCWPW1 peaks in three subtypes within the TSS ±1000 kb region.

Supplemental File 1 Summary of all ChIP-seq experiments indicating antibodies, samples, replicates, genotype and data source

Supplemental File 2 ZCWPW1 peaks in WT and *Prdm9* knockout mice

Supplemental File 3 Down-regulated and up-regulated genes in *Zcwpw1*$^{-/-}$ (vs WT) whose promoters overlap with ZCWPW1 peaks

Figure 1-source data 1 Number of DMC1 and RAD51 foci in Figure 1D and 1F

Figure 1 figure supplemental 2–source data 1 Number of synapsed chromosome pairs per cell in Figure 1 figure supplemental 2C

Figure 1 figure supplemental 3-source data 1 Number of MSH4 and RNF212 foci in Figure 1 figure supplemental 3B and 3D

Figure 1

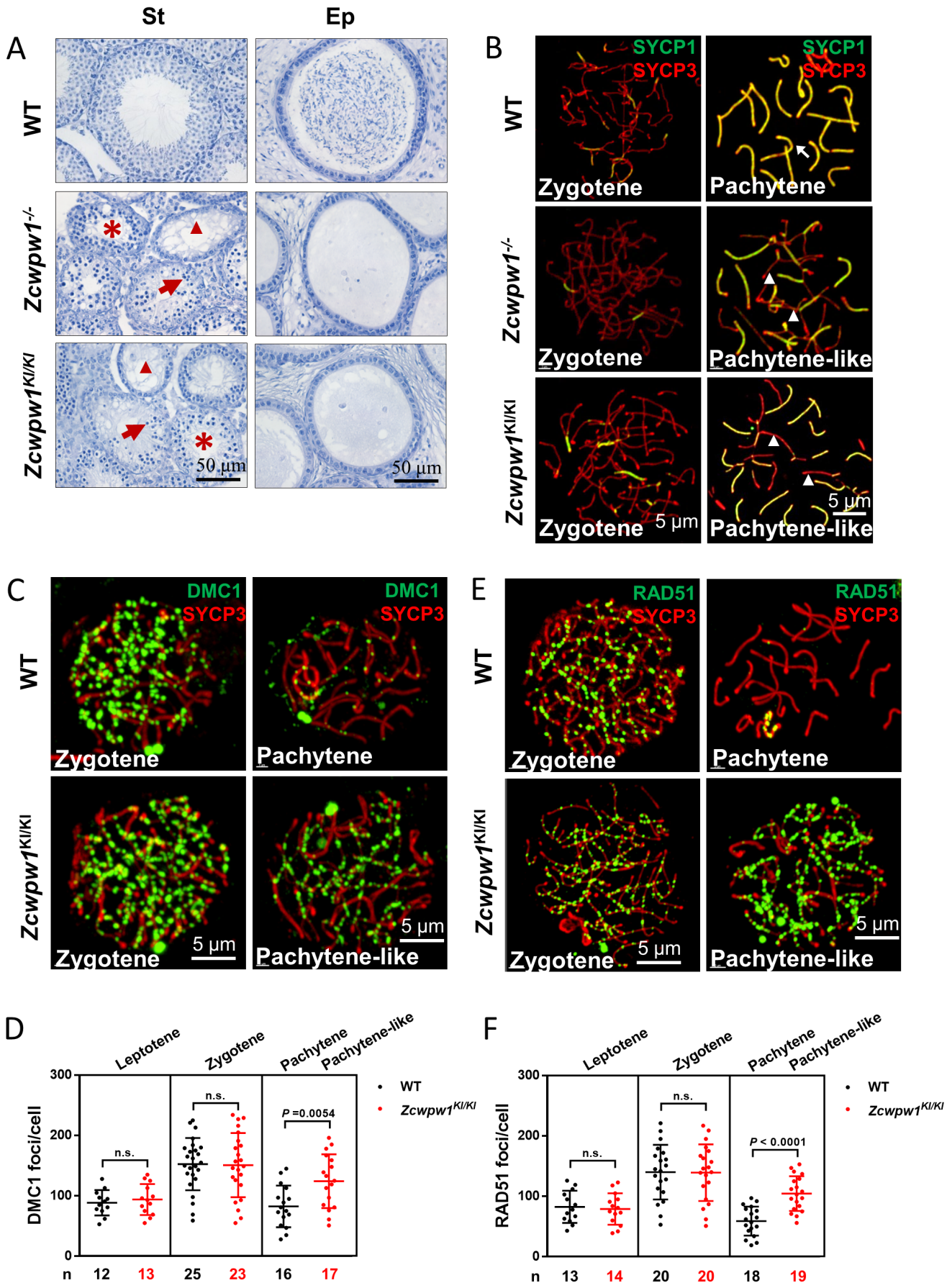

A
St
Ep
WT
Zcwpw1-/-
Zcwpw1KI/KI
50 μm
50 μm
B
SYCP1
SYCP3
Zygotene
Pachytene
Pachytene-like
5 μm
C
DMC1
SYCP3
Zygotene
Pachytene
Pachytene-like
5 μm
D
Leptotene
Zygotene
Pachytene
Pachytene-like
DMC1 foci/cell
n.s.
P =0.0054
WT
Zcwpw1KI/KI
n 12 13 25 23 16 17
E
RAD51
SYCP3
F
RAD51 foci/cell
P < 0.0001
n 13 14 20 20 18 19

Figure 2

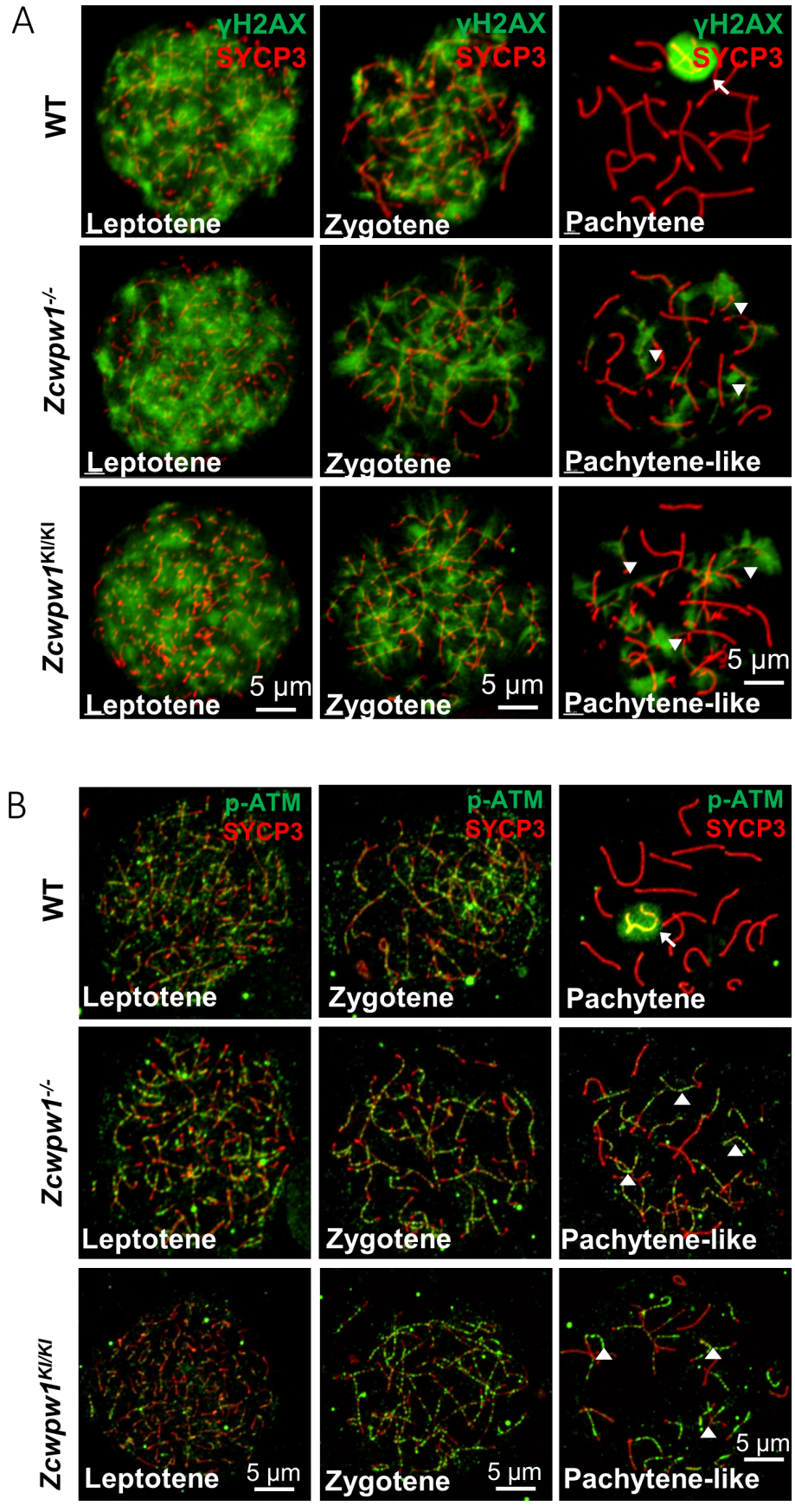

A
γH2AX
SYCP3
WT
Leptotene
Zygotene
Pachytene
Zcwpw1-/-
Leptotene
Zygotene
Pachytene-like
Zcwpw1KI/KI
Leptotene
5 μm
Zygotene
5 μm
Pachytene-like
5 μm
B
p-ATM
SYCP3
WT
Leptotene
Zygotene
Pachytene
Zcwpw1-/-
Leptotene
Zygotene
Pachytene-like
Zcwpw1KI/KI
Leptotene
5 μm
Zygotene
5 μm
Pachytene-like
5 μm

Figure 3

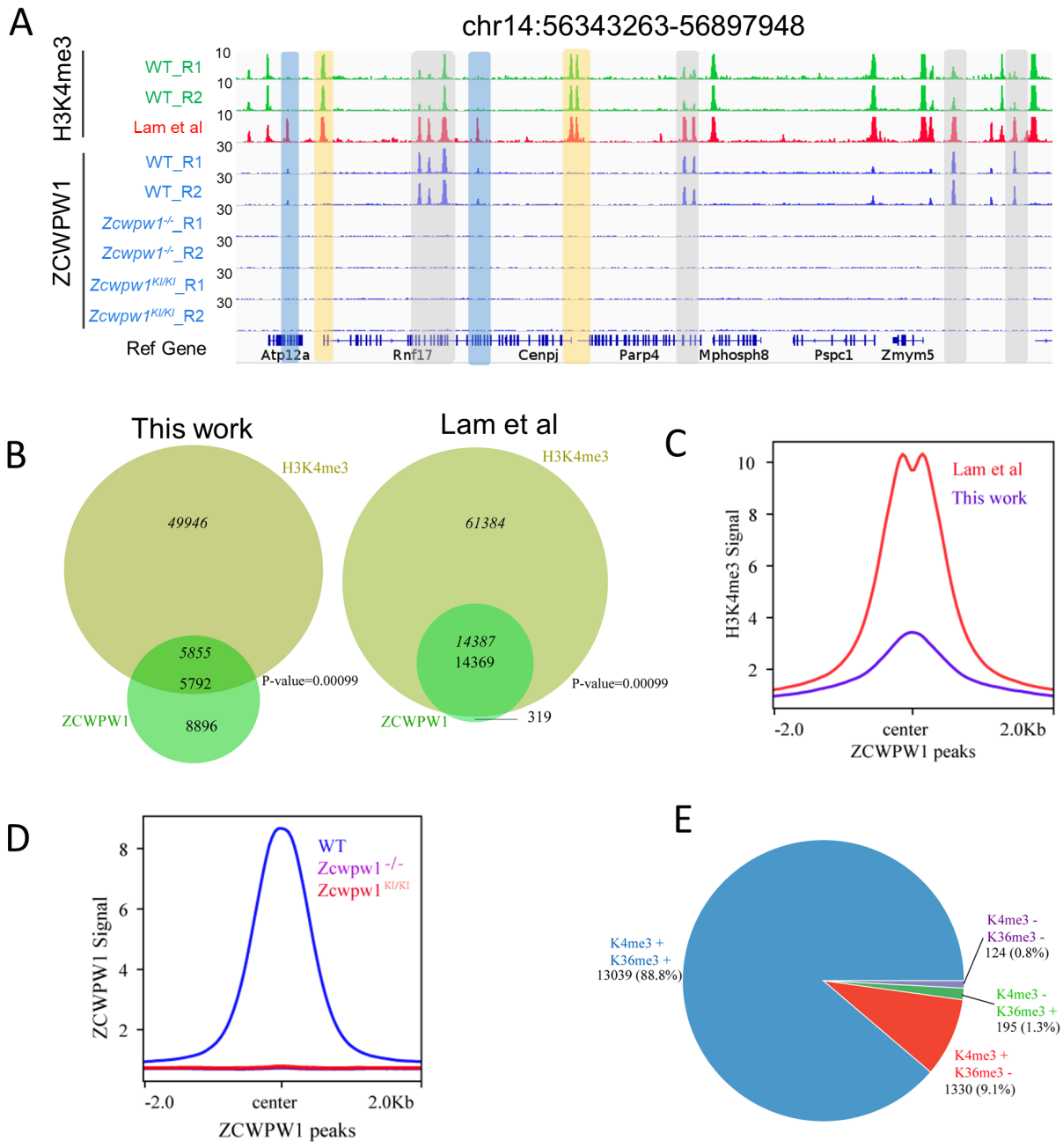

A
chr14:56343263-56897948
H3K4me3
WT_R1
WT_R2
Lam et al
ZCWPW1
WT_R1
WT_R2
Zcwpw1-/-_R1
Zcwpw1-/-_R2
Zcwpw1KI/KI_R1
Zcwpw1KI/KI_R2
Ref Gene
Atp12a
Rnf17
Cenpj
Parp4
Mphosph8
Pspc1
Zmym5
B
This work
H3K4me3
49946
5855
5792
8896
ZCWPW1
P-value=0.00099
Lam et al
H3K4me3
61384
14387
14369
319
ZCWPW1
P-value=0.00099
C
H3K4me3 Signal
Lam et al
This work
-2.0
center
2.0Kb
ZCWPW1 peaks
D
ZCWPW1 Signal
WT
Zcwpw1 -/-
Zcwpw1 KI/KI
-2.0
center
2.0Kb
ZCWPW1 peaks
E
K4me3 +
K36me3 +
13039 (88.8%)
K4me3 -
K36me3 -
124 (0.8%)
K4me3 -
K36me3 +
195 (1.3%)
K4me3 +
K36me3 -
1330 (9.1%)

Figure 4

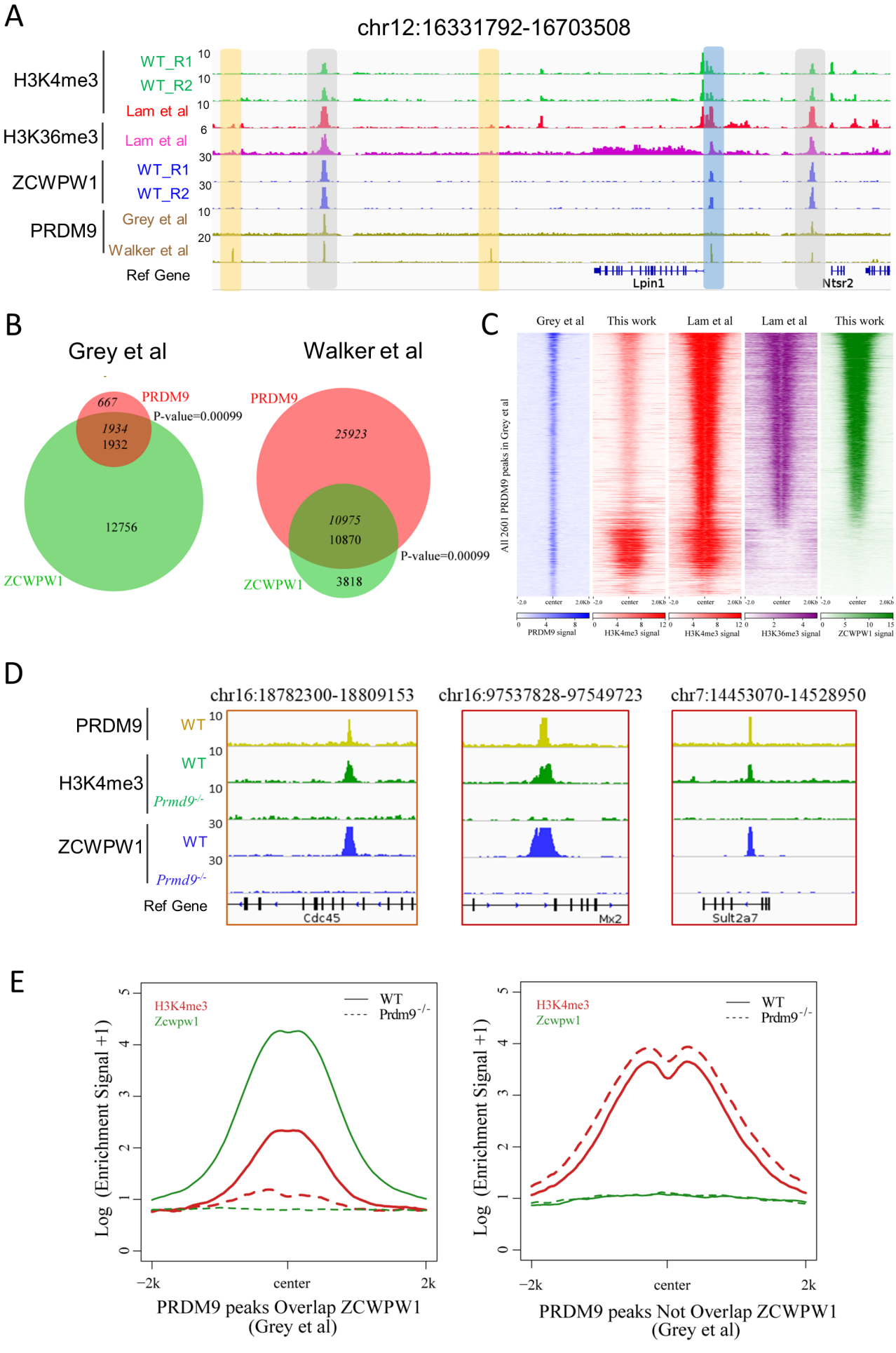

A
chr12:16331792-16703508
H3K4me3
WT_R1
WT_R2
Lam et al
H3K36me3
Lam et al
ZCWPW1
WT_R1
WT_R2
PRDM9
Grey et al
Walker et al
Ref Gene
Lpin1
Ntsr2
B
Grey et al
Walker et al
PRDM9
667
1934
1932
12756
ZCWPW1
P-value=0.00099
PRDM9
25923
10975
10870
3818
ZCWPW1
P-value=0.00099
C
Grey et al
This work
Lam et al
Lam et al
This work
All 2601 PRDM9 peaks in Grey et al
PRDM9 signal
H3K4me3 signal
H3K4me3 signal
H3K36me3 signal
ZCWPW1 signal
D
chr16:18782300-18809153
chr16:97537828-97549723
chr7:14453070-14528950
PRDM9
WT
H3K4me3
WT
Prmd9-/-
ZCWPW1
WT
Prmd9-/-
Ref Gene
Cdc45
Mx2
Sult2a7
E
H3K4me3
Zcwpw1
WT
Prdm9-/-
Log (Enrichment Signal +1)
−2k
center
2k
PRDM9 peaks Overlap ZCWPW1
(Grey et al)
PRDM9 peaks Not Overlap ZCWPW1
(Grey et al)

Figure 5

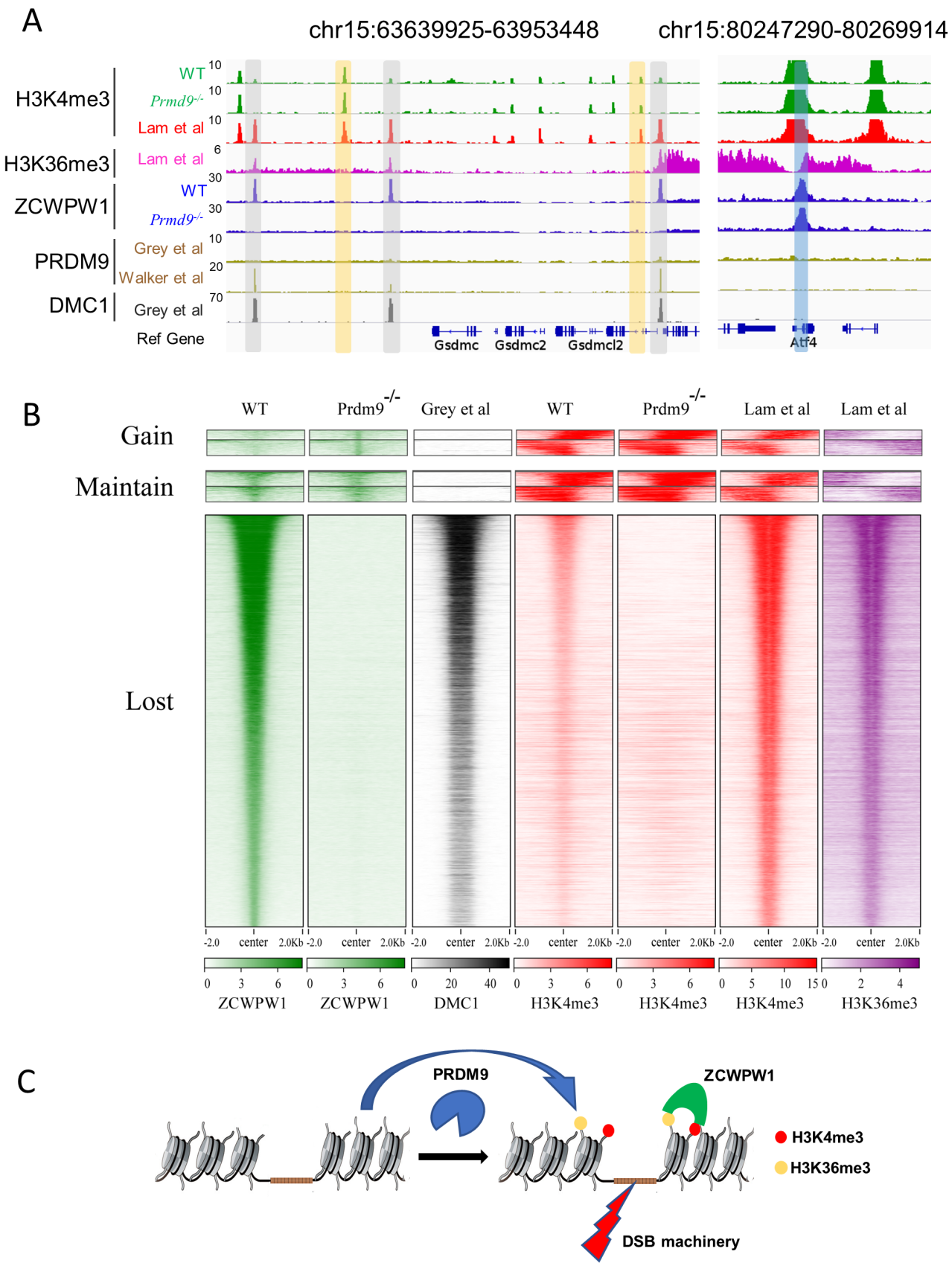

A
chr15:63639925-63953448
chr15:80247290-80269914
H3K4me3
WT
Prmd9-/-
Lam et al
H3K36me3
Lam et al
ZCWPW1
WT
Prmd9-/-
PRDM9
Grey et al
Walker et al
DMC1
Grey et al
Ref Gene
Gsdmc
Gsdmc2
Gsdmcl2
Atf4
B
WT
Prdm9-/-
Grey et al
WT
Prdm9-/-
Lam et al
Lam et al
Gain
Maintain
Lost
-2.0
center
2.0Kb
ZCWPW1
ZCWPW1
DMC1
H3K4me3
H3K4me3
H3K4me3
H3K36me3
C
PRDM9
ZCWPW1
H3K4me3
H3K36me3
DSB machinery

Figure 1–figure supplement 1

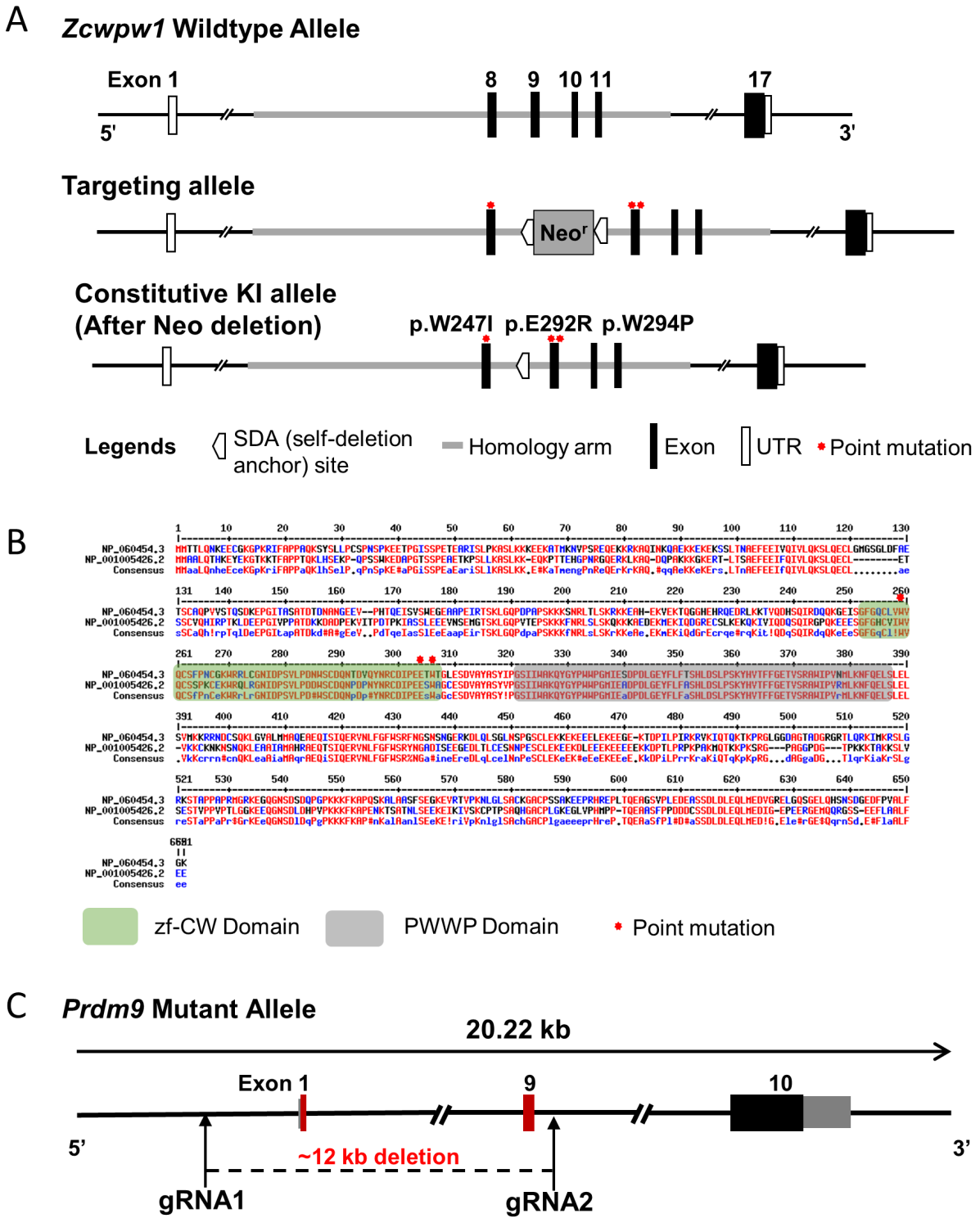


Figure 1–figure supplement 2

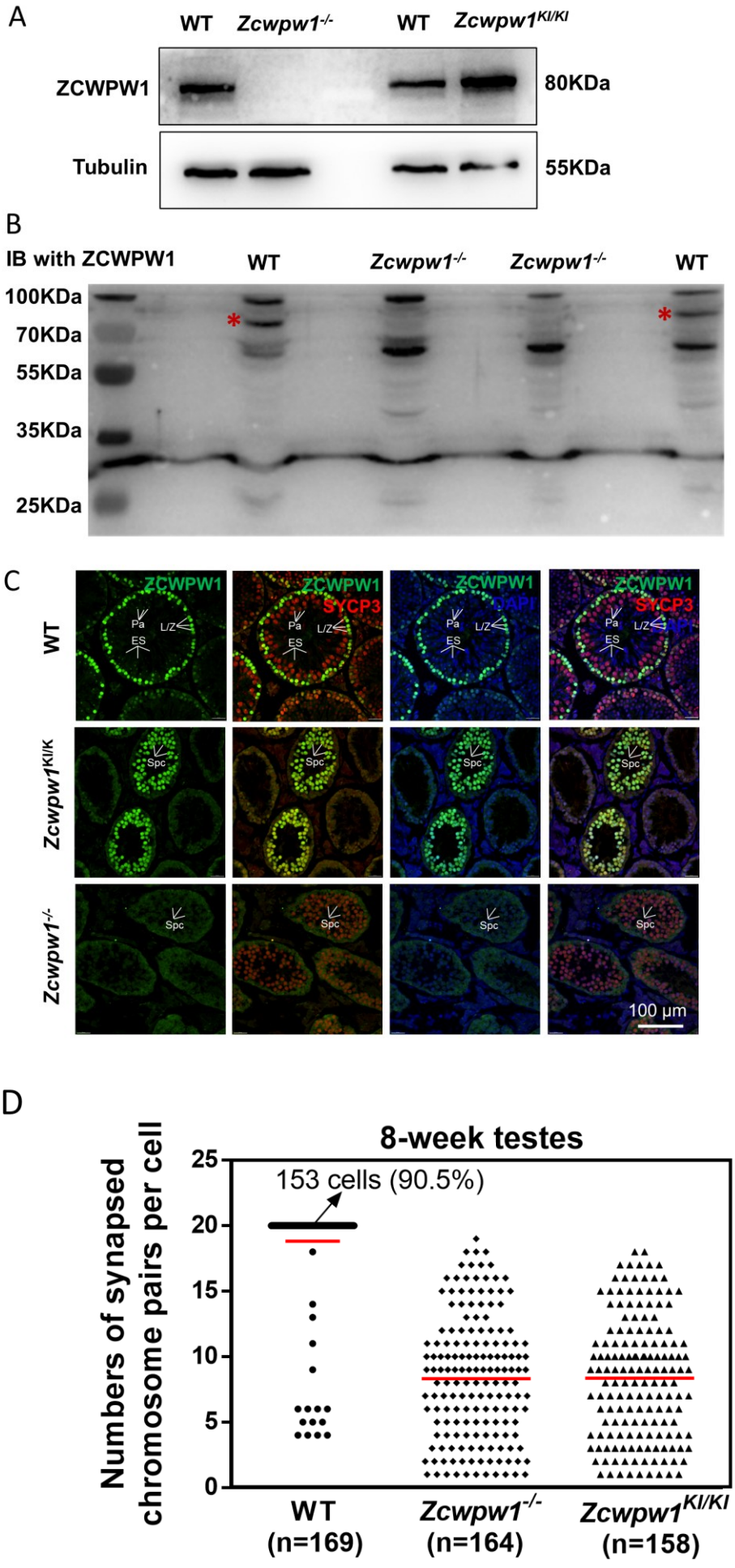


Figure 1–figure supplement 3

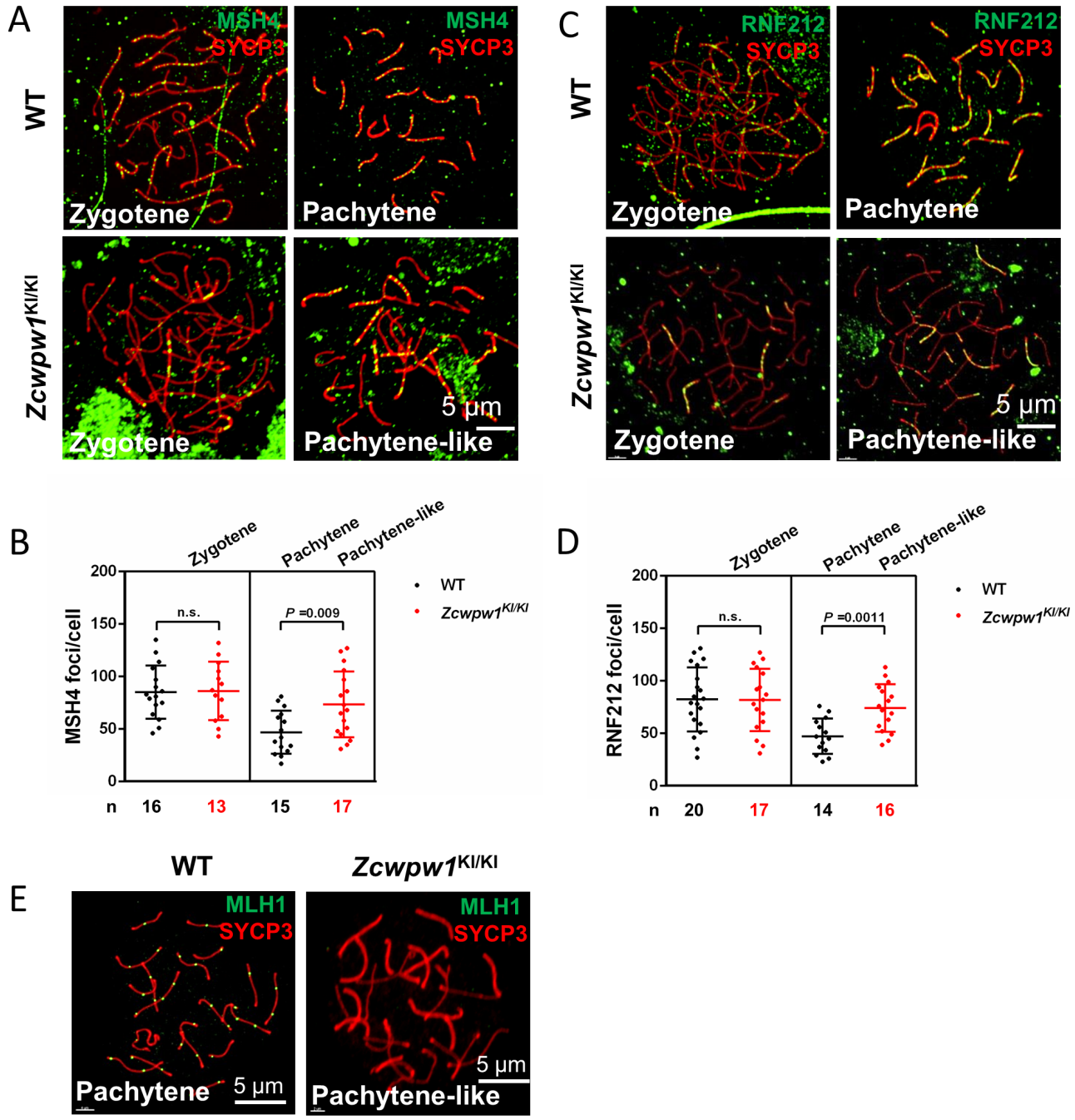


Figure 3–figure supplement 1

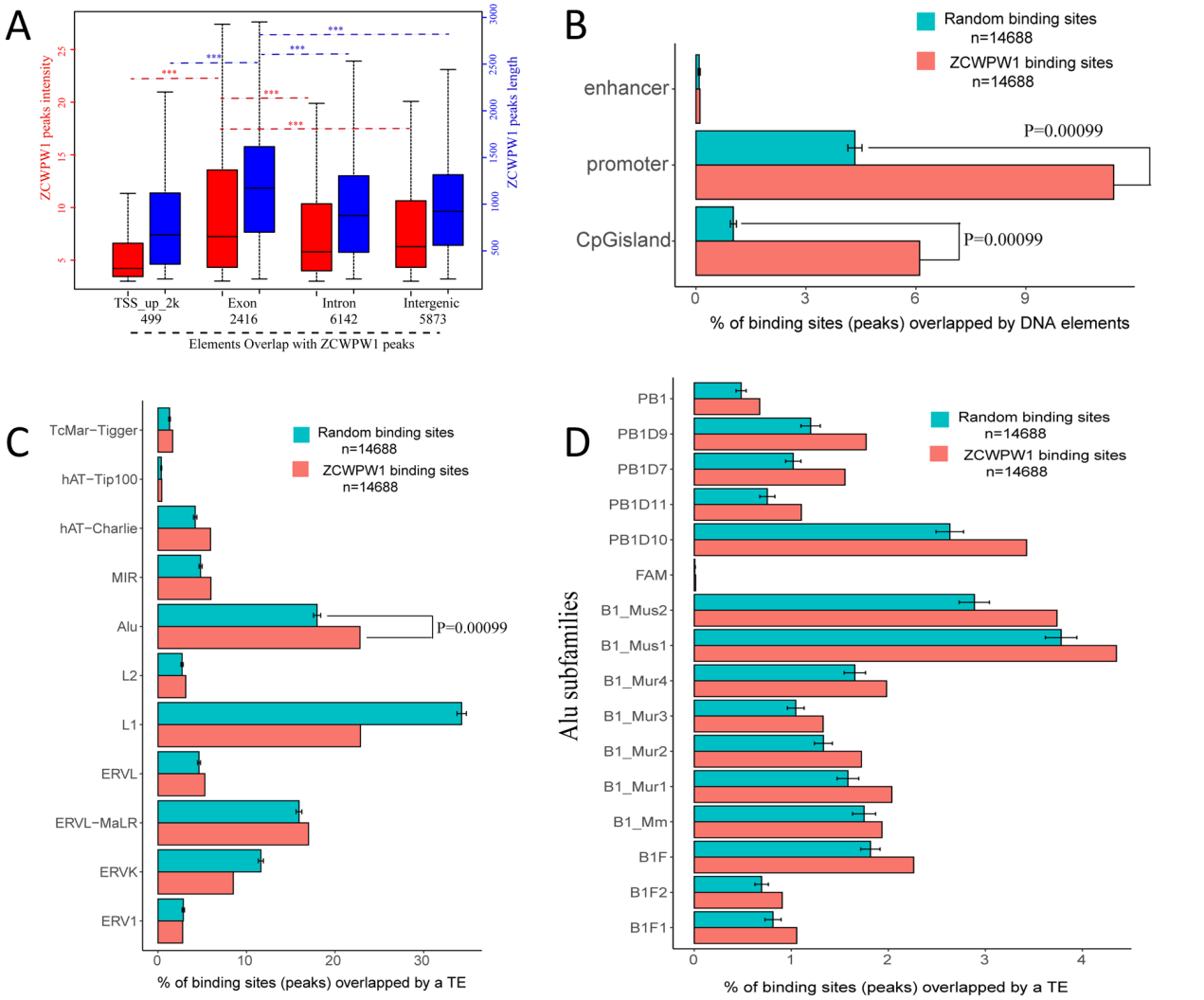


Figure 3–figure supplement 2

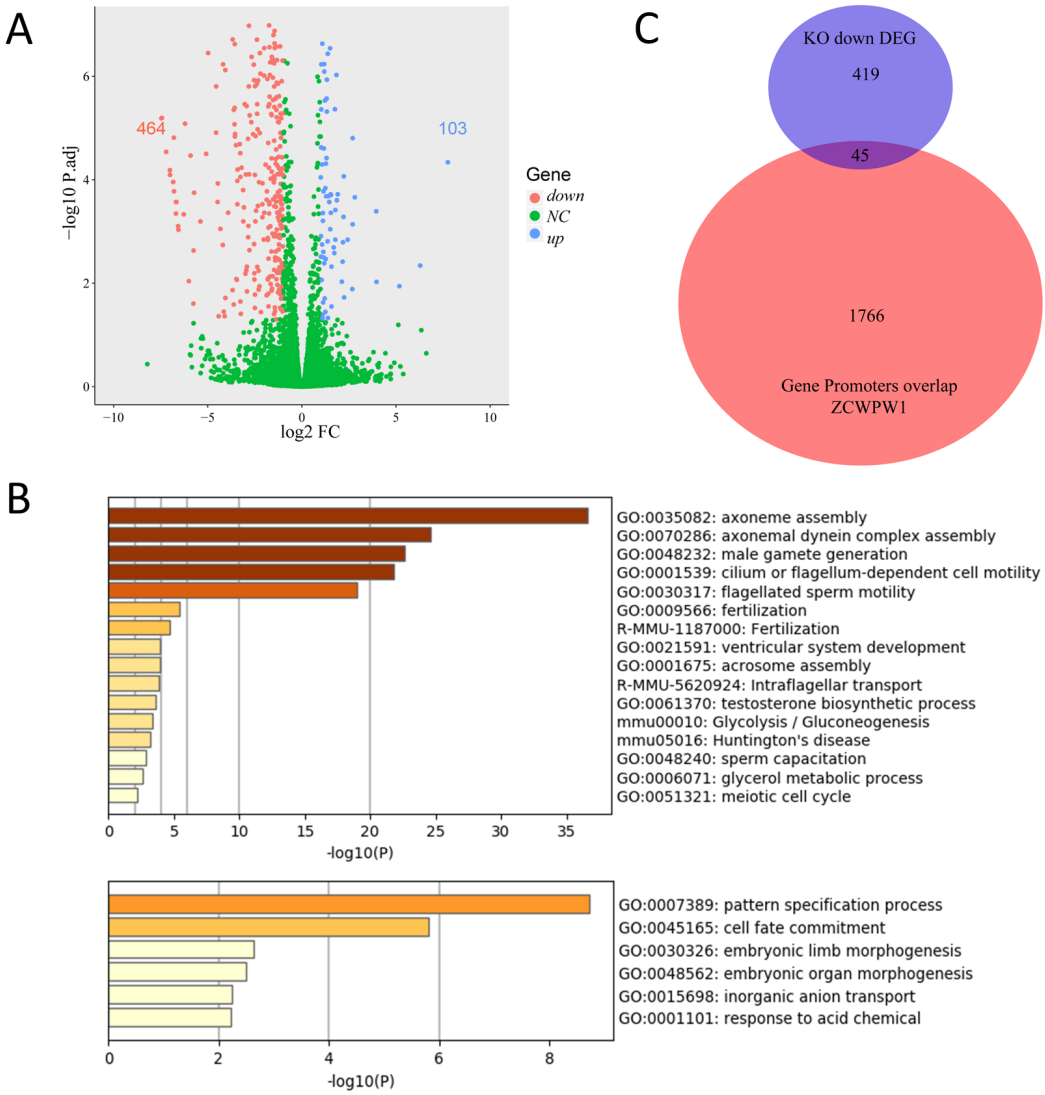


Figure 3–figure supplement 3

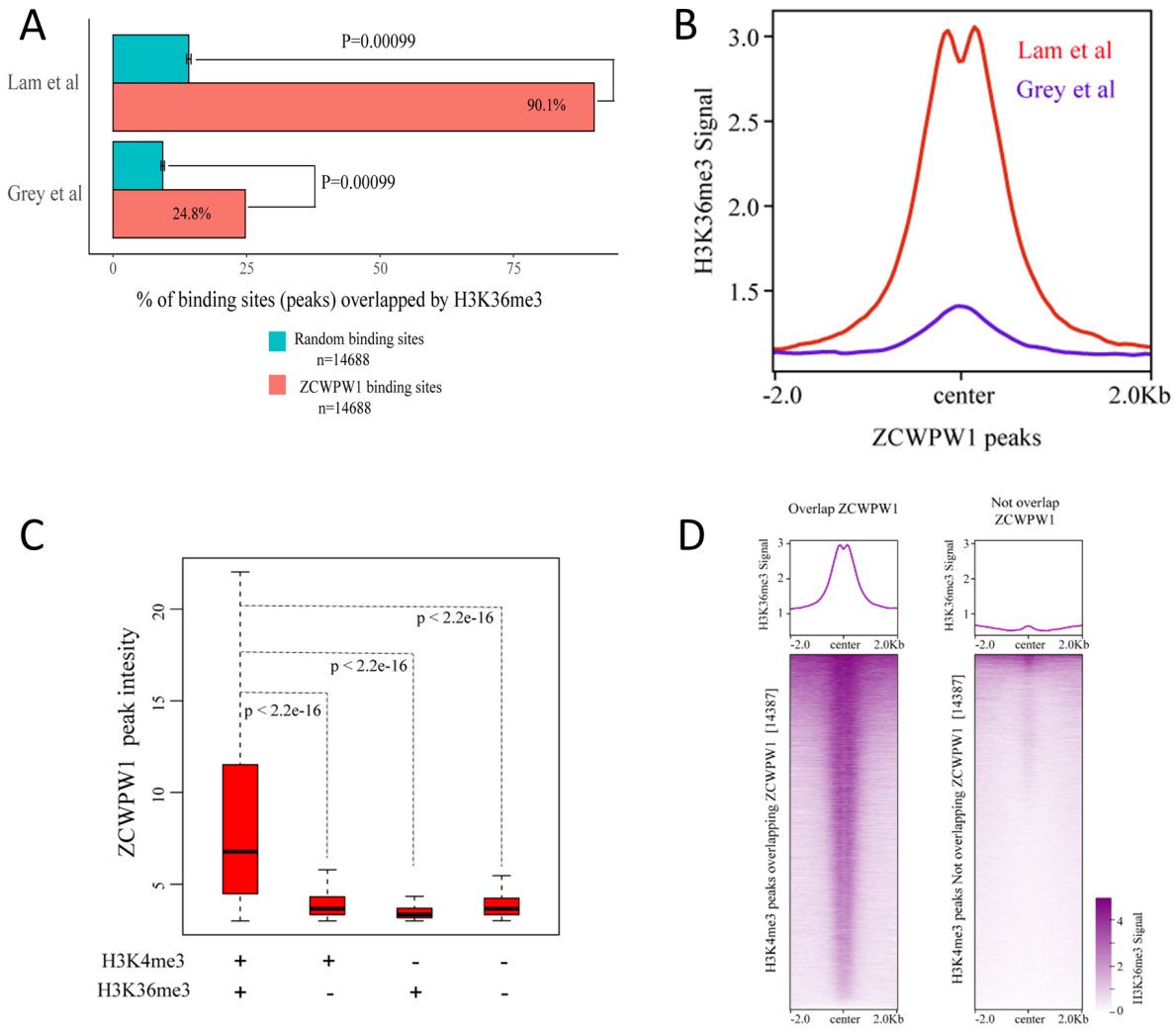


Figure 3–figure supplement 4

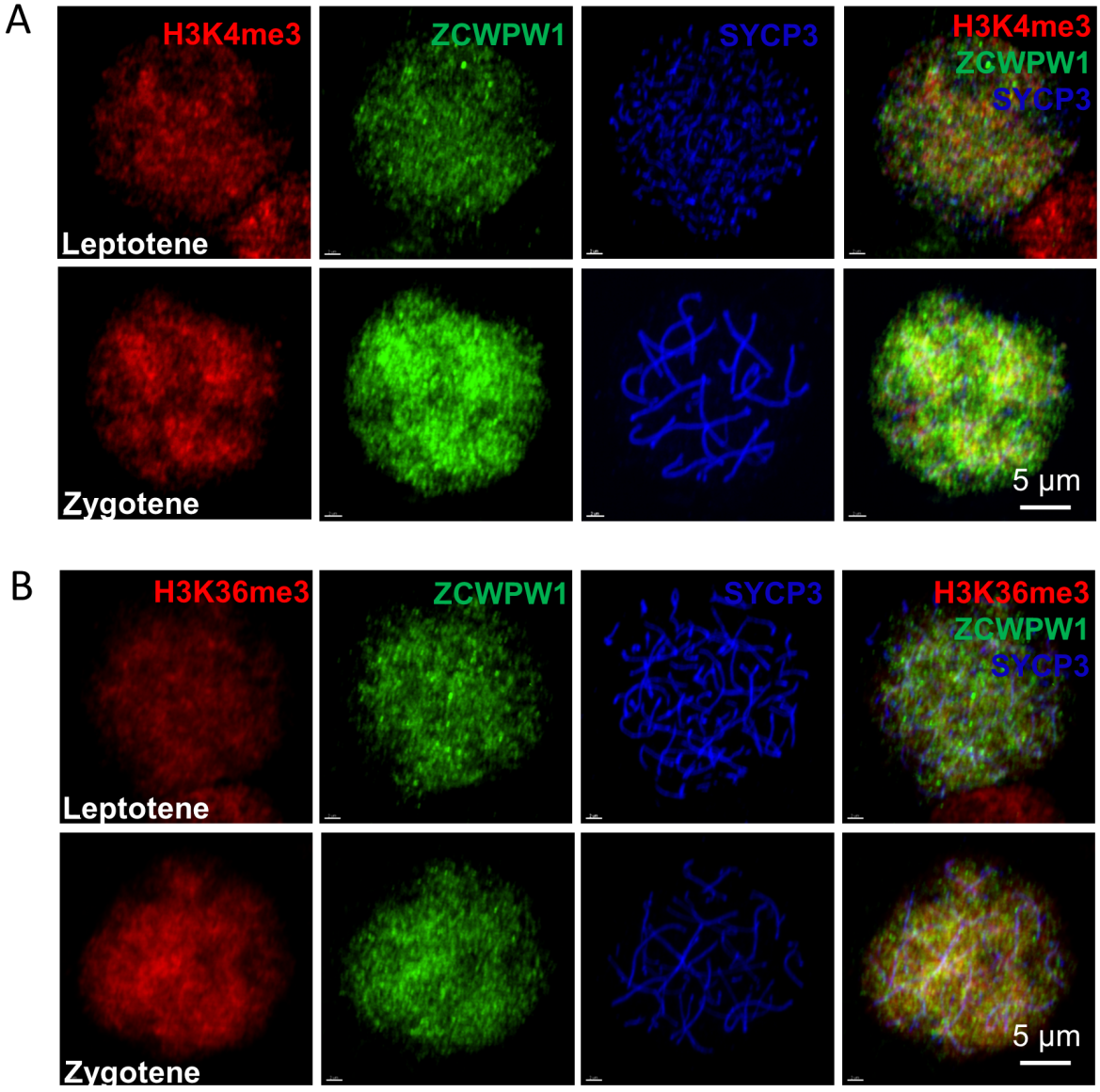


Figure 4–figure supplement 1

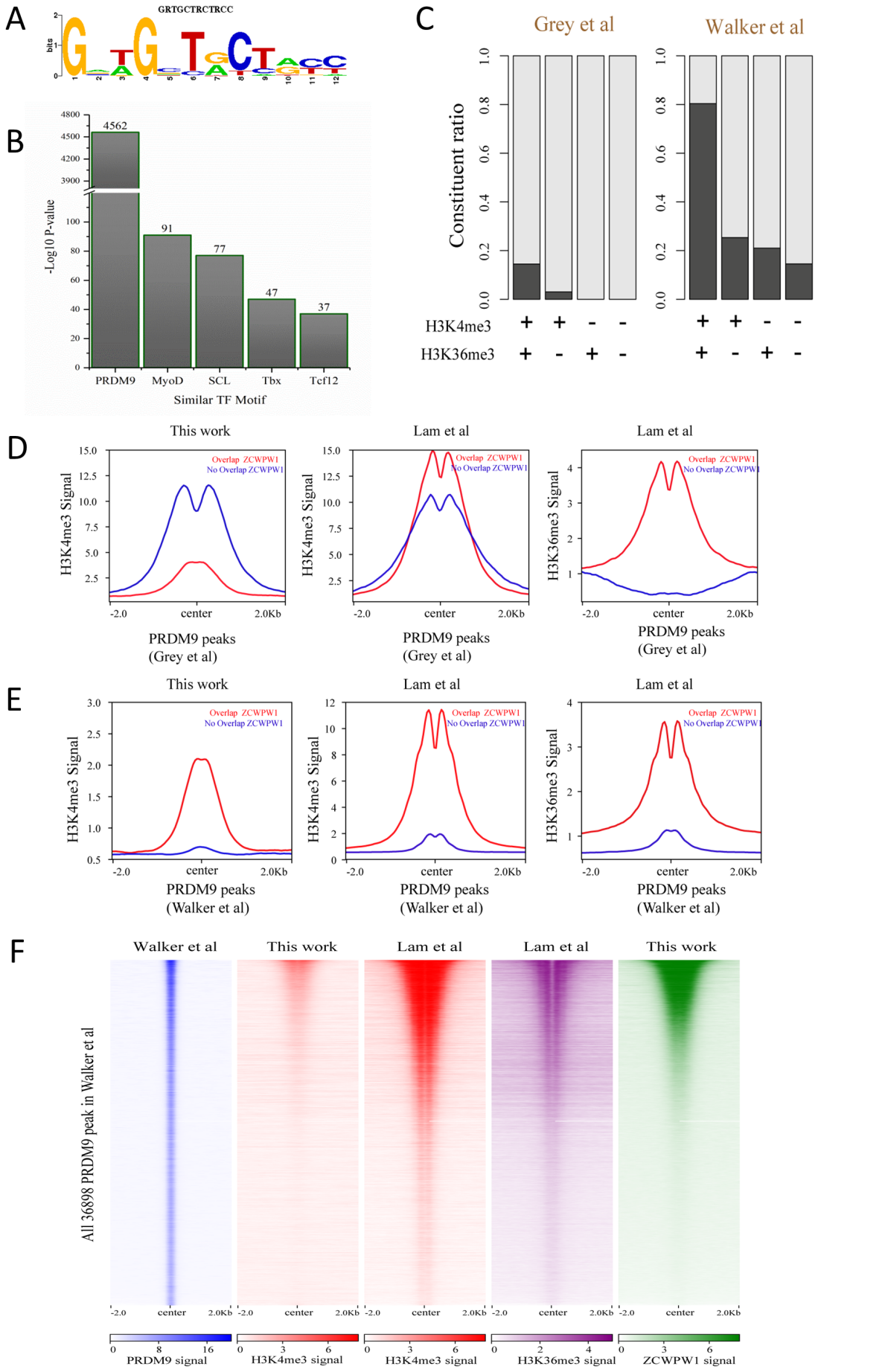


Figure 4–figure supplement 2

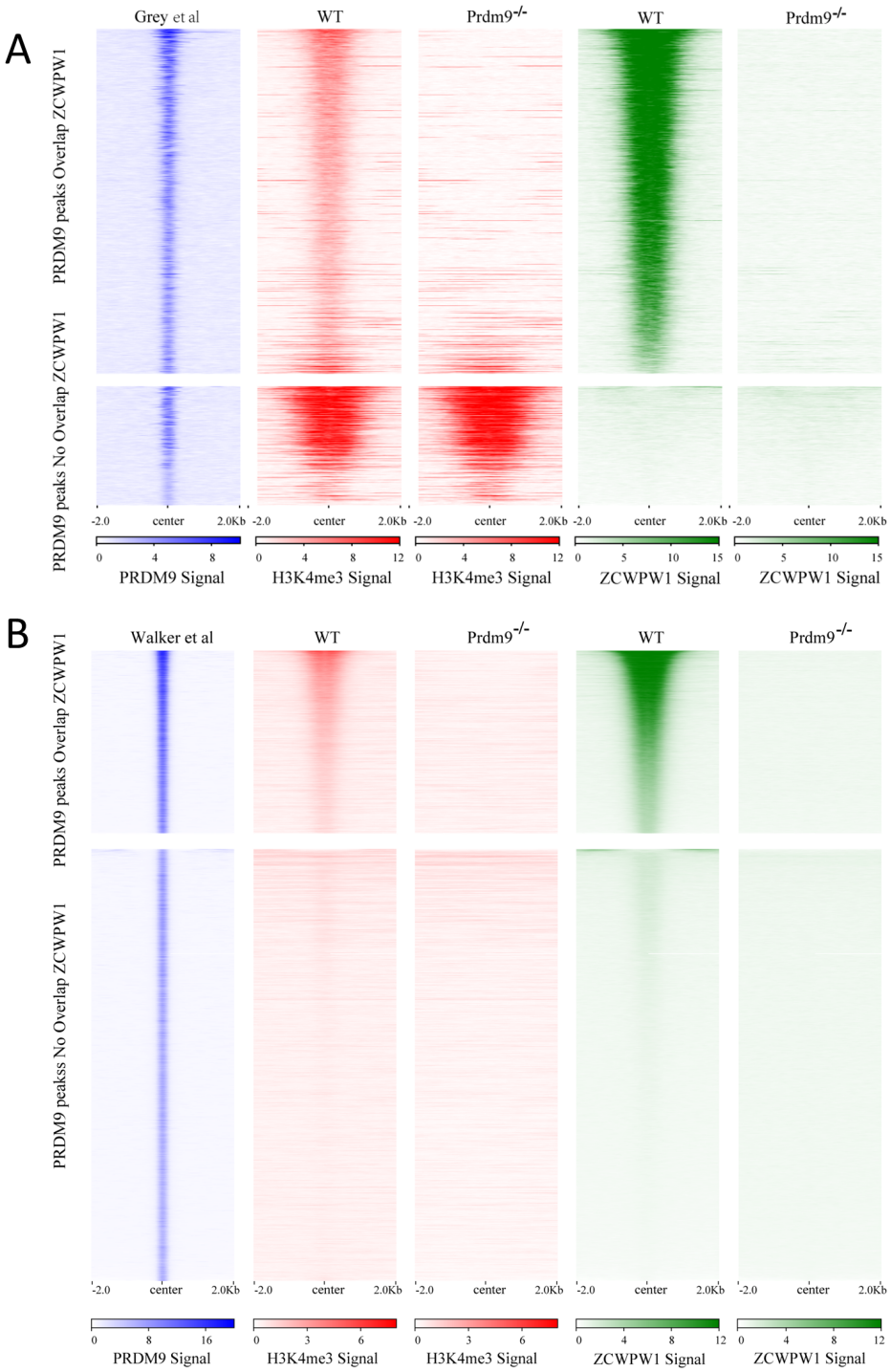
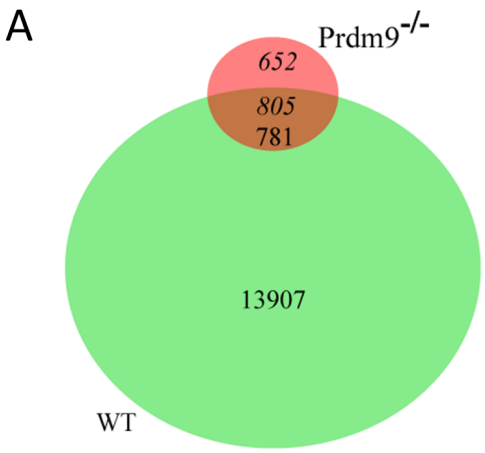
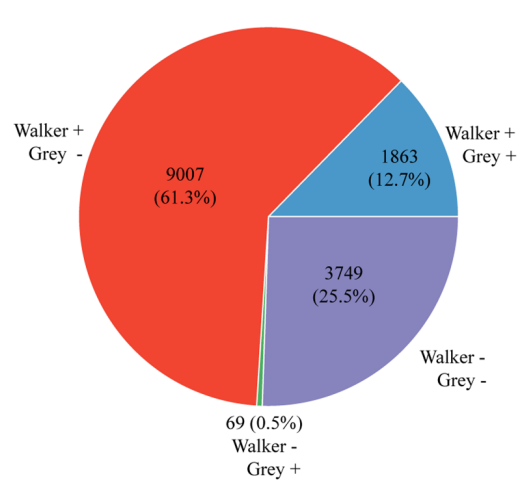
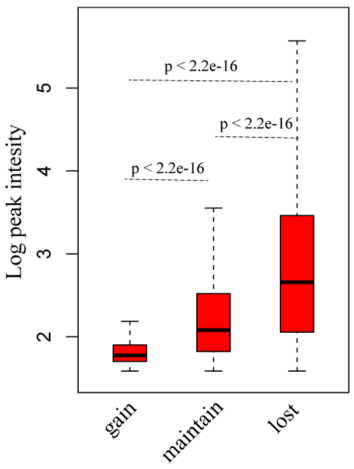
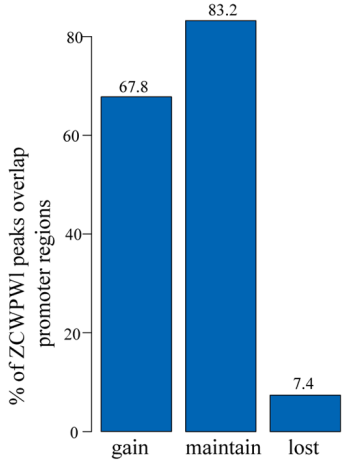


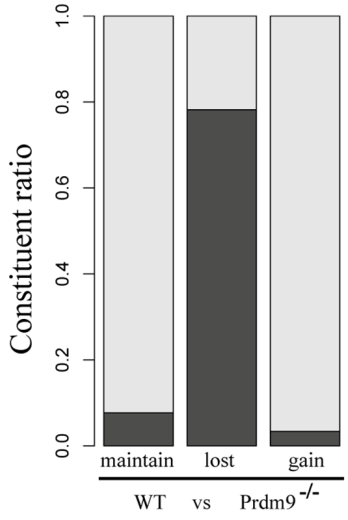
Figure 4–figure supplement 3











F

3028 ZCWPW1 peaks: Lost but not overlap with PRDM9

	Name	P-value	Target Sequences with Motif	Background Sequences with Motif
	PRDM9	1e-606	994(32.8%)	117(3.88%)

Figure 5–figure supplement 1