PRDM9 forms a multiprotein complex tethering recombination hotspots to the chromosomal axis

Emil D. Parvanov[1,2], Hui Tian[1], Timothy Billings[1], Ruth L. Saxl[1], Rakesh Aithal[3], Lumir Krejci[2,3*], Kenneth Paigen[1], Petko M. Petkov[1*]

[1]Center for Genome Dynamics, The Jackson Laboratory, Bar Harbor, ME, USA, 04609

[2]Department of Biology, Masaryk University, Brno, Czech Republic

[3]National Centre for Biomolecular Research, Masaryk University, Brno, Czech Republic 625 00

[*]Corresponding authors: petko.petkov@jax.org and lkrejci@chemi.muni.cz

Correspondence:

Petko M. Petkov, Ph.D.

Email petko.petkov@jax.org

Running title: PRDM9 brings hotspots to axis

Keywords: meiosis/multiprotein complex/PRDM9/recombination/synaptonemal complex

SUMMARY STATEMENT

This paper shows how the recombination regulator PRDM9 brings the hotspot DNA to the chromosomal axis to facilitate meiotic recombination initiation

ABSTRACT

In mammals, meiotic recombination occurs at 1-2 kb genomic regions termed hotspots, whose positions and activities are determined by PRDM9, a DNA-binding histone methyltransferase. We now show that the KRAB domain of PRDM9 binds additional proteins into complexes that bring hotspots into the next phase of recombination. By a combination of yeast-two hybrid assay, *in vitro* binding, and co-immunoprecipitation from mouse spermatocytes, we identified four proteins that directly interact with PRDM9's KRAB domain, CXXC1, EWSR1, EHMT2, and CDYL. These proteins are co-expressed in spermatocytes at the early stages of meiotic prophase I to which PRDM9 expression is restricted. We also detected association of PRDM9-bound complexes with the meiotic cohesin REC8 and the synaptonemal complex proteins SYCP3 and SYCP1. Our results suggest a model in which PRDM9-bound hotspot DNA is brought to the chromosomal axis by the action of these proteins, which also ensure proper chromatin and spatial environment for the subsequent recombination events. Double-strand breaks are initiated on the chromosomal axis; homology search and their subsequent repair are restricted to the synaptonemal complex space.

INTRODUCTION

Genetic recombination assures the proper segregation of homologous chromosomes at the first meiotic division, preventing aneuploidy. It also plays an important evolutionary role by facilitating the creation of new, favorable combinations of alleles and the removal of deleterious mutations by unlinking them from surrounding sequences. In mammals, as in yeast, higher plants, and birds, recombination occurs at specialized sites along chromosomes known as hotspots (Baudat et al., 2013; Paigen and Petkov, 2010), typically a kilobase or so in length, separated by tens to hundreds of kilobases that lack recombination. Hotspot locations and relative activity determine the recombination landscape.

Mammals are distinct from other organisms exhibiting hotspots in that a meiosis-specific protein, PRDM9, first identified by our group and others (Baudat et al., 2010; Myers et al., 2010; Parvanov et al., 2010), is the primary determinant of recombination hotspot locations (Baudat et al., 2010; Hinch et al., 2011; Myers et al., 2010; Parvanov et al., 2010; Smagulova et al., 2011). PRDM9 combines domains from two large families of proteins - KRAB- zinc finger (ZnF) (Lupo et al., 2013) and PR domain proteins (Fumasoni et al., 2007) and is the only protein known to contain all three characteristic domains of these families – a KRAB domain implicated in protein-protein interactions, a PR/SET domain with a histone methylation activity, and a ZnF domain for DNA recognition and binding. Recombination begins when the C-terminal ZnF domain of PRDM9 recognizes and binds to hotspot specific DNA sequence (Baudat et al., 2010; Billings et al., 2013; Grey et al., 2011; Myers et al., 2010; Parvanov et al., 2010). The PR/SET domain then locally trimethylates histone H3 on lysine 4 (H3K4me3), which results in rearrangement of the local nucleosome pattern, creating a central nucleosome-depleted region (Baker et al., 2014) where the double strand breaks (DSBs) required for the exchange of DNA sequences between homologous chromatids occur (Brick et al., 2012). The extent of trimethylation of local nucleosomes delimits the span over which subsequent migration of the Holliday junction, a cross-strand structure that forms as a key intermediate in genetic recombination, can occur,

thereby setting the limits of where the final genetic crossovers can take place (Baker et al., 2014). In the absence of PRDM9, DSB are formed at otherwise available H3K4me3 sites, mainly promoters, but they cannot be repaired properly and the germ cells undergo apoptosis (Smagulova et al., 2011). One possible explanation of this observation is that PRDM9 is involved in protein-protein interactions necessary for correct positioning and repair of future DSBs.

We now show that the PRDM9 KRAB domain plays a crucial role in binding and recruiting additional proteins into multi-protein complex that brings hotspots into the next phase of recombination; these proteins include CXXC1, a DNA-binding protein with a CXXC domain found in CpG-binding proteins (Illingworth et al., 2010), EWSR1, which binds single stranded RNA and DNA (Fisher, 2014; Li et al., 2007; Oakland et al., 2013), EHMT2, a histone methyltransferase catalyzing formation of H3K9me1,2 and H3K56me1 (Tachibana et al., 2007; Tachibana et al., 2005), and CDYL, a methyl reader of H3K9me2 and H3K9me3 (Escamilla-Del-Arenal et al., 2013) and H3K27me3 (Zhang et al., 2011), with a role in chromatin silencing, as well as a putative histone acetylase (Lahn et al., 2002).

A combination of yeast two-hybrid screens and *in vitro* pull-downs of purified proteins showed that the four proteins directly bind to the N-terminal region of PRDM9 where the KRAB domain is located. We further confirmed the *in vivo* interaction between PRDM9 and EWSR1, EHMT2, and CDYL in early prophase spermatocytes by co-immunoprecipitation (co-IP), and cytological colocalization. In spermatocytes, we also detected association of PRDM9-bound complexes with meiotic cohesins such as REC8 and chromosomal axis/synaptonemal complex proteins SYCP3 and SYCP1.

These results suggest that PRDM9-bound hotspot DNA is brought to the chromosomal axis by interaction with other proteins serving as a link between PRDM9 and cohesins/SC proteins, thereby assuring proper spatial environment for DSB initiation and their subsequent repair.

RESULTS

Yeast two-hybrid assay identifies direct PRDM9 interactors

To search for proteins directly interacting with PRDM9, we performed a yeast two-hybrid (Y2H) screen using cloned full-length *Prdm9* as bait and a 6-month old mouse testis cDNA library as prey. Screening ~6x10^6 colonies, we isolated a total of 329 positive clones, which after sequencing coalesced to 118 individual open reading frames. Four clones, representing C-terminal portions of *Ehmt2* (amino acids 376-1263), *Cxxc1* (amino acids 217-660), *Ewsr1* (amino acids 479-655), and *Cdyl* (amino acids 74-593) genes, were confirmed as interacting strongly with full-length PRDM9 by pairwise Y2H under the most discriminating conditions (Fig. 1). To identify the positions of their binding sites on the PRDM9 molecule, shorter PRDM9 fragments were cloned as bait constructs and tested by pairwise Y2H with each interacting clone (Fig. 1A and B). All four clones were found to interact with PRDM9 fragments representing the isolated KRAB and PR/SET domains, and the *Ehmt2* and *Ewsr1* clones also interacted with the intervening region between the KRAB and PR/SET domains as well (Fig. 1B). In turn, the sequence of the clones showing positive interactions with PRDM9 can give us clues regarding the binding sites of these proteins. The only known functional domain in the cloned 479-655 amino acid fragment of EWSR1 is a C-terminal RanBP2-type zinc finger, which has been implicated in protein binding (Steggerda and Paschal, 2002). The cloned portion of CXXC1 lacks its N-terminal PHD1 and CXXC-type zinc fingers. The remaining C-terminal part involved in PRDM9 binding includes acidic, basic, coiled coil domains, a Set1-interacting domain, and a PHD2 domain (Tate et al., 2009). The known functional domains of EHMT2 and CDYL are all present in their respective cloned fragments.

EWSR1, EHMT2, CDYL and CXXC1 directly bind with PRDM9 *in vitro*

Each of the interacting proteins detected by Y2H also bound to PRDM9 *in vitro*. For these tests, we cloned, expressed and purified N-terminally tagged versions of these proteins from *E.*

coli. Full-length PRDM9 was cloned as a fused construct with an N-terminal MBP tag. The other four were cloned separately as fused constructs with HALO and GST tags. The tagged proteins were expressed in *E. coli* and purified by three-step procedure including ion exchange and affinity chromatography. All four proteins bound PRDM9 *in vitro* (Fig. 1C). Interestingly, one of the two histone modifiers, EHMT2, showed relatively weaker binding to PRDM9 (Fig. 1C, second row) compared to EWSR1, CDYL and CXXC1.

Binding of PRDM9 and its interacting proteins in spermatocytes

We next tested whether these proteins also interact with PRDM9 and with each other in mouse spermatocyte lysates by co-IP in germ cells of 14-dpp (days post-partum) juvenile mice, which are enriched for the leptotene through early pachytene stages of meiotic prophase I (Fig. 2A, left panel). The immunoprecipitate brought down with antibodies against PRDM9 also contained EWSR1, EHMT2, and CDYL as interactors. *In vivo* binding of PRDM9 to EWSR1 was stronger than with the two histone modifiers, EHMT2 and CDYL. We could not test CXXC1 for the lack of specific antibodies. We also carried out reverse co-IP experiments in which antibodies against EWSR1 pulled down PRDM9, EHMT2, and CDYL (Fig. 2A, central panel). Antibodies against CDYL pulled down EHMT2 but failed to show signal with PRDM9 or EWSR1 (Fig. 2A, right panel). Unfortunately, available anti-EHMT2 antibodies proved unsuitable for co-IP.

Caution is required in the interpretation of reverse co-IP data when two proteins are present in very different amounts, which can result from expression differences within germ cells or between Sertoli cells and the leptotene-zygotene stage spermatocytes where PRDM9 is expressed. In such cases only a limited fraction of the more abundant protein is available for interaction with spermatocyte specific proteins present in lesser amounts. This can result in apparent weak or negative interaction when pulling down with the antibody against the dominant protein, but not the reverse. Additionally, binding of antibodies recognizing individual epitopes

might result in interference with protein interaction and reduction or loss of detected signal. As a result, failure to detect interaction in the reverse direction does not necessarily negate the evidence of positive interaction *in vitro* and in one direction *in vivo*.

We conclude that PRDM9 interacts with EHMT2, CDYL and EWSR1 *in vivo* and that the interactions most probably represent direct binding as demonstrated *in vitro*. We also find evidence of strong interaction between EHMT2 and CDYL *in vivo.* Such interaction has been reported previously in HeLa cells (Mulligan et al., 2008) and in mouse ES cells (Escamilla-Del-Arenal et al., 2011). Lacking evidence of direct interaction between EWSR1 and these other proteins (see more in the analysis of mutants below), we cannot yet distinguish whether PRDM9 forms a single complex with all three proteins, and likely CXXC1, or separate complexes with EHMT2 and CDYL on the one hand and with EWSR1 on the other.

Interactions with the synaptonemal complex

Because the DSBs initiated after PRDM9 activates recombination are subsequently formed and repaired on the chromosomal axis and in the synaptonemal complex (SC), we tested whether PRDM9 and its interacting proteins bind to proteins known to be part of, or bound to the chromosomal axis, such as the meiotic-specific cohesin REC8, HORMA domain protein HORMAD2, chromosomal axis/SC lateral element protein SYCP3, and the SC central element protein SYCP1. Co-IP with anti-PRDM9 pulled down REC8, SYCP3, and SYCP1, but not HORMAD2 (Fig. 2B, left panel). Antibodies against EWSR1 also pulled down REC8 and SYCP3 and showed very faint signal with SYCP1 (Fig. 2B, central panel). Co-IP with antibodies against CDYL did not detect interaction with any of these proteins (Fig. 2B, right panel). Again, unfortunately, we were unable to carry out pulldown with antibody to EHMT2.

In reverse, antibodies against the lateral element protein SYCP3, in addition to pulling down the other SC proteins REC8 and SYCP1, pulled down PRDM9 itself, its partners EHMT2 and CDYL, but interestingly, not EWSR1, which may be explained by concentration differences in

the cell types present in 14-dpp testis (see immunofluorescence results below) (Fig. 2C, left panel). Antibodies against the central element protein SYCP1, in addition to pulling down the other chromosomal axis/SC proteins, REC8, and SYCP3, albeit only weakly for the latter, pulled down PRDM9, but showed only a very faint or no signal for EHMT2, CDYL, and EWSR1 (Fig. 2C, right panel). HORMAD2 was not detected in any of these co-IP experiments. Unfortunately, all available commercial antibodies against REC8 were not suitable for co-IP.

Although these data provide strong evidence that PRDM9 and its interacting proteins associate with the SC *in vivo*, we were unable to detect direct interaction between PRDM9 and the chromosomal axis/SC proteins, by Y2H or *in vitro* after expression of the proteins in *E. coli* (data not shown). Consequently, to provide additional confirmation of interactions with the chromosomal axis/SC *in vivo*, we used a HEK293 mammalian cell expression system and showed co-IP of REC8 with PRDM9 when the two were co-expressed in the same cells, but not when they were expressed separately (Fig. 2D). The lack of interaction when these proteins were expressed and purified from *E.coli* could be explained either because interaction requires a presently unknown protein intermediate naturally present in mammalian cells, or because REC8, which is known to be phosphorylated in meiocytes (Fukuda et al., 2012), must be post-translationally modified to interact with PRDM9.

PRDM9 binding to other proteins is not dependent on its own binding to DNA

To investigate whether PRDM9 binding to other proteins is dependent on its zinc finger domain and the ability to bind to DNA, we created a new PRDM9 functional knockout mouse model, B6(Cg)-Prdm9$^{tm3.1Kpgn}$/Kpgn (designated here as *Prdm9$^{tm3.1Kpgn}$*). In this mouse, we placed a mutation creating an alternative splice acceptor site 44 bp inside the 5'-end of exon 12, which codes for the entire zinc finger domain. This acceptor site is used instead of the natural one, resulting in a spliced mRNA containing a frame shift with a stop codon shortly after the splicing site. As a result, the mutant protein contains amino acids 1-381 of PRDM9 but lacks all

of its zinc fingers and the ability to bind DNA at specific sequences (Walker et al., 2015). It is stable and appears as a 48-kDa band on Western blots in testes from both heterozygous and homozygous mutant animals (Fig. S1A). Both male and female mice homozygous for this mutation are sterile. Spermatogonia and spermatocytes were found in mutant testes but no post-meiotic spermatids were observed (Fig. S1B). Cytological staining of spreads of mutant spermatocytes show cells arrested at an aberrant pachytene-like stage, with inappropriate γH2AX staining on autosomes, and asynapsis of homologous chromosomes (Fig S1C), indicating that pairing of homologous chromosomes was impaired. Lacking the zinc finger domain of PRDM9 also led to increased apoptosis during meiosis compared to wild type (Fig S3D, $p < 0.05$). This phenotype is very similar to the one described in *Prdm9*^{TM1Ymat} KO mice which have no PRDM9 protein (Hayashi et al., 2005; Sun et al., 2015).

The truncated 48 kDa version of PRDM9 was immunoprecipitated from testis lysates using anti-PRDM9 antibodies directed against the remaining portion of the molecule (Fig. 3, left panel). These co-IP experiments pulled down all of the proteins that associate with full length PRDM9: EWSR1, EHMT2, CDYL, REC8, SYCP3, and SYCP1, but not HORMAD2, which also does not associate with full length PRDM9 (Fig. 3, central panel). An additional band corresponding to non-phosphorylated SYCP3 (Fukuda et al., 2012) was detected by the pull-down although it was not detectable in the cell lysate (Fig. 3, left panel, asterisk). These results entirely matched the co-IP experiments in wild type testis (Fig. 2A and 2B), confirming that PRDM9 binds to its interactors *in vivo* through its N-terminal portion where the KRAB and PR/SET domains are located. These results also make it clear that PRDM9 binding to other proteins does not depend on its ability to bind DNA.

EWSR1, EHMT2 and CDYL bind to chromosomal axis in the absence of PRDM9

To determine if binding of EWSR1, EHMT2 and CDYL occurs entirely through PRDM9 or whether these proteins can bind to the SC independently, we performed co-IP with anti-SYCP3

antibodies in testes of *Prdm9*^{tm1Ymat} knockout mice (Hayashi et al., 2005) that do not express any PRDM9 protein (Sun et al., 2015). All of the other three proteins – EWSR1, EHMT2, and CDYL – were pulled down by SYCP3 (Fig. 3, central panel), indicating that they can bind to the synaptonemal complex independently of PRDM9. Co-IP with EWSR1 pulled down SYCP3 but showed only very weak signals with EHMT2 and CDYL (Fig. 3, right panel). Together, these results show that EWSR1, EHMT2, and CDYL bind to SC independently of their interaction with PRDM9, and that EWSR1 interacts with EHMT2/CDYL predominantly through PRDM9.

PRDM9 colocalizes with its interactors in mid- to late zygotene

Our previous work has established the presence of PRDM9 in pre-leptonema, leptonema, and early-to mid-zygonema, a period of roughly 48 hours (Sun et al., 2015). To determine whether PRDM9 co-expresses and co-localizes with its interactors, we performed double staining with combinations of antibodies against PRDM9 and its interactors in seminiferous tubules and spermatocyte spreads. We first determined the spatial and temporal co-localization of PRDM9 and EWSR1 in seminiferous tubules of 14-dpp juvenile mice (Fig. 4A, top panel, and Fig. S2A, top panel). EWSR1 showed high expression in spermatogonia (Fig. 4A, Fig. S2A and B, top panel, arrowhead) and Sertoli cells (marked by GATA4, Fig. S2B, arrow) located at the base of seminiferous tubule, and remained present in the nuclei of spermatocytes located in the tubule lumen (Fig 4A, top panel, and Fig. S2A, top panel, arrow). At this stage, when most spermatocytes are in leptonema and zygonema (Sun et al., 2015), PRDM9 and EWSR1 are clearly co-expressed in those cells (Fig. 4A, top panel, and Fig. S2A, top panel, arrow). Because the EWSR1 signal is weaker in PRDM9-positive cells, we sought to confirm that these proteins co-localize in spermatocyte spreads of 14-dpp mice by performing triple staining with EWSR1, PRDM9, and SYCP3, using the chromosomal axis protein SYCP3 as a marker of meiotic progression (Fig. 4B). PRDM9 and EWSR1 were clearly co-expressed in pre-leptonema to mid-zygonema nuclei showing mostly diffuse and punctate appearance. The EWSR1 signal

increased dramatically in pachynema when PRDM9 has disappeared from the germ cell nuclei. Interestingly, the EWSR1 signal was excluded from the sex body in pachynema as demonstrated by its lack of co-localization with BRCA1 (Fig. 4B, yellow arrow).

Similar EWSR1-PRDM9 co-localization pattern is seen in seminiferous tubules of *Prdm9*$^{tm3.1Kpgn}$ mice lacking the PRDM9 zinc finger domain (Fig. 4A, second row, and Fig. S2A, second row). As noted above, this mutant undergoes meiotic arrest around the zygonema-pachynema transition and the seminiferous tubule lumens contains very few of the EWSR1-positive, PRDM9-negative germ cell nuclei characteristic for pachynema. A very similar EWSR1 staining is found in *Prdm9*^{tm1Ymat} KO mice which do not express any PRDM9 protein (Fig. 4A, third row, and Fig. S2A, third row).

We conclude that EWSR1 and PRDM9 are co-expressed in early meiotic prophase for the entire time of PRDM9 expression.

EHMT2 and CDYL signals were not detectable on spermatocyte spreads. For this reason, we tested for co-expression of PRDM9, CDYL, and EHMT2 in seminiferous tubules of 14-dpp mice. CDYL signals were found in different type of cells in both the cytoplasm and the nucleus, whereas EHMT2 signals, like PRDM9, were restricted to the nucleus. Strong nuclear and cytoplasmic CDYL signal was detected in some cells close to the basal membrane which were also strongly EHMT2-positive but PRDM9 negative (Fig. 4C, top row, arrowhead). These are primarily spermatogonia, based on the lack of GATA4 signal (Fig. S3, arrowhead). Sertoli cells, marked by the presence of GATA4, showed weaker EHMT2 signal (Fig. S3, arrow). Weak nuclear CDYL and weak EHMT2 signals were detected in germ cells attached to the basal membrane that also showed weak PRDM9 signal (Fig. 4C, top row, open arrow). We consider these cells to be spermatocytes in pre-leptonema or leptonema based on their tubule staging and PRDM9 positivity (Sun et al., 2015). The increase of PRDM9 signal was accompanied by disappearance of the EHMT2 signal and translocation of the CDYL signal to the cytoplasm (Fig. 4C, top row and second row, short arrows). These cells are most probably in leptonema to

zygonema. Pachynema cells abundant in the lumen showed strong cytoplasmic CDYL but neither EHMT2 nor PRDM9 signals (Fig. 4C, third row, long arrow). Triple staining in testes of *Prdm9*$^{tm3.1Kpgn}$ and *Prdm9*^{tm1Ymat} mutant mice confirmed the meiotic arrest phenotype found by EWSR1-PRDM9 staining. *Prdm9*$^{tm3.1Kpgn}$ testes showed CDYL-EHMT2-PRDM9-positive cells extending to the lumen, and appearance of a few CDYL cytoplasmic-positive, PRDM9-EHMT2-negative cells (Fig. 4C, fourth and fifth rows). Similarly, *Prdm9*^{tm1Ymat} testes which lack PRDM9 protein did not contain any EHMT2-positive cells in the lumen but did contain some CDYL cytoplasmic-positive, EHMT2-negative cells (Fig. 4C, sixth and seventh rows). Therefore, neither the presence of truncated PRDM9 nor the complete loss of PRDM9 affected the localization patterns of CDYL or EHMT2 from pre-leptonema through zygonema/early pachynema-like stages (Fig 4C and Fig. S3).

It appears that the EWSR1, CDYL, and EHMT2 coincide with each other and with PRDM9 only in pre-leptonema to early zygonema, when all four of them are present at low concentrations. After that, CDYL is translocated to the cytoplasm; EHMT2 disappears altogether, whereas both EWSR1 and PRDM9 increase in expression until the zygonema-pachynema transition. At pachynema, PRDM9 disappears whereas EWSR1 expression dramatically increases. These findings can explain why co-IP results show low interaction between PRDM9 and either EHMT2 or CDYL, as the three proteins coincide in germ cells for only a brief time window, and the bulk of proteins are expressed either in other cell types, as is the case for EHMT2 and CDYL expression in Sertoli cells and spermatogonia, or in later meiotic stages as in the case of PRDM9. EWSR1 shows low interaction with either CDYL or EHMT2 by co-IP (Fig. 3, right panel), even though the three proteins are strongly expressed in Sertoli cells.

PRDM9 and its interactors co-localize with REC8 and SC proteins

First, we sought to determine whether co-localization between meiotic-specific cohesins such as REC8 and PRDM9 could be detected in early meiotic prophase on spermatocyte

spreads which can be staged by the appearance of the REC8 signal (Ishiguro et al., 2014). In pre-leptonema, REC8 staining shows both diffuse and punctate patterns. This is similar to the pattern of PRDM9 at the same stage. In leptonema, REC8 starts forming rod-like structures (Fig. 5A, top row) which become predominant in zygonema as elements of the chromosomal axes (Fig. 5A, second row). The co-localization of these axis structures with PRDM9 at late zygonema was apparent (Fig. 5A, third row). *Prdm9*$^{tm3.1Kpgn}$ testes had a REC8-PRDM9 pattern similar to wild type zygonema but showed less evidence for co-localization (Fig. 5A, fourth row). It should be noted that PRDM9 signal in this mutant has a more diffuse and less punctate pattern than in wild type spermatocytes suggesting that the punctate pattern reflects PRDM9 binding to hotspot DNA (Fig. 5A, fourth row; see also Fig. 5B and 5C, fourth rows). *Prdm9*^{tm1Ymat} testes contained mostly cells with diffuse, punctate and rod-like REC8 patterns reminiscent of pre-leptonema to early zygonema, indicating that progression of spermatocytes into zygonema requires an intact PRDM9-bound hotspot translocation to the chromosomal axis (Fig. 5A, fifth row).

Next, we tested whether PRDM9 co-localizes with SC proteins in early meiotic prophase by double staining with a combination of PRDM9 and either SYCP3 or SYCP1 antibody on spermatocyte spreads. SYCP3 and SYCP1 show punctate or rod-like staining in early meiosis, similar to that of REC8 (Fig. 5B and 5C). Both SC protein signals overlapped with PRDM9, with a clearer pattern in zygonema (Fig. 5B and 5C, second rows). Because this pattern makes it difficult to determine whether the proteins truly co-localize, we determined the statistical significance of their colocalization by computer analysis of the overlap between staining sites after inverting the image of one of the signals 180^{O} relative to the other (Kumar et al., 2015). In wild type mice, PRDM9 showed co-localization with SYCP3 in zygotene (Li's ICQ value p=0.00006, Fig. 5D, left panel). PRDM9 also showed co-localization with SYCP1 when the latter first appeared in mid-zygotene (p=0.0041, Fig. 5D, right panel).

Importantly, we compared these results with the colocalization pattern of the *Prdm9*$^{tm3.1Kpgn}$ mutant to determine whether the colocalization of PRDM9 with chromosomal axis proteins is dependent on its binding to DNA. At best, we found very weak evidence of significant co-localization between PRDM9 and either SYCP3 or SYCP1 in this mutant (Fig. 5D, left panel, p=0.13, and Fig. 5D, right panel, p=0.16). Moreover, the difference in the colocalization patters between wild type and mutant spermatocytes was statistically significant for both combinations ($p<0.05$).

The combined co-IP and cytological evidence indicates that binding of PRDM9 to DNA either promotes its eventual translocation to the chromosomal axis, or else stabilizes the PRDM9-chromosomal axis protein complexes so that they are more easily detected.

DISCUSSION

PRDM9 directly interacts with EWSR1, CXXC1, EHMT2, and CDYL

Yeast two-hybrid screens and *in vitro* binding assays show that EWSR1, CXXC1, EHMT2, and CDYL can directly bind to PRDM9 via its KRAB domain and by using additional contact points extending further to the PR/SET domain. With the exception of CXXC1, for which adequately specific antibodies are not currently available, we confirmed that these interactions also occur in mouse spermatocytes. However, the spatial and temporal pattern of expression of each of these proteins *in vivo* suggests that these interactions may occur at different stages of meiotic progression and could even involve two distinct PRDM9 complexes, one with EHMT2 and CDYL and another with EWSR1 and possibly CXXC1.

EHMT2 and CDYL are expressed strongly in Sertoli cells and in spermatogonia, before PRDM9 appears at the onset of meiosis, and they remain present in the subsequent pre-leptotene and leptotene stages when PRDM9 is first detected in the nuclei of the germ cells (Sun et al., 2015). In early zygonema, EHMT2 disappears from the nucleus, whereas CDYL is translocated to the cytoplasm and remains there into pachynema. PRDM9 increases its

presence in nucleus in zygonema, and then starts to disappear, being absent from the germ cell nuclei at pachynema and later meiotic stages. CDYL interacts strongly with EHMT2 as detected by co-IP experiments in testes, suggesting that their interactions take place in any cell type where they are present together – Sertoli cells, spermatogonia, and early spermatocytes. The two proteins show similar weak co-IP patterns with PRDM9, due to their limited presence overlap in leptonema-early zygonema. These data confirm previous results showing that CDYL readily complexes EHMT2 in mouse embryonic stem cells where the two proteins are found bound together in heterochromatin regions, including the inactivated X chromosome in females (Escamilla-Del-Arenal et al., 2013). This finding contrasts with the fact that PRDM9 binding to DNA results in chromatin activation by catalyzing H3K4 trimethylation of the nearby nucleosomes (Baker et al., 2014), suggesting that a triple PRDM9-CDYL-EHMT2 complex may well have a different molecular function than the double CDYL-EHMT2 complex lacking PRDM9.

EWSR1 expression coincides with PRDM9 in leptonema and zygonema. Although their punctate appearance made it hard to determine whether that the two proteins truly co-localize cytologically, these data, together with the results of Y2H, *in vitro* binding, and the strong mutual co-IP of the two proteins, make a strong case for the likelihood that the two proteins physically interact with each other from pre-leptonema to late zygonema. However, unlike PRDM9 and EHMT2, which disappear by the end of zygonema, and CDYL, which is translocated to the cytoplasm in mid-zygonema, EWSR1 dramatically increases in nuclei at pachynema (Fig. 4B). The continued presence of EWSR1 expression at later stages suggests that EWSR1 likely provides two distinct functions: an earlier one when complexed with PRDM9 and a later one when PRDM9 has disappeared. In this context it is notable that, EWSR1 showed weak co-IP signal with EHMT2 and CDYL in wild type testis, reminiscent of the co-IP results of PRDM9 with either of EHMT2 and CDYL. Moreover, the co-IP signals of EWSR1 with EHMT2 and CDYL all but disappeared in *Prdm9*^{tm1Ymat} testes lacking PRDM9 protein but remained equally strong in

the *Prdm9*$^{tm3.1Kpgn}$ mutant expressing PRDM9 without its DNA-binding domain. This suggests the possibility for the presence of a short-term and possibly weak PRDM9-EWSR1-CDYL-EHMT2 complex in and around leptonema, followed by separation of CDYL-EHMT2 and their subsequent disappearance from the nucleus, leaving PRDM9 and EWSR1 bound throughout zygonema, or alternatively, existence of separate PRDM9-EHMT2-CDYL and PRDM9-EWSR1 complexes. In humans EWSR1 interacts with BARD1 (Spahn et al., 2002), which apart from other recombination factors (BRCA1, RAD51 etc.) interacts with SETDB1 (Goehler et al., 2004) – a H3K9 histone methyltransferase, which in turn interacts with CDYL-EHMT2 (Mulligan et al., 2008). PRDM9 may have similar role like SETDB1 by taking central place as a shared binding partner of EWSR1 and CDYL-EHMT2 complex. This possibility is further corroborated by the recent findings that PRDM9, like SETDB1, is capable of trimethylating H3K9 *in vitro* (Powers et al., 2016; Wu et al., 2013) although the presence of H3K9ac at hotspots (Buard et al., 2009) might prevent this activity *in vivo*.

CXXC1 binds directly to PRDM9 in Y2H and in *in vitro* experiments. Unfortunately, the lack of appropriate antibodies against CXXC1 prevented us from detecting its interactions with other mammalian meiotic proteins *in vivo*. What is known is that in mammals CXXC1 is part of the Set1 complex responsible for most of H3K4 trimethylation in somatic cells (Lee and Skalnik, 2005; Shilatifard, 2012). In ES cells in particular, it is required for both H3K4me3 deposition after DNA damage and the subsequent acetylation of H3K9 at the same nucleosomes (Clouaire et al., 2014). With respect to its possible functions in meiosis, studies of Spp1, the yeast homolog of CXXC1, provide clues to its possible role in mammalian meiosis. In *S. cerevisiae*, DSBs initiate at H3K4me3 sites found near promoters (Borde et al., 2009). DSB formation at these sites is promoted when they become tethered to the chromosomal axis by Spp1/CXXC1 (Sommermeyer et al., 2013). We now show evidence that CXXC1 binds to PRDM9 directly, and that REC8 interaction with PRDM9 is indirect, possibly requiring the mediation of other proteins. This suggests a possible role for CXXC1, linking PRDM9-bound and subsequently

H3K4-trimethylated sites to the chromosomal axis, and promoting DSBs initiation. In contrast to its yeast homologue Spp1, which links H3K4me3 sites to chromosomal axis in yeast (Acquaviva et al., 2013; Sommermeyer et al., 2013), which lack any protein functioning like PRDM9, direct binding of CXXC1 to PRDM9 in mammals could be either the triggering event of such translocation, or enhance the likelihood that CXXC1 will bind H3K4-trimethylated sites specifically at PRDM9-dependent hotspots ensuring DSB initiation at sites where they could be properly repair.

PRDM9 interacts and colocalizes with meiotic-specific cohesin REC8 and synaptonemal complex proteins SYCP1 and SYCP3

We found no evidence for direct binding between PRDM9 and the meiotic cohesin REC8 in either our Y2H screen or when the two proteins were tested *in vitro* after expression in *E. coli.* However, the two proteins showed ample evidence of interaction in spermatocytes both by pull-down experiments and cytological co-localization. The two proteins also interacted when they were expressed together in HEK 293 cell cultures, suggesting that if they require an additional molecule to provide a link, this hypothetical partner need not to be meiosis-specific. One potential candidate is EWSR1, a generally “sticky” protein (Schwartz et al., 2015) which pulls down both PRDM9 and REC8 in spermatocytes; perhaps a more likely candidate, given its known role in bring hotspots to the chromosome axis in yeast, is CXXC1. These are not mutually exclusive possibilities.

This opens the possibility that complexes including PRDM9, EWSR1, and other proteins, such as CXXC1, can play a key role in homolog recognition by bringing the homologous hotspot DNA sequences bound to PRDM9 in contact with cohesins and subsequently to the chromosomal axis where they can find each other.

Our findings that the synaptonemal complex proteins REC8 and SYCP3, and to some extent SYCP1, not only interact with each other, but also with PRDM9, EHMT2 and EWSR1, and to a

lesser extent with CDYL support the concept that these proteins play important roles in bringing activated hotspots from out in the DNA loops down to the chromosome axis, stabilizing hotspot location there, and then participating in DSB formation and repair. Interestingly, both SYCP3 and SYCP1 showed stronger binding with PRDM9 than with each other (Fig. 2C). This cannot be explained solely by concentration differences as all three proteins are prominent in leptonema and zygonema (Fig.5B and 5C). Since SYCP1 and SYCP3 do not bind directly to each other but through SYCP2, and then SYCP1 participates in the SC central element formation interacting with intermediates such as SYCP2, SYCE1, SYCE2 and SYCE3, and TEX12 (see (Bolcun-Filas and Schimenti, 2012) for review), this suggests that PRDM9 might come to a closer contact with SYCP1 rather than just bringing the hotspot DNA to the chromosomal axis, probably because it remains bound to DNA for some time after DSB are initiated and homology search occurs in the internal space of the synaptonemal complex.

Judging by the properties of its yeast homolog Spp1, CXXC1 may well function in combination with the histone modifications placed by PRDM9 at hotspots in the transport of hotspots to the chromosome axis. Once DSBs have formed there, EWSR1 is known to promote ssDNA invasion into dsDNA (Guipaud et al., 2006; Schwartz et al., 2015) as occurs in the formation of the meiotic Holliday Junctions involved in DNA exchange between homologs, and has been shown to contribute to DSB repair(Li et al., 2007). And EHMT2 and CDYL are known to be involved in the establishment of closed chromatin states (Escamilla-Del-Arenal et al., 2013; Leung et al., 2011; Xu and Price, 2011) such as those evidenced by the presence of γH2AX around as yet unrepaired meiotic DNA lesions (Chicheportiche et al., 2007). Nucleosomes in the immediate vicinity of DSB sites are modified to create an open chromatin state required for repair (Baker et al., 2014; Xu and Price, 2011). In this regard, EHMT2-CDYL temporary presence at PRDM9-bound complexes could help establish the necessary closed chromatin configuration before DSBs are initiated by SPO11.

Protein interactions are not dependent of PRDM9 binding to DNA

Our characterization of protein-protein interactions in the *Prdm9*$^{tm3.1Kpgn}$ mutant showed that PRDM9 binding to its partners is not dependent on its binding to DNA. This is in agreement with the Y2H data showing that PRDM9 binding to EWSR1, EHMT2, CDYL, and CXXC1 occurs through its N-terminal part including the KRAB and PR/SET domains. However, co-localization scores of both PRDM9-SYCP3 and PRDM9-SYCP1 in this mutant are not statistically significant and generally lower than in wild type testis (Fig. 5D). This suggests that PRDM9 binding to DNA may stabilize the complexes, by itself or by recruiting additional proteins, on the chromosomal axis to ensure proper DSB initiation and their subsequent repair.

EWSR1, EHMT2, and CDYL bind to synaptonemal complex in the absence of PRDM9

PRDM9 presence by itself is not necessary for binding of EWSR1 and EHMT2-CDYL complexes to chromosomal axis as evidenced by our SYCP3 co-IP results in testes of *Prdm9*^{tm1Ymat} KO mutant (Fig. 3, central panel). EHMT2 and CDYL complexes have not been previously studied in the context of meiotic prophase, however, their association with closed chromatin and X inactivation has been demonstrated (Escamilla-Del-Arenal et al., 2013; Zhang et al., 2011). EWSR1 binding to SYCP3 seems to be stabilized in this mutant suggesting that its subsequent release may be due to the course of events occurring after PRDM9-DNA complexes are brought to and/or stabilized on the chromosomal axis.

Model of events occurring before recombination initiation

Our results show extensive protein-DNA and protein-protein interactions as part of recombination-related events in early meiotic prophase. Our data suggest the following working model for the mechanisms and dynamics of these events (Fig. 6). In early leptonema, PRDM9 binds to hotspots where it eventually trimethylates histone 3 at lysine-4 and lysine-36 in the adjacent nucleosomes. The KRAB domain of PRDM9 complexes CXXC1 and possibly EWSR1

around the same time. The presence of CXXC1 (and possibly EWSR1) in combination with a signal resulting from PRDM9 histone methylation causes hotspots to be translocated from out in chromosome loops down to the chromosomal axis where DSBs can form. EHMT2 and CDYL, already present on the chromosomal axis, bind to PRDM9 and participate in establishing the surrounding closed chromatin state necessary for DSB initiation. They disappear from the nucleus shortly after fulfilling this function, in early zygonema. PRDM9 remains bound to the hotspots until DSB are initiated, and probably later, bringing the hotspot DNA into the internal SC space, where the complexes come into contact with SYCP1. PRDM9 then disappears from the nucleus in late zygonema, whereas EWSR1 further participates in the formation and resolution of Holliday junctions at pachynema, and carries out additional functions.

As interesting as this model may be in suggesting new experimental directions, it does lack at least one essential element: what brings the homologous hotspot down to the chromosome axis for precise pairing and DNA exchange. We know from previous studies (Baker et al., 2015) that precise recombination can occur between homologous hotspots even when one has lost the ability to bind PRDM9, removing PRDM9 binding or its consequences as the requisite signal. The model proposed here brings one homolog down to the chromosome axis, but we still lack a model for what brings its homologous partner into apposition.

MATERIALS AND METHODS

Constructs

pBAD-Prdm9 was described in (Billings et al., 2013).

pGBKT7-Prdm9 – whole length Prdm9 ORF from pBAD-Prdm9 was amplified with primers pR638+pR916 and inserted into pGBKT7 using BamHI-PstI restriction sites.

pGBKT7-Prdm9 KRAB–PR-SET – pR638+pR608 .

pGBKT7-Prdm9 KRAB – pR790+pR1382.

pGBKT7-Prdm9 ID–PR-SET – pR1658+ pR1661.

pGBKT7-Prdm9 ID/2 – pR1658+pR1383

pGBKT7-Prdm9 ID – pR1658+pR1659

pGAD-T7- Prdm9 construct was created by restriction excision by EcoRI-SalI from pBAD-Prdm9 and insertion by same sites in pGAD-T7 empty vector.

Antibodies

Anti-PRDM9	Guinea pig custom made		1:200
Anti-EHMT2	PP-A8620A-00	Perseus Proteomics	1:500
Anti-EWSR1	ab54708	Abcam	1:200
Anti-SYCP3	NB300-231	Novus Biologicals	1:500
Anti-SYCP1	NB 300-229	Novus Biologicals	1:500
Anti-CDYL	ab5188-50	Abcam	1:50
Anti-REC8	ab38372	Abcam	1:200
Anti-MBP	E8038S	New England Biolabs	1:1000
Anti-HORMAD2	sc-82192	Santa Cruz	1:1000

Goat anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor® 594 conjugate (Life Technologies, Cat. A-11037, Carlsbad, CA), 1:1000

Goat anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor® 488 conjugate (Life Technologies Cat. ab150077), 1:1000

Goat anti-Mouse IgG (H+L) Secondary Antibody, Alexa Fluor® 488 conjugate (Life Technologies, Cat. ab150113), 1:1000

Goat anti-Mouse IgG (H+L) Secondary Antibody, Alexa Fluor® 647 conjugate (Life Technologies, Cat. A-21236), 1:1000

Goat anti-Guinea pig IgG (H+L) Secondary Antibody, Alexa Fluor® 594 conjugate (Life Technologies, Cat. A-11076), 1:1000

Goat anti-Guinea pig IgG (H+L) Secondary Antibody, Alexa Fluor® 488 conjugate (Life Technologies, Cat. A-11073), 1:1000

Yeast two-hybrid screen

The yeast two-hybrid screen was performed by co-transformation of pGBKT7-Prdm9 construct and mouse testes cDNA library (3x106 to 107 clones) (#638848, Clontech) in PJ69-4α strain. The positive interactions were selected by plating the transformants on –Trp, -Leu, -His plates supplemented with 3-Amino-1,2,4 triazole (3-AT) chemical for detection of the strongest interactions. Totally 329 colonies showing positive interactions were isolated. Each colony was re-streaked, and plasmid DNA was isolated and sequenced to identify the individual cDNA library clone.

Yeast two-hybrid screen validation and Prdm9 domain mapping

The detected open-reading frames interacting with Prdm9 were isolated and transformed in PJ69-4A strain. The pGBKT7-Prdm9 constructs were transformed in PJ69-4 α strain. The PJ69-4 α strain was crossed with PJ69-4A strain carrying the clone of interest and the plated on selective media plates for confirmation of positive interactions.

Protein-protein *in vitro* pull-downs

Expression of PRDM9 was performed in Arctic DE3 cells. Preculture was grown overnight at 30^{o}C. Next day the cells were re-innocuilated, the culture was grown for 3-4 hours and shifted for 16-24 hours at 14 ^{o}C at 200 rpm. The resulted bacterial pellet was ground by SPEX™ SamplePrep 6870 Freezer/Mill™and dissolved in CBB+150mM KCl buffer 1X (see the recipe bellow) in roughly 10ml per 1g pellet, supplied with protease inhibitor cocktail (aprotinin, chymostatin, leupeptin, pepstatin (4XPI)) - Applichem, 1mM PMSF - Applichem, 1mM mercaptoethanol, 0.01% NP40.

The purification of each protein was done in three steps - Sp-Sepharose, amylose beads, and FPLC salt gradient on MonoS column. The pull-down of the purified PRDM9 protein and the protein of interest was done by amylose beads, or affinity beads corresponding to the tag of the second protein. The pull-down was performed by mixing of 1mcg of each protein for 30min at 34 degrees, followed by addition of 30mcl beads washed by K+150mM KCl buffer and incubated for another 30min. The bound proteins were eluted with SDS loading buffer.

Co-immunoprecipitation assays

Testis material from twenty 14-dpp old C57BL/6J male mice was extracted in cold PBS, homogenized by Dounce-homogenizer, passed through 40 µm cell strainer (Falcon BD # 352340), centrifuged for 5 min at 2000rpm (Beckman Coultier centrifuge). The cell pellet was re-suspended in 1 ml Pierce IP buffer with 1mM PMSF and 1X protease inhibitor cocktail EDTA-free (Roche). The sample was incubated for 30 min on rotation and centrifuged at 13,200g. One-tenth of the sample was used for Input, and the rest for the pull-down. The pulldown was done by protein A, or G-beads (Dynabeads, Lifesciences) dependent on the antibody source. As a negative control, IgG from the same animal species as the pulldown antibody was used. The extract was incubated overnight at 4^{o}C with rotation. After washing the beads three times with 1ml Pierce IP buffer, the complexes were eluted with 200 µl GST buffer (0.2 M glycine, 0.1% SDS, 1% Tween 20, pH 2.2) for 20 min at room temperature with rotation. The sample was neutralized with 40 µl 1M Tris-HCl pH 8, then 40 µl SDS-loading buffer was added and the samples were subjected to SDS-PAGE and Western Blotting.

Chromosome spreads

For preparation of nuclear spreads, the drying-down technique (Peters et al., 1997) was used in C57BL/6J, *Prdm9*^{TM1Ymat} and C57BL/6J-*Prdm9*$^{tm3.1Kpgn}$ male mice by double or

consecutive immunolabeling with PRDM9/SYCP3, PRDM9/SYCP1, PRDM9/EWSR1/SYCP3, EWSR1/BRCA1/SYCP3, SYCP3/γH2AX/CREST, and PRDM9/REC8.

PAS staining

For histological evaluation, tissues were dissected out, fixed with Bouin's solution, embedded in paraffin wax and sectioned at 5 µm. Sections were stained with Periodic acid–Schiff–diastase (PAS) using standard techniques.

Immunofluorescent staining

For protein immunolocalization, tissues were dissected out, fixed with 4% PFA solution, embedded in paraffin wax and sectioned at 5 µm. Sections were microwave heated in 10 mM sodium citrate buffer, pH 6.0 for 10 min after deparaffinization, and then treated with PBS containing 0.1% Triton X-100. After blocking non-specific binding sites with 10% normal donkey serum (Jackson ImmunoResearch Labs, Inc., Cat.# 017-000-121, West Grove, PA, USA), sections were then incubated with primary antibodies overnight at 4 ºC and secondary antibodies for 2 h at room temperature, respectively. Finally, slides were rinsed in PBS, stained for 3 min with DAPI (4', 6-diamidino-2-phenylindole) (Sigma, Cat.# 28718-90-3, St. Louis, MO, USA), rinsed three times in PBS for 5 min each and mounted in Antifade reagent (Life Techonologies, Cat.# S-2828, Carlsbad, CA). Sections were observed and photographed with Microscope Axio Imager.Z2 (Zeiss, Jena, Germany).

TUNEL assay

For detection of apoptosis in tissues, sections were subjected to fluorescence labelling of DNA strand breaks by Terminal deoxy-nucleotidyl transferase-mediated digoxigenin-dUTP nick-end labeling (TUNEL) assay, using the In situ cell death detection kit (Roche, Cat.#

11684795910, Basel, Switzerland) according to the manufacturer's protocol. As a negative control, the TdT enzyme was omitted in parallel reactions.

ACKNOWLEDGMENTS

The authors thank to Anita Hawkins and Catrina Spruce for technical help, and to Mary Ann Handel for critical reading and helpful suggestions.

CONFLICT OF INTEREST

The authors declare no conflict of interests.

AUTHOR CONTRIBUTIONS

EDP, KP and PMP designed the experiments; EDP, HT, TB, RS and RA performed the experiments; LK provided material support and guidance; EDP, HT, KP and PMP analyzed the data; PMP and KP wrote the paper.

FUNDING

This work was supported by NIH grants R01 GM078452 to PMP, P50 GM076468 to Gary Churchill/project B to PMP, R01 GM078643 to KP, Cancer Core grant CA34196 to The Jackson Laboratory; Czech Science Foundation grants GACR13-26629S and GACR207/12/2323 to LK, and South Moravian Programme grant 2SGA2773 to EDP.

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**, 215-8.

Baker, C. L., Kajita, S., Walker, M., Saxl, R. L., Raghupathy, N., Choi, K., Petkov, P. M. and Paigen, K. (2015). PRDM9 drives evolutionary erosion of hotspots in Mus musculus through haplotype-specific initiation of meiotic recombination. *PLoS genetics* **11**, e1004916.

Baker, C. L., Walker, M., Kajita, S., Petkov, P. M. and Paigen, K. (2014). PRDM9 binding organizes hotspot nucleosomes and limits Holliday junction migration. *Genome research* **24**, 724-32.

Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G. and de Massy, B. (2010). PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* **327**, 836-40.

Baudat, F., Imai, Y. and de Massy, B. (2013). Meiotic recombination in mammals: localization and regulation. *Nature Rev Genet* **14**, 794-806.

Billings, T., Parvanov, E. D., Baker, C. L., Walker, M., Paigen, K. and Petkov, P. M. (2013). DNA binding specificities of the long zinc-finger recombination protein PRDM9. *Genome biology* **14**, R35. PMCID: PMC4053984.

Bolcun-Filas, E. and Schimenti, J. C. (2012). Genetics of meiosis and recombination in mice. *International review of cell and molecular biology* **298**, 179-227.

Borde, V., Robine, N., Lin, W., Bonfils, S., Geli, V. and Nicolas, A. (2009). Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites. *The EMBO journal* **28**, 99-111.

Brick, K., Smagulova, F., Khil, P., Camerini-Otero, R. D. and Petukhova, G. V. (2012). Genetic recombination is directed away from functional genomic elements in mice. *Nature* **485**, 642-5.

Buard, J., Barthes, P., Grey, C. and de Massy, B. (2009). Distinct histone modifications define initiation and repair of meiotic recombination in the mouse. *The EMBO journal* **28**, 2616-2624.

Chicheportiche, A., Bernardino-Sgherri, J., de Massy, B. and Dutrillaux, B. (2007). Characterization of Spo11-dependent and independent phospho-H2AX foci during meiotic prophase I in the male mouse. *Journal of cell science* **120**, 1733-42.

Clouaire, T., Webb, S. and Bird, A. (2014). Cfp1 is required for gene expression-dependent H3K4 trimethylation and H3K9 acetylation in embryonic stem cells. *Genome biology* **15**, 451.

Escamilla-Del-Arenal, M., da Rocha, S. T. and Heard, E. (2011). Evolutionary diversity and developmental regulation of X-chromosome inactivation. *Human genetics* **130**, 307-27.

Escamilla-Del-Arenal, M., da Rocha, S. T., Spruijt, C. G., Masui, O., Renaud, O., Smits, A. H., Margueron, R., Vermeulen, M. and Heard, E. (2013). Cdyl, a new partner of the inactive X chromosome and potential reader of H3K27me3 and H3K9me2. *Molecular and cellular biology* **33**, 5005-20.

Fisher, C. (2014). The diversity of soft tissue tumours with EWSR1 gene rearrangements: a review. *Histopathology* **64**, 134-50.

Fukuda, T., Pratto, F., Schimenti, J. C., Turner, J. M., Camerini-Otero, R. D. and Hoog, C. (2012). Phosphorylation of chromosome core components may serve as axis marks for the status of chromosomal events during mammalian meiosis. *PLoS genetics* **8**, e1002485.

Fumasoni, I., Meani, N., Rambaldi, D., Scafetta, G., Alcalay, M. and Ciccarelli, F. D. (2007). Family expansion and gene rearrangements contributed to the functional specialization of PRDM genes in vertebrates. *BMC evolutionary biology* **7**, 187.

Goehler, H., Lalowski, M., Stelzl, U., Waelter, S., Stroedicke, M., Worm, U., Droege, A., Lindenberg, K. S., Knoblich, M., Haenig, C. et al. (2004). A protein interaction network

links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease. *Molecular cell* **15**, 853-65.

Grey, C., Sommermeyer, V., Borde, V. and de Massy, B. (2011). [What defines the genetic map? The specification of meiotic recombination sites]. *Medecine sciences : M/S* **27**, 63-9.

Guipaud, O., Guillonneau, F., Labas, V., Praseuth, D., Rossier, J., Lopez, B. and Bertrand, P. (2006). An in vitro enzymatic assay coupled to proteomics analysis reveals a new DNA processing activity for Ewing sarcoma and TAF(II)68 proteins. *Proteomics* **6**, 5962-72.

Hayashi, K., Yoshida, K. and Matsui, Y. (2005). A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. *Nature* **438**, 374-8.

Hinch, A. G., Tandon, A., Patterson, N., Song, Y., Rohland, N., Palmer, C. D., Chen, G. K., Wang, K., Buxbaum, S. G., Akylbekova, E. L. et al. (2011). The landscape of recombination in African Americans. *Nature* **476**, 170-5.

Illingworth, R. S., Gruenewald-Schneider, U., Webb, S., Kerr, A. R., James, K. D., Turner, D. J., Smith, C., Harrison, D. J., Andrews, R. and Bird, A. P. (2010). Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS Genet* **6**.

Ishiguro, K., Kim, J., Shibuya, H., Hernandez-Hernandez, A., Suzuki, A., Fukagawa, T., Shioi, G., Kiyonari, H., Li, X. C., Schimenti, J. et al. (2014). Meiosis-specific cohesin mediates homolog recognition in mouse spermatocytes. *Genes & development* **28**, 594-607.

Kumar, R., Ghyselinck, N., Ishiguro, K., Watanabe, Y., Kouznetsova, A., Hoog, C., Strong, E., Schimenti, J., Daniel, K., Toth, A. et al. (2015). MEI4 - a central player in the regulation of meiotic DNA double-strand break formation in the mouse. *Journal of cell science* **128**, 1800-11.

Lahn, B. T., Tang, Z. L., Zhou, J. X., Barndt, R. J., Parvinen, M., Allis, C. D. and Page, D. C. (2002). Previously uncharacterized histone acetyltransferases implicated in

mammalian spermatogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 8707-8712.

Lee, J. H. and Skalnik, D. G. (2005). CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex. *The Journal of biological chemistry* **280**, 41725-31.

Leung, D. C., Dong, K. B., Maksakova, I. A., Goyal, P., Appanah, R., Lee, S., Tachibana, M., Shinkai, Y., Lehnertz, B., Mager, D. L. et al. (2011). Lysine methyltransferase G9a is required for de novo DNA methylation and the establishment, but not the maintenance, of proviral silencing. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 5718-23.

Li, H., Watford, W., Li, C., Parmelee, A., Bryant, M. A., Deng, C., O'Shea, J. and Lee, S. B. (2007). Ewing sarcoma gene EWS is essential for meiosis and B lymphocyte development. *The Journal of clinical investigation* **117**, 1314-23.

Lupo, A., Cesaro, E., Montano, G., Zurlo, D., Izzo, P. and Costanzo, P. (2013). KRAB-Zinc Finger Proteins: A Repressor Family Displaying Multiple Biological Functions. *Current genomics* **14**, 268-78.

Mulligan, P., Westbrook, T. F., Ottinger, M., Pavlova, N., Chang, B., Macia, E., Shi, Y. J., Barretina, J., Liu, J., Howley, P. M. et al. (2008). CDYL bridges REST and histone methyltransferases for gene repression and suppression of cellular transformation. *Molecular cell* **32**, 718-26.

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**, 876-9.

Oakland, T. E., Haselton, K. J. and Randall, G. (2013). EWSR1 binds the hepatitis C virus cis-acting replication element and is required for efficient viral replication. *Journal of virology* **87**, 6625-34.

Paigen, K. and Petkov, P. (2010). Mammalian recombination hot spots: properties, control and evolution. *Nature reviews. Genetics* **11**, 221-33.

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

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 research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* **5**, 66-8.

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 genetics*.

Schwartz, J. C., Cech, T. R. and Parker, R. R. (2015). Biochemical Properties and Biological Functions of FET Proteins. *Annual review of biochemistry* **84**, 355-79.

Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annual review of biochemistry* **81**, 65-95.

Smagulova, F., Gregoretti, I. V., Brick, K., Khil, P., Camerini-Otero, R. D. and Petukhova, G. V. (2011). Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. *Nature* **472**, 375-8.

Sommermeyer, V., Beneut, C., Chaplais, E., Serrentino, M. E. and Borde, V. (2013). Spp1, a member of the Set1 Complex, promotes meiotic DSB formation in promoters by tethering histone H3K4 methylation sites to chromosome axes. *Molecular cell* **49**, 43-54.

Spahn, L., Petermann, R., Siligan, C., Schmid, J. A., Aryee, D. N. and Kovar, H. (2002). Interaction of the EWS NH2 terminus with BARD1 links the Ewing's sarcoma gene to a common tumor suppressor pathway. *Cancer research* **62**, 4583-7.

Steggerda, S. M. and Paschal, B. M. (2002). Regulation of nuclear import and export by the GTPase Ran. *International review of cytology* **217**, 41-91.

Sun, F., Fujiwara, Y., Reinholdt, L. G., Hu, J., Saxl, R. L., Baker, C. L., Petkov, P. M., Paigen, K. and Handel, M. A. (2015). Nuclear localization of PRDM9 and its role in meiotic chromatin modifications and homologous synapsis. *Chromosoma* **124**, 397-415.

Tachibana, M., Nozaki, M., Takeda, N. and Shinkai, Y. (2007). Functional dynamics of H3K9 methylation during meiotic prophase progression. *The EMBO journal* **26**, 3346-59.

Tachibana, M., Ueda, J., Fukuda, M., Takeda, N., Ohta, T., Iwanari, H., Sakihama, T., Kodama, T., Hamakubo, T. and Shinkai, Y. (2005). Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes & development* **19**, 815-26.

Tate, C. M., Lee, J. H. and Skalnik, D. G. (2009). CXXC finger protein 1 contains redundant functional domains that support embryonic stem cell cytosine methylation, histone methylation, and differentiation. *Molecular and cellular biology* **29**, 3817-31.

Walker, M., Billings, T., Baker, C. L., Powers, N., Tian, H., Saxl, R. L., Choi, K., Hibbs, M. A., Carter, G. W., Handel, M. A. et al. (2015). Affinity-seq detects genome-wide PRDM9 binding sites and reveals the impact of prior chromatin modifications on mammalian recombination hotspot usage. *Epigenetics & chromatin* **8**, 31.

Wu, H., Mathioudakis, N., Diagouraga, B., Dong, A., Dombrovski, L., Baudat, F., Cusack, S., de Massy, B. and Kadlec, J. (2013). Molecular Basis for the Regulation of the H3K4 Methyltransferase Activity of PRDM9. *Cell reports* **5**, 13-20.

Xu, Y. and Price, B. D. (2011). Chromatin dynamics and the repair of DNA double strand breaks. *Cell cycle* **10**, 261-7.

Zhang, Y., Yang, X. H., Gui, B., Xie, G. J., Zhang, D., Shang, Y. F. and Liang, J. (2011). Corepressor Protein CDYL Functions as a Molecular Bridge between Polycomb

Repressor Complex 2 and Repressive Chromatin Mark Trimethylated Histone Lysine 27. *Journal of Biological Chemistry* **286**, 42414-42425.

FIGURE LEGENDS

Figure 1. CXXC1, EWSR1, EHMT2, and CDYL directly interact with PRDM9. A. Scheme of the domain structure of PRDM9 (top) showing the positions of the KRAB, PR/SET, and the zinc finger domains. Below are schematic representations of the full-length and deletion constructs used in yeast two-hybrid assay to map the contact points of PRDM9 interacting with the other proteins. B. Results of the yeast two-hybrid assay showing that KRAB is the major protein contact domain of PRDM9. The extents of the cloned and expressed fragments of the interactor proteins are in brackets. C. Purified PRDM9 and its interactors bind to each other *in vitro.* Purified HALO-tagged EWSR1 and CDYL, or GST-tagged EHMT2 and CXXC1 were immobilized on amylose beads alone (left half), or mixed with purified full-length MBP-tagged PRDM9 (right half). Specific interactions are detected by the immobilization of interactor protein on amylose beads only in the presence of MBP-tagged PRDM9. All four proteins specifically bind to PRDM9. SN – supernatant; B – beads.

Figure 2. PRDM9, EWSR1, and CDYL co-immunoprecipitate each other and chromosomal axis/synaptonemal complex proteins other from wild type 14-dpp spermatocytes. A. PRDM9 and its interactors co-IP each other. Left panel, Co-IP with anti-PRDM9; central panel, co-IP with anti-EWSR1; right panel, co-IP with CDYL. In each panel, Input – 1% of the amount of total testis extract used for co-IP with specific antibodies; IgG – co-IP with non-immune guinea pig (PRDM9), goat (EWSR1), or rabbit (CDYL) IgG (negative control); αPRDM9, αEWSR1, αCDYL – co-IP with the respective antibody. The antibodies used to detect specific proteins on western blots are shown on the left. B. PRDM9 and its interactors also interact with

chromosomal axis/SC proteins. Co-IP with the same antibodies used in A, but probed with antibodies against chromosome axis/SC proteins as marked on the left. C. Reciprocal co-IP shows that SC proteins co-IP PRDM9 and its interactors. Co-IP with SYCP3 (left panel) or SYCP1 (right panel). Input – 1% total testis extract; IgG – negative control; αSYCP3 and αSYCP1 – specific co-IP. The antibodies used for detection on western blots are on the left. D. Interaction between REC8 and PRDM9 in cultured cells. REC8 was cloned under a V5 tag, and PRDM9 was cloned under a FLAG-tag. The two proteins were either co-expressed or separately expressed in HEK 293 cells. Specific interaction between the two proteins was detected in extracts of cells where the two proteins were co-expressed (Co, left) but not in mixed extracts of cells where the two proteins were expressed separately (Mix, right).

Figure 3. Protein-protein interactions in 14-dpp spermatocytes of PRDM9 mutant mice. Left panel: truncated PRDM9 shows similar interaction pattern as full-length PRDM9. Co-IP with anti-PRDM9 in testes of *Prdm9*$^{tm3.1Kpgn}$ mice lacking its DNA-binding ZnF domain. Left lane – 1% input, right lane – co-IP. Central panel: SYCP3 interacts with the other proteins in the absence of PRDM9. Co-IP with anti-SYCP3 in testes of *Prdm9*^{tm1Ymat} mice lacking PRDM9 protein. Left lane – 1% input, right lane – co-IP. Right panel: EWSR1 interacts with SYCP3 in the absence of PRDM9. Co-IP with anti-EWSR1 in testes of *Prdm9*^{tm1Ymat} mice. Left lane – 1% input, central lane, co-IP with normal IgG (negative control), right lane – co-IP with specific antibody.

Figure 4. EWSR1, CDYL, and EHMT2 are co-expressed with PRDM9 in early meiotic prophase. A. Immunofluorescence analysis of EWSR1 (red) – PRDM9 (green) co-expression in tissue sections from testis tubules of 14-dpp mice. Top panel: wild type B6 mice. Middle panel: *Prdm9*$^{tm3.1Kpgn}$. Co-expression in both cases is detected in spermatocytes found in the lumen but not in spermatogonia and Sertoli cells lining the tubule basal membrane. Arrowhead, spermatogonia with high EWSR1 expression. Arrow, spermatocytes with EWSR1 and PRDM9 positive signals. Lower panel: *Prdm9*^{tm1Ymat}. EWSR1 shows similar expression pattern as in wild

type. B. Co-localization analysis of EWSR1 and PRDM9 in spermatocyte spreads. Triple staining for EWSR1 (red), SYCP3 (white), and PRDM9 (green) in leptonema, zygonema (top panels) and pachynema (lower left panel) or EWSR1 (red), SYCP3 (white), and BRCA1(green) in pachynema (lower right panel) of B6 mice. EWSR1 is excluded from the sex body marked by BRCA1 in pachynema (yellow arrow). C. Co-expression of CDYL (red), EHMT2 (blue), and PRDM9 (green) expression in tissue sections from testis tubules of 14-dpp mice. Top three rows, co-expression analysis in B6 mice. Fourth and fifth rows, expression in *Prdm9*$^{tm3.1Kpgn}$ mutant. Sixth and seventh row, expression in *Prdm9*^{tm1Ymat} mutant. Arrowhead, Sertoli cells showing strong nuclear and cytoplasmic CDYL and strong EHMT2 signal PRDM9 negative. Open arrow, pre-leptonema to leptonema cells with weak nuclear CDYL, EHMT2, and PRDM9 signals. Short arrow, late leptonema to early zygonema cells with strong PRDM9, cytoplasmic CDYL, and lack of EHMT2 signals. Long arrow, pachynema cells with strong cytoplasmic CDYL but lack of EHMT2 and PRDM9 signals.

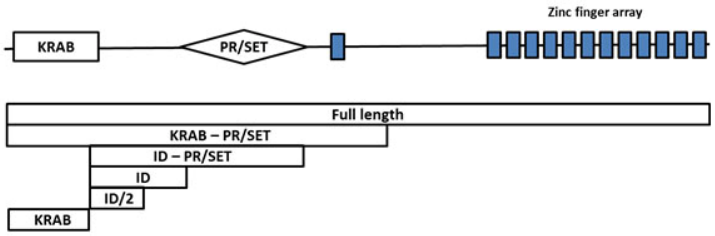
Figure 5. REC8, SYCP3, and SYCP1 co-localize with PRDM9 in early meiotic prophase. A. Co-localization of REC8 (red) and PRDM9 (green) in spermatocyte spreads. First three rows, B6 spermatocytes in leptonema (top row), early zygonema (second row), and late zygonema (third row). Arrows show co-localization pattern of the two proteins becoming more prominent at later stages. Fourth row, co-localization pattern in late zygonema-like cells in the *Prdm9*$^{tm3.1Kpgn}$ mutant. Note the more diffuse pattern of the truncated PRDM9 lacking its DNA-binding domain compared to wild type in the third row. Fifth row, leptonema or early zygonema- like cell in the *Prdm9*^{tm1Ymat} mutant. B. Co-localization of SYCP3 (red) and PRDM9 (green) in spermatocyte spreads. Disposition of the panels is the same as in A. Arrows show co-localization pattern of the two proteins. C. Co-localization of SYCP1 (red) and PRDM9 (green) in spermatocyte spreads. Disposition of the panels is the same as in A. D. Quantitation of SYCP3/PRDM9 (left panel) and SYCP1/PRDM9 (right panel) co-localization in B6 and *Prdm9*$^{tm3.1Kpgn}$ mutant

compared to the co-localization pattern of images where the two signals were inverted to 180^{o} relative to each other. Asterisk, p <0.05.

Figure 6. Model of events occurring before recombination initiation (see text for details)**.**

Fig.1

A



B

	KRAB	ID/2	ID	ID - PR/SET	KRAB - PR/SET	Full-length	Interacting PRDM9 domain
CXXC1 (217-660 aa)	+	-	-	-	+	+	KRAB
EHMT2 (376-1263 aa)	+	+	+	+	+	+	KRAB and multipoint
CDYL (74-593 aa)	+	-	-	+	+	+	KRAB and PR-SET
EWSR1 (479-655 aa)	+	-	+	+	+	+	KRAB and multipoint

C

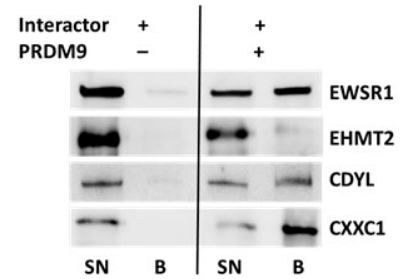


Fig. 2

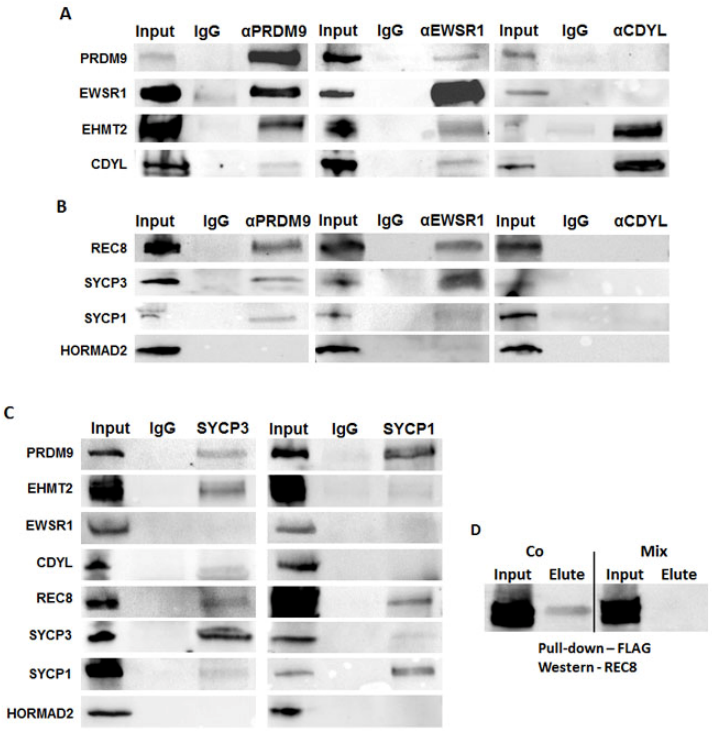


Fig.3

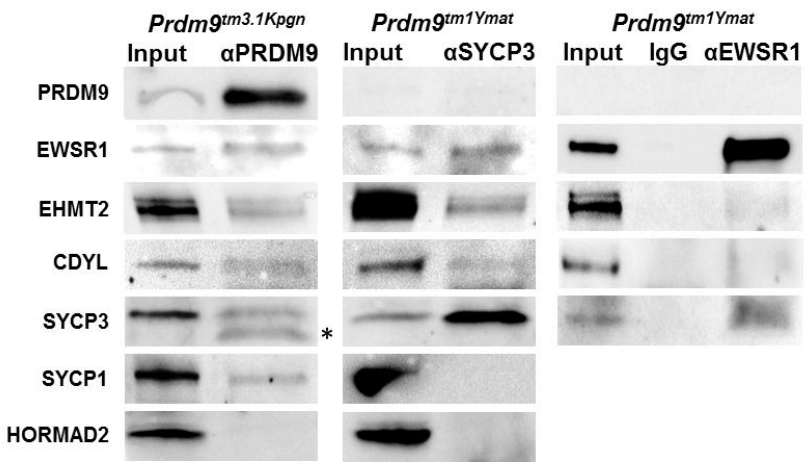


Fig. 4

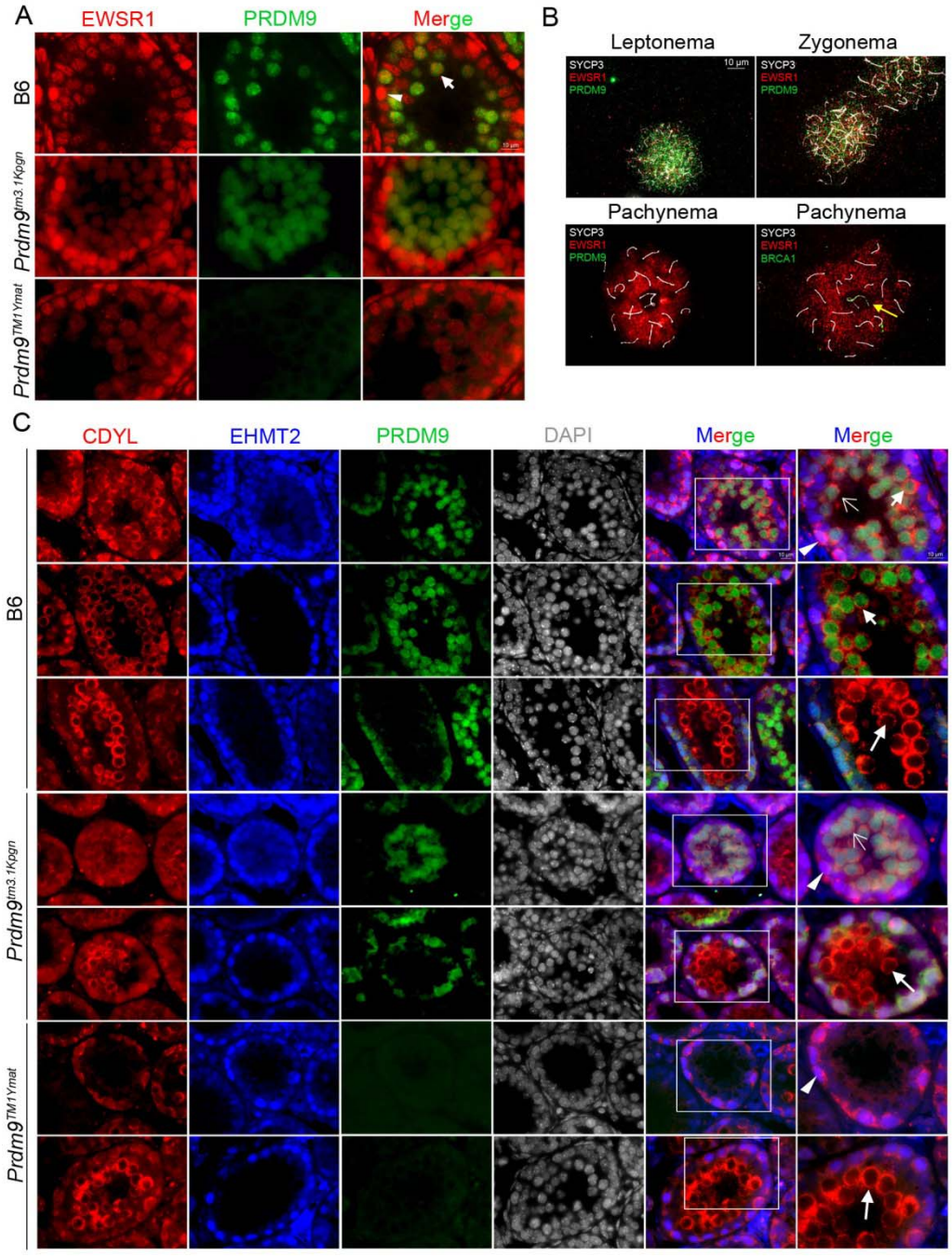


Fig. 5

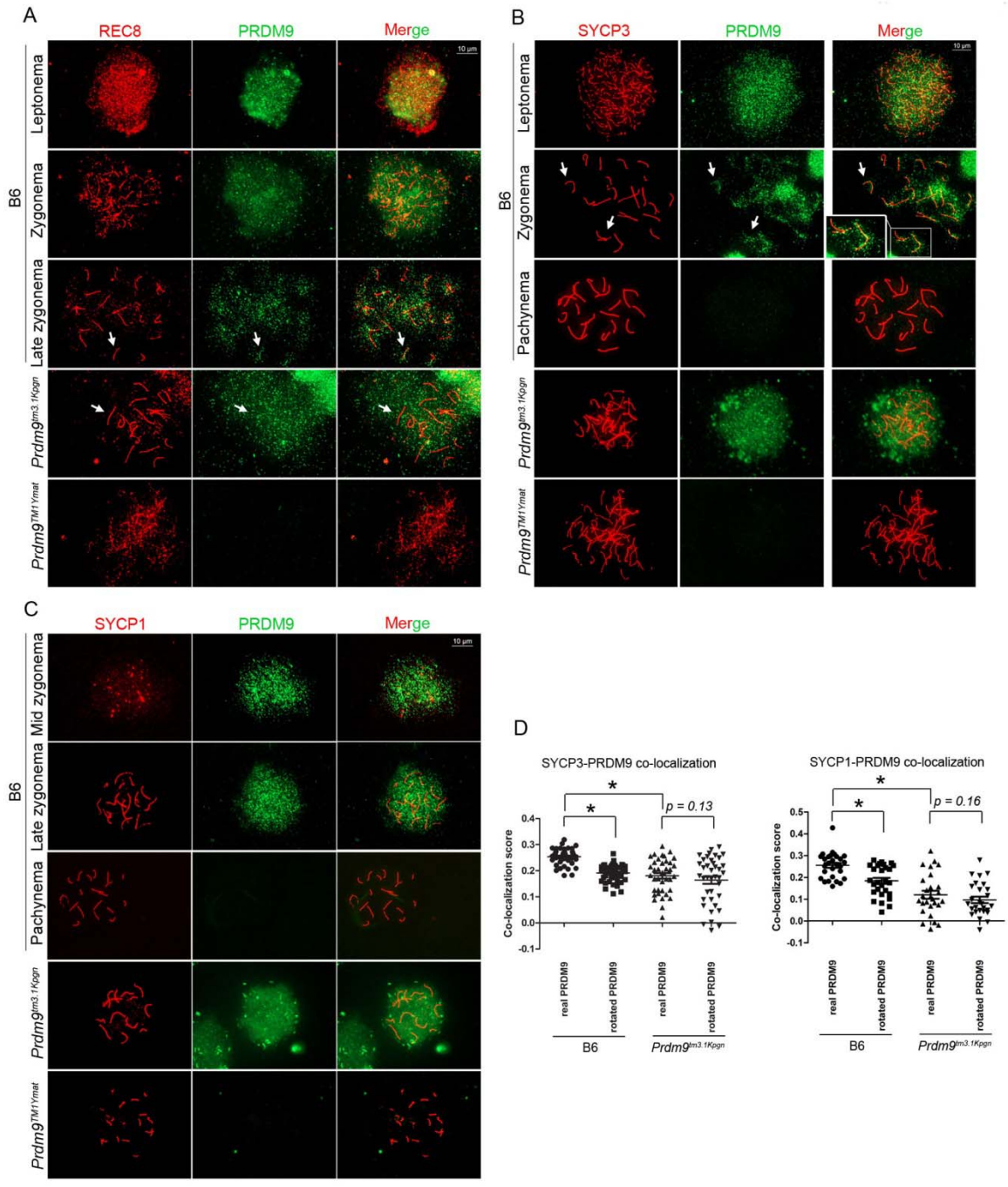


Fig. 6