Natural variation in the zinc-finger-encoding exon of Prdm9 affects hybrid sterility phenotypes in mice

Khawla FN AbuAlia¹, Elena Damm¹, Kristian K Ullrich¹, Amisa Mukaj², Emil Parvanov²,³,⁴, Jiri Forejt² & Linda Odenthal-Hesse*#

Affiliations:
¹Department Evolutionary Genetics, Research Group Meiotic Recombination, and Genome Instability, Max Planck Institute for Evolutionary Biology, Plön D-24306, Germany
²Laboratory of Mouse Molecular Genetics, Institute of Molecular Genetics, Czech Academy of Sciences, Vestec CZ-25250, Czech Republic
³Department of Translational Stem Cell Biology, Research Institute of the Medical University of Varna, 9002 Varna, Bulgaria.
⁴Ludwig Boltzmann Institute for Digital Health and Patient Safety, Medical University of Vienna, 1090 Vienna, Austria.

#Corresponding author:
E-mail: odenthalhesse@evolbio.mpg.de (LOH)

ORCID IDs: 0000-0002-7534-4014 (KFNA), 0000-0002-6261-6076 (ED), 0000-0003-4308-9626 (KKU), 0000-0001-6744-0949 (AM), 0000-0001-6744-0949 (EP), 0000-0002-2793-3623 (JF) & 0000-0002-5519-2375 (LOH).

Running Title:
Variability in Prdm9 affects hybrid sterility phenotypes

Keywords:
reproductive isolation, Mus musculus, fertility, asynapsis, Prdm9, Hstx2, t-haplotype
Abstract

PRDM9-mediated reproductive isolation was first described in the progeny of Mus musculus musculus (MUS) PWD/Ph and Mus musculus domesticus (DOM) C57BL/6J inbred strains. These male F1-hybrids fail to complete chromosome synapsis and arrest meiosis at prophase I, due to incompatibilities between the Prdm9 gene and hybrid sterility locus Hstx2. We identified fourteen alleles of Prdm9 in Exon 12, encoding the DNA-binding domain of the PRDM9 protein in outcrossed wild mouse populations from Europe, Asia, and the Middle East, eight of which are novel. The same Prdm9 allele was found in all mice bearing introgressed t-haplotypes, encompassing Prdm9 and inversions preventing recombination with wildtype Chr 17. We asked whether seven novel Prdm9 alleles in MUS populations and the t-haplotype allele in one MUS and three DOM populations induce Prdm9-mediated reproductive isolation. The results show that only combinations of the dom2 allele of DOM origin and the MUS msc1 allele ensure complete infertility of intersubspecific hybrids outside the context of inbred mouse strains. The results further indicate that the erasure of PRDM9 msc1 binding motifs may be shared by MUS mice from populations with different Prdm9 alleles, implicating that erased PRDM9 binding motifs may be uncoupled from their corresponding PRDM9 zinc finger arrays at the population level. Our data corroborate the model of Prdm9-mediated hybrid sterility beyond inbred strains of mice and suggest that sterility alleles of Prdm9 may be rare.

Introduction

Hybrid sterility is an evolutionary concept of reproductive isolation in which hybrid zygotes develop into healthy adults that fail to produce functional gametes and are thus sterile. In the “Bateson-Dobzhansky-Muller model of incompatibilities”, hybrid sterility occurs when two or more independently evolved genes are incompatible when interacting within an individual (BATESON 1909; DOBZHANSKY 1936; MULLER 1942). The first hybrid sterility locus identified in mammals was Hybrid sterility 1 (Hstf1) on chromosome 17 (FOREJT AND
At the Hst1 locus, the Prdm9 gene is responsible for the observed hybrid sterility and encodes PR domain-containing protein 9 (PRDM9) (Mihola et al. 2009). The PRDM9 protein is expressed in testicular tissue and fetal mouse ovaries during the early phases of meiotic prophase I when recombination is initiated (Hayashi et al. 2005; Lawson et al. 2011). PRDM9 has three conserved domains, an N-terminal KRAB domain that promotes protein-protein binding (Imai et al. 2017; Parvanov et al. 2017; Wang et al. 2021), an NLS/SSXRD repression domain with nuclear localization signal, and a central PR/SET domain that confers methyltransferase activity (Powers et al. 2016). The C-terminal domain is highly polymorphic and comprises an array of C2H2-type zinc fingers (ZNFs), which differ among PRDM9 variants in both type and number (Oliver et al. 2009; Baudat et al. 2010; Berg et al. 2010; Parvanov et al. 2010; Berg et al. 2011; Baudat et al. 2013; Buard et al. 2014; Kono et al. 2014). Variation among ZNFs is most pronounced in the amino acids at positions -1, 3, and 6 of the ZNF α-Helix, which are responsible for recognizing specific DNA target motifs (Billings et al. 2013; Baker et al. 2014; Walker et al. 2015; Altemose et al. 2017a; Altemose et al. 2017b; Patel et al. 2017). Additional amino acid substitutions in positions -5, -2, and 1 of the α-Helix are rarely seen (Parvanov et al. 2010; Kono et al. 2014). Since only amino acids at positions -1, 3, and 6 are involved in protein-DNA interactions, all other positions are not predicted to affect DNA binding affinity (Persikov and Singh 2014). Upon interaction with its specific DNA motif, the PR/SET domain of PRDM9 tri-methylates the adjacent nucleosomes on histone-3 by lysine-4 (H3K4) and lysine-36 (H3K36) (Hayashi et al. 2005; Wu et al. 2013; Eram et al. 2014), thereby triggering a cascade of events that initiate recombination, as reviewed in (Damm and Odenthal-Hesse 2022).

PRDM9-mediated reproductive isolation was discovered in male F1-hybrid progeny of Mus musculus (MUS) and Mus musculus domesticus (DOM) strains that differ in Prdm9 alleles (Forejt and Ivanyi 1974; Mihola et al. 2009). The PWD/Ph (hereafter PWD) MUS strain possesses the msc1 allele, and the C57BL6/J (hereafter B6) DOM
strain possesses the *dom2* allele. F₁ hybrid males from the cross between the PWD female and the B6 male do not complete chromosome synapsis and spermatogenesis arrests at meiotic prophase I, which prevents them from forming gametes and results in a postzygotic isolation barrier (MIHOLA *et al.* 2009; DZUR-GEJDOŠOVA *et al.* 2012; FLACHS *et al.* 2012). An X-linked locus *Hstx2*, located in a 2.7 Mb region on the proximal part of the X chromosome, modifies the effect of Prdm9 on the fertility of intersubspecific hybrids. *Hstx2* is structurally distinct between PWD and B6 mice and causes complete hybrid sterility only when the maternal *HstX2PWD* is active (BHATTACHARYYA *et al.* 2014; BALCOVA *et al.* 2016). The interaction between Prdm9 and the MUS *Hstx2PWD* allele promotes asynapsis of homologous chromosomes, ultimately leading to meiotic arrest (BHATTACHARYYA *et al.* 2013; BHATTACHARYYA *et al.* 2014; BALCOVA *et al.* 2016). In contrast, male F₁ hybrids of the reciprocal cross (B6 × PWD) carrying the *Hstx2B6* allele retain a low level of fertility. The *Hstx2* locus behaves as a recombination cold spot in crosses (BALCOVA *et al.* 2016), has reduced Prdm9-mediated H3K4me3 recombination initiation sites, and lacks DMC1-decorated DNA DSB hotspots (LUSTYK *et al.* 2019).

Defective pairing and high levels of chromosomal asynapsis are observed in hybrids with ineffective double-stranded break (DSB) repair (MIHOLA *et al.* 2009; DAVIES *et al.* 2016). It has been hypothesized that the molecular mechanism of PRDM9 action is related to the evolutionary divergence of homologous genomic sequences in DOM and MUS subspecies (DAVIES *et al.* 2016) and, more specifically, to the phenomenon of historical erosion of genomic binding sites of PRDM9 ZNF domains (DAVIES *et al.* 2016; FOREJT 2016; ZELAZOWSKI AND COLE 2016; FOREJT *et al.* 2021). Nucleotide polymorphisms within genomic target motifs may affect the binding affinity of PRDM9 in heterozygous individuals. The preferential formation of DSBs occurs on the haplotype with the motif with a stronger binding affinity (BAKER *et al.* 2015). However, since the uncut strand provides the template for repair, the less efficient motif is preferentially transmitted to the next generation, which can lead to the erosion of PRDM9 binding sites over time (JEFFREYS AND NEUMANN 2002). Therefore,
polymorphisms that reduce the binding affinity of a given PRDM9 variant are predicted to become enriched within populations over time, resulting in the attenuation of the hotspots (BOULTON et al. 1997). Direct evidence for over-transmission has been observed in human and mouse hotspots (JEFFREYS AND NEUMANN 2005; BERG et al. 2011; COLE et al. 2014; ODENTHAL-HESSE et al. 2014). In mouse hybrids, higher affinity binding sites for a given PRDM9 variant are four times more likely to be found on the chromosome of the other species, with which Prdm9 did not coevolve, suggesting that binding site erosion is a predominant factor driving hotspot loss in several mouse lineages (SMAGULOVA et al. 2016).

In mice, about 17.5% of hotspots have been eroded in the time it took for the PWD and B6 strains to diverge – averaging to roughly one PRDM9 binding site lost every 700 to 1500 generations (SMAGULOVA et al. 2016). Indeed, in sterile hybrids of inbred strains PWD (MUS) x B6 (DOM), a bias in initiation efficiency between diverged homologs is mainly driven by functional dom2 binding sites found on the PWD genome that are eroded on the B6 genome and vice versa for msc1 sites (DAVIES et al. 2016), with recombination being initiated at a large number of asymmetric sets of breaks (MIHOLA et al. 2009; DAVIES et al. 2016). In this cross, fertility was restored when the B6 PRDM9 zinc-finger array was replaced with the human variant B, making symmetric recombination hotspots predominant (DAVIES et al. 2016), further supporting the hypothesis that hybrid sterility is under an oligogenic control, with PRDM9 as the main factor.

Hybrid sterility also occurs outside of laboratory models as Prdm9 alleles found among MUS and DOM wild-derived inbred strains (PIALEK et al. 2008) showed fertility disruption in about one-third of the intersubspecific male hybrids (MUKAJ et al. 2020). Mice with Prdm9 alleles that were closely related to previously identified hybrid sterility alleles showed reduced sperm counts and low paired testes weights that were associated with high asynapsis rates of homologous chromosomes in meiosis I and early meiotic arrest (MUKAJ et al. 2020). Replacing Prdm9dom2 with the ‘humanized’ targeted Prdm9tm1(PRDM9)wthg (DAVIES et al. 2016) restored fertility in these mouse hybrids, supporting the role of PRDM9 as the
leading player in wild-derived inbred strains (MUKAJ et al. 2020). Furthermore, although
the exon 12 sequence of Prdm9msc5 had identical nucleotides in several strains, the
degree of fertility reduction observed differed between strains, and the effect of the
heterozygosity between genomic backgrounds remained unknown.

The relationship between the degree of chromosome asynapsis, meiotic arrest, and the
number of expected symmetric DSB hotspots per chromosome was reported in PWD × B6
hybrids (GREGOROVA et al. 2018). Asynapsis was shown to operate in-cis, depending on the
increased heterozygosity of homologs from evolutionarily divergent subspecies. Introducing
at least 27 Mbs of sequence homology belonging to the same subspecies (con-subspecific
homology) fully restored the synopsis of a given autosomal pair (GREGOROVA et al. 2018).
Prdm9msc1/dom2 also displayed a sterilizing effect in MUS x CAS hybrids, where the rate of
synapsis was proportional to the level of non-recombining MUS genetic background
(VALISKOVA et al. 2022).

To date, complete sterility has been observed in hybrids of inbred strains where hotspot
erosion is exacerbated and the same strain-specific allele activates all hotspots (DAVIES et al.
2016; SMAGULOVA et al. 2016). Through the process of inbreeding, a genome that initially
possessed wildtype levels of heterozygosity becomes sequentially homozygous, and this
process should equally affect PRDM9 binding sites. However, as soon as a strain is fully
inbred, there should be no more heterozygous binding motifs, and erosion should cease
totally or at least slow down substantially, as it is now based only on rare mutations with
stochastic placement within binding sites. Substantial erosion of binding sites must have,
therefore, occurred before inbreeding, likely as a result of the high frequency of msc1 alleles
in wild populations (BUARD et al. 2014; KONO et al. 2014; FOREJT et al. 2021). In natural
populations, Prdm9 evolves rapidly, with protein variants behaving like the predator and
specific motifs as prey following Red Queen dynamics to avoid negative selection of a
complete loss of recombination hotspots over time (ABE et al. 2004; LATRILLE et al. 2017).
However, while the \textit{Prdm9} gene shows remarkable natural allelic divergence, with more than 150 alleles found in mouse populations to date (Buard et al. 2014; Kono et al. 2014; Vara et al. 2019; Mukaj et al. 2020), little is known about how many of these alleles are hybrid-sterility inducing, nor about their DNA binding motifs and their level of erosion.

Furthermore, the fertility of F\textsubscript{1} hybrids could thus also be modified by additional hybrid sterility loci. At least three autosomal polymorphic hybrid sterility factors exist between PWD and STUS strains (Bhattacharyya et al. 2014). Five \textit{Prdm9}-dependent quantitative trait loci have been identified in intersubspecific (MUS x CAS) hybrids, segregating on DOM background (Valiskova et al. 2022). However, not only do laboratory intercrosses between wild-derived inbred strains differ from the pure form of hybrid sterility observed in (PWDxB6) laboratory crosses, but contrasting patterns are also observed in wild mice.

The natural hybrid zone is a relatively recent secondary contact zone across Europe, where only a third of all house mouse males exhibit fertility traits below the range of the pure subspecies. Complex polygenic control of hybrid sterility has been observed, and several interchangeable autosomal loci have been proposed to be sufficient to activate the Dobzhansky-Muller incompatibility in wild mouse hybrids (Dzur-Gejdosova et al. 2012; Turner and Harr 2014). A genome-wide association study (GWAS) revealed strong interactions between Ch17 and Chr X, but most of these loci were located outside of \textit{Prdm9} and \textit{Hstx2} (Turner and Harr 2014).

Naturally occurring chromosome 17 haplotypes, the \textit{t} haplotypes (Silver 1985) also strongly influence male fertility in wild mice. Males heterozygous for the \textit{t}-haplotype pass it on to more than half of their offspring, with some variants presenting transmission rates over 90%, while females transmit the \textit{t}-haplotype within the expected Mendelian ratio (Lyon 2003). Despite a strong drive, \textit{t}-haplotypes are only present in 10-40% of all populations of wild house mice, presumably because they also include genes causing male infertility and embryonic lethality (Olds-Clarke 1997; Planchart et al. 2000;
SCHIMENTI et al. 2005; KELEMN AND VICOSO 2018). Several distorter loci of the t-haplotype act in trans to impair motility in wildtype spermatozoa. They over-activate a signaling pathway controlling sperm motility kinase (SMOK), resulting in abnormal flagellar movements and loss of sperm motility (HERRMANN et al. 1999). To prevent the distorter locus from affecting t-haplotype-bearing sperm, a responder locus, consisting of a fusion gene of a SMOK member and the 3’-UTR of the triple ribosomal s6 gene, confers partial cis-resistance to overactivation by the distorters, thereby restoring sperm motility (HERRMANN AND BAUER 2012). The t-haplotype consists of 30 Mb of introgressed sequence transferred from an unidentified Mus ancestor in the Mus musculus subspecies over one million years ago (HAMMER AND SILVER 1993) and encompasses the Prdm9 locus (TRACHTULEC et al. 2008), with the most diverse allele of the Mus musculus subspecies identified to date (Kono et al. 2014). However, it remains unknown whether Prdm9 contributes to the observed reduction in fertility associated with t-haplotypes.

In summary, the low incidence of sterile wild-mouse hybrids in DOM/MUS natural hybrid zone (TURNER et al. 2012) together with the reported large number of hybrid sterility loci in intersubspecific backcrosses and intercrosses contrasts with the F1 hybrid sterility model based on the Prdm9 allelic incompatibility, Hstx2, and background heterozygosity in PRDM9 binding sites. Further experimental evidence is needed to understand the mechanism of Prdm9-driven hybrid sterility and its role in wild mouse populations. Here, we ask whether hybrid sterility is under the PRDM9 control in wild mice beyond the context of inbred strains of mice. If the Prdm9-driven hybrid sterility is linked to the asymmetric erosion of PRDM9 binding motifs, then in a simple scenario, the sterility-inducing alleles would be expected to be ancestral alleles situated closest to the common ancestor on the phylogenetic tree. To test this hypothesis, we examine the evolutionary relationship between the known Prdm9 hybrid sterility alleles and any newly identified allele.

**Material and Methods**
Mice

All work involving experimental mice was performed according to approved animal protocols and institutional guidelines of the Max Planck Society and with permits obtained from the local veterinary office ‘Veterinäramt Kreis Plön’ (permit number: 1401-144/PLÖ-004697). Mice, including strains of PWD/Ph strain, C57/Bl6 strain with transgene Prdm9tm1.(PRDM9)Wthg strain and consomic C57BL/6J-Chr X.1sPwD/Ph/ForeJ mice, as well as several wild mice populations were all maintained in the mouse facilities of the Max Planck Institute for Evolutionary Biology in Plön, following FELASA guidelines and German animal welfare law. We analyzed three outcrossed populations of DOM mice; first, the French Massif-Central (MCF) population, founded in December 2005 with a starting population size of sixteen breeding pairs, with additional wild-caught animals introduced into the breeding population at the beginning of April 2010. These mice were in generation sixteen at the start of this experiment, nine generations since crossing in with the second set of new wild-caught animals. The German Cologne-Bonn (CBG) population was founded in August 2006 with ten breeding pairs and maintained as an outcross for fourteen generations at the start of this experiment. In November 2012, new wild-caught breeding pairs were crossed in. The Iranian population from Ahvaz (AHI) was started in December 2006, with six founding breeding pairs. It has been maintained in an outcross for fifteen generations at the start of this experiment, after two rounds of reduction due to inbreeding depression. Seventeen breeding pairs of mice initially trapped in Almaty, Kazakhstan (AKH) in December 2008, founded the Mus musculus musculus population, which had been maintained for thirteen generations at the beginning of this experiment. Whole-genome sequencing and transcriptomic data of multiple individuals from each population are publicly available in the European Nucleotide Archive and as custom tracks in the UCSC genome browser (HARR et al. 2016).

Organ withdrawal
Organ withdrawal after euthanasia is not legally considered an animal experiment according to §4 of the German Animal Welfare Act. It, therefore, does not need to be approved by the competent authority (Ministerium für Landwirtschaft, ländliche Räume, Europa und Verbraucherschutz). F1 hybrid males were euthanized after being first rendered unconscious by deliberately introducing a specific CO2/O2 mixture ratio, then sacrificed using CO2 euthanasia followed by cervical dislocation. To reduce loose hair contaminating the organs during the dissection of the animal, their coat was sprayed with 75% EtOH before organ withdrawal. Spleen, a liver lobe, and both testes were extracted, and epididymides were removed. One epididymis was placed in 500 µl of cold phosphate-buffered saline for sperm counting, and all other organs were immediately snap-frozen in liquid nitrogen and stored at -70°C.

Fertility phenotyping

We collected three fertility parameters, body weight (BW) and paired testes weight (TW), and spermatozoa released from epididymal tissues, counted in Million/ml (SC). One epididymis, including caput, corpus, and cauda, was repeatedly cut in 1 ml of cold phosphate-buffered saline to release spermatozoa. The tube was vigorously shaken for 2 minutes, and spermatozoa in the solution were diluted to 1:40 in PBS. We counted 10 µl of diluted spermatozoa in a Bürker chamber (0,1 mm chamber height), where two replicates of 25 squares were counted. In cases when only a few (<10) spermatozoa were found, additional dilutions were prepared and counted. We added the two replicated 25 squares counts (A25 + B25) from spermatozoa released from a single epididymis to approximate spermatozoa released from a pair of epididymides. The epidydimal spermatozoa count released in 1 ml PBS was then calculated by taking the paired counts, the volume of 25 squares (V25=0,02*0,02*0,01*25=0,0001 cm³), and the dilution factor into account.

Spreading and Immunofluorescence analyses of spermatocytes
Spermatocyte nuclei were spread for immunohistochemistry as described in (Anderson et al. 1999), with the following modifications. Firstly, a single-cell suspension of spermatogenic cells from the whole testis was prepared in 0.1 M sucrose solution. The sucrose-cell slurry to which protease inhibitors (Roche 11836153001) were added, then dropped onto paraformaldehyde-treated glass slides. Glass slides were kept in a humidifying chamber for 3 hours at 4 ºC to allow cells to spread and fix. Slides were briefly washed in distilled water and transferred to pure PBS before blocking in PBS with 5-vol% goat serum. Primary antibodies HORMAD2 (a gift from Attila Toth, rabbit polyclonal antibody 1:700), SYCP3 (mouse monoclonal antibody, Santa Cruz, #74569, 1:50), yH2AX (ab2893. 1:1000), and CEN (autoimmune serum, AB-Incorporated, 15-235) were used for immunolabelling. Secondary antibodies goat anti-Mouse IgG-AlexaFluor568 (MolecularProbes, A-11031), goat anti-Rabbit IgG- AlexaFluor647 (MolecularProbes, A-21245), goat anti- Human IgG-AlexaFluor647 (MolecularProbes, A-21445), goat anti- Rabbit IgG-AlexaFluor488 (MolecularProbes, A-11034) were used at 1:500 concentration at room temperature for one hour. A Nikon Eclipse 400 microscope with a motorized stage control was used for image acquisition with a Plan Fluor objective, 60x (MRH00601). Images were captured with a DS-QiMc monochrome CCD camera and the NIS-Elements program (from Nikon). Image J software was used to process the images.

**Prdm9 Genotyping**

We used ear clips taken at weaning to identify *Prdm9* allelic variation in the wild mouse populations. All F1 and F2 hybrid offspring used in the experiments were instead genotyped from the counted sperm sample taken from one of the epididymides. Furthermore, we confirmed initial parental PRDM9 genotyping after successful mating (> 5 male offspring) by sacrificing all F0 males. All genotyping was done on individual mouse IDs, but in such a way that the experimenter was blind to the matching fertility phenotypes.
DNA extraction

DNA was extracted from ear clips or whole ears using salt extraction. Briefly, cells were lysed in SSC/0.2 % SDS, and proteins were digested using Proteinase K (20 mg/µl), incubating at 55 °C overnight. We salted out the DNA using 4.5 M NaCl solution, followed by two consecutive rounds of Chloroform extraction. The DNA was then Ethanol precipitated and washed twice with 70 % ethanol, and the pellet was then dried at room temperature and finally dissolved in 30 µl Tris-EDTA pH 8.0. The DNA samples were stored at 4 °C for short-term and - 70 °C long-term storage. The slurry of isolated spermatozoa with epididymal tissues was processed similarly; however, to lyse sperm heads and remove Protamines, we increased the SDS concentration to 1 % and added not only Proteinase K (20 mg/µl) but also TCEP (Thermo Scientific 77720, 0.5 M) to a final concentration of 0.01 µM. This extraction method produces a mixture of DNA extracted from somatic and sperm cells.

Amplification of the minisatellite coding for the zinc-finger array of PRDM9

The ZNF arrays of each mouse were PCR amplified similarly as in (Buard et al. 2014) on 10-30 ng of genomic DNA in 12 µl reactions of the PCR buffer “AJJ” from (Jeffreys et al. 1990) using a two-polymerase system with Thermo Taq-Polymerase (EP0405) and Stratagene Pfu Polymerase (600159) to ensure high-fidelity PCR. When offspring are heterozygous for two alleles of different lengths (in most cases), we separated heterozygous bands after gel electrophoresis on Low Melting agarose (Thermo Fischer #R0801) by excising the bands and eluting the DNA using Agarase (Thermo Fischer #EO0461). If two heterozygous bands were apparent, excised and eluted product was immediately used in sequencing reactions after estimating the amount of DNA from the gel. If only one band was evident, alleles were not separated by size. Therefore, the purified PCR products were cloned using TOPO TA Cloning Kit for Sequencing (Life
Technologies no. 450030), following the manufacturers’ specifications before sequencing.

We analyzed at least eight clones per sample.

Sequencing

Sequencing reactions of either eluted PCR product or picked clones were set up using BigDye 3.0, according to the manufacturer’s protocol, then purified using X-terminator, and finally sequenced using 3130x/ Genetic Analyzer. Only PRDM9 variants with less than 12 ZNFs could be sequenced to their ends in both directions. Exon 12 of Prdm9 was fully sequenced for all alleles; however, forward and reverse sequences overlapped along the entire length of the exon only in alleles smaller than <1000 bp, such that larger alleles had sequence stretches only covered by either forward or reverse sequencing. Nevertheless, the sequencing products of all alleles still provided sufficient overlap for full-length assembly. We assembled the forward and reverse sequences based on the estimates of fragment sizes from PCR products on gels to accurately assemble the coding minisatellite using Geneious Software (Version 10.2-11). After sequencing and alignment, assembled minisatellites were conceptually translated into the amino acid sequence of the ZNF domain, and HMMER scores were computed using a Polynomial SVM (PERSIKOV AND SINGH 2014).

Phylogenetic Analyses

The phylogeny on all alleles tested for hybrid sterility phenotypes in (MUKAJ et al. 2020) and this publication was computed using the R package “repeatR” from https://mpievolbio-it.pages.gwdg.de/repeatr/(DAMM et al. 2022). Briefly, minisatellite-like repeats within the gene are identified, extracted, and filtered for incomplete sequences before matrices based on minimum edit distance (Hamming) were computed using weighting costs \( w_{\text{mut}} = 1, w_{\text{indel}} = 3.5 \) and \( w_{\text{slippage}} = 1.75 \) as given in (VARA et al. 2019). These Minimum edit distances represent a metric on the set of changes between Prdm9 minisatellite repeat units of 84
bp in length. As such, it can be used as a measure of genetic distance. We computed two
distance matrices for each type of repeat, as in (DAMM et al. 2022). The first distance
matrix included all nucleotides, while the second matrix excluded nucleotides known to be
under positive selection, which are coding for the hypervariable amino acids responsible
for DNA binding specificity (-1, +3, +6). Two phylogenetic reconstructions of
the Prdm9 hypervariable region were then computed separately from both matrices, using
a neighbor-joining approach with the “bionj” function of the R package ape 5.0 (PARADIS
AND SCHLIEP 2019) and rooted on the “humanized” Prdm9 allele from (DAVIES et al. 2016).

Genotyping for Chr17 t-haplotype and X-chromosomal haplotypes near Hstx2.

The presence of the t-haplotype was tested using markers Tcp1 and Hpa-4ps (PLANCHART
et al. 2000), and X-chromosomal haplotypes across the refined Hstx2 interval were tested
using primers in Table S2 from (LUSTYK et al. 2019). Each forward primer was labeled
with either HEX or FAM and amplified using the ABI Multiplex Kit according to the
manufacturer’s protocol. Fragment lengths were then analyzed by capillary
electrophoresis using a 3730 DNA Analyzer. Allele sizes were scored and binned using
the Microsatellite plugin in Geneious v.10.2.

PRDM9 in-silico DNA binding predictions

For in-silico DNA motif binding predictions, the nucleotide sequence was first conceptually
translated into a protein sequence, and the C2H2 zinc-finger binding predictions were
computed using a polynomial kernel with the method of (Persikov and Singh 2014). We
converted the matrices from (Persikov and Singh 2014) to tab MEME, JASPAR, and
STAMP input files by using the RSAT matrix conversion tool (Santana-Garcia et al.
2022)(http://rsat.sb-roscoff.fr/convert-matrix_form.cgi), choosing the reverse complement
option. We used MEME input files for TomTom and STAMP input files for STAMP and
JASPAR input files for PWMScan. The thus computed JASPAR files were inputted into
PWMScan (Ambrosini et al. 2018) to find binding site predictions on the *Mus musculus* reference genome mm10 (which most closely resembles the genome of the C57Bl6/J strain). The .bed files containing the genome-wide putative binding sites of each PRDM9 variant were compared using bedtools intersect (Quinlan and Hall 2010), reporting each incident where bed files overlapped for at least one base pair.

**Statistical Analyses**

The majority of graphs, calculations, and statistical analyses were performed using GraphPad Prism software version 9.4.1 for Mac (GraphPad Software, San Diego, CA, USA). Statistical tests are stated in the text and the Figure legends. Briefly, pairwise comparisons were performed using unpaired t-tests with Welch correction, and as we did not assume equal sample variances *a priori*, these were compared using F-tests. Similarly, multiple comparisons of fertility parameters were first evaluated for differences in sample variances using Brown-Forsythe ANOVA tests. If significant differences between means were observed, we performed Welch’s ANOVA with Dunnett’s T3 multiple comparisons test. When there was no indication of unequal sample variance, we performed ordinary one-way ANOVA instead, which we evaluated with Bonferroni multiple comparisons tests. Asynapsis data were compared between genotypes using unpaired t-tests with Welch correction, and linear correlation was assessed using the Pearson correlation coefficient (r). Using the binomial probability calculator on the VassarStats: Website for Statistical Computation (http://vassarstats.net), we tested for transmission ratio distortion.

**Results and Discussion**

Previous analyses had identified four alleles of *Prdm9* that induced hybrid sterility in wild-derived inbred strains of mice initially trapped in Europe (MUKAJ et al. 2020; FOREJT et al. 2021). We screened additional European wild mice further away from the hybrid zone and
mice from Asia and the Middle East for novel Prdm9 alleles. Mice were initially caught by
(HARR et al. 2016), MUS were initially trapped in Almaty, Kazakhstan (43°16'N, 76°53'E),
and DOM in three different locations: the city of Ahvaz, Iran (31°19' N, 48°42' E), the
Massif-Central area in France (45°32'N, 2°49'E) and the Cologne-Bonn area in Germany
(50°52'N, 7°8'E). Previous observations of diverse haplotypes in the Iranian basin
(HARDOUIN et al. 2015) and demographic analyses of the source populations also confirmed
the AHI population as the most ancestral (FUJIWARA et al. 2022). These populations have
been housed and maintained as outcrosses for many generations before this study (see
Materials and Methods). Despite high degrees of relatedness, these populations have
maintained a much larger genomic diversity than inbred strains (LAWAL et al. 2021). These
outcrossed populations show low levels of introgression between MUS/DOM (Ullrich et al.
2017) and moderate levels of bidirectional introgression patterns from Mus spretus in all
three DOM populations (Banker et al. 2022). These outcrossed populations of mice with
inter-individual genetic diversity have high average SNP densities compared to the
C57BL/6 strain, with the number of population-private variants and genomic introgression
for each population collected in (Table S1). The distribution of original trapping locations of
all mice tested for hybrid sterility phenotypes in this study and in (MUKAJ et al. 2020) are
shown in Figure S1. Testis mRNA expression levels are available for multiple individuals
of each outcrossed population (Harr et al. 2016) and have demonstrated a robust Prdm9
expression (KELEMEN et al. 2022). In summary, these genetically diverse individuals from
several outcrossed populations provide a unique resource to evaluate the Prdm9 allelic
incompatibility-mediated hybrid sterility model beyond the context of inbred strains of
mice. We screened all populations for individual Prdm9 alleles by sequencing exon12 of
Prdm9 containing the minisatellite coding for the C2H2 zinc-finger domain of PRDM9 as
described in (BUARD et al. 2014; KONO et al. 2014). The amino-acid variation between ZNF
domains was determined by conceptually translating the nucleotide sequence of satellite
repeats into the amino-acid sequence of individual zinc fingers (Figure S2 and Figure S3).
We defined the ZNF variation based on the amino acids at positions −1, +3, and +6 of the
alpha-helix, representing the DNA contact residues (depicted in the cartoon of Figure 3A).

Based on variation in nucleotide repeats and the composition, order, and number of repeats, we have identified eight full-length Prdm9 alleles in MUS mice from Kazakhstan and six in DOM from France, Germany, and Iran (shown in Figure S4B together with previously identified alleles from (MUKAJ et al. 2020)). We named these Prdm9 alleles according to the International Committee of Standardized Genetic Nomenclature for mice (MGI) (in Figure 4 and Figure S4B), registered them at JAX, and submitted their sequences to GenBank under Accession numbers (OQ055171-OQ055188).

We found a single, peculiar Prdm9 allele in MUS and DOM subspecies in all four original trapping locations associated with t-haplotypes. In this allele, nine nucleotides are deleted in the translated amino-acid sequence of the first ZNF of the array, removing three amino acids, including one of the zinc-binding Cysteine ligands (Figure S2). In addition, distinct amino acids are seen in positions −1, +3, and +6, not present in any other PRDM9 variants, such as TDK and ASQ, with additional differences between the two types of ASQ ZNFs in positions -8 and +5. Similarly, the ANQ ZNF found in mice with t-haplotypes differs from the ANQ found in mice without t-haplotypes at position -2, a pattern previously seen in mice with t-haplotypes (KONO et al. 2014). We, therefore, tested additional mice with t-haplotypes for Prdm9, including Mus musculus castaneous (CAS) from Taiwan and mouse strain T/t^d (FOREJ T et al. 1988), confirming that all possessed the same Prdm9 allele.

Given that a single Prdm9 allele was found in all mice carrying a t-haplotype, regardless of subspecies, we consider it an intraspecies Mus musculus t-haplotype Prdm9 allele "Prdm9^mmt1". Together with t-haplotypes, the mmt1 allele was always present in a heterozygous state and occurred in all of our outcrossed populations at high frequencies. We found a t-haplotype in 50% of the MUS population from Almaty (Kazakhstan), in 88% of the DOM population from Ahvaz (Iran), in 90% of the mice from Cologne-Bonn (Germany), and 100% of the mice from the Massif-central (France) population. Even though initial population
frequencies may have gotten heavily distorted due to the t-haplotype over-transmission, and
the alleles may not reflect the initial population frequencies in which they occurred in the wild.
However, the alleles we identified in the outcrossed population should still reflect the Prdm9 alleles in the wild. While most alleles are novel, the mmt1 allele (KONO et al. 2014) and three MUS alleles which we found in mice from Kazakhstan had been identified previously, such as msc6, previously named Ma8 and identified in Grozny, Russia (KONO et al. 2014), msc12 allele identified as 7mus1 in MUS from (BUARD et al. 2014), and msc9 found in strains CHD and BLG2, named Ma12 (KONO et al. 2014). A single DOM allele has been previously identified as 16dom1 (BUARD et al. 2014) in the DOT strain, originally from Tahiti (French Polynesia), which we found in Ahvaz, Iran, as dom8. All alleles in this study, except the mmt1 allele, are present only in one population, yet some closely resemble alleles of other subspecies. For example, msc11 closely resembled two alleles from the CAS subspecies, firstly, the classical cst1 allele, which possesses Serine at position -1 of the alpha-helix of the 8th ZNF (PARVANOV et al. 2010); in this position, msc11 has a single amino-acid change to Asparagine, and secondly, a CAS trapped in Nowshahr, Iran, Ca1 (KONO et al. 2014), differs only by a single amino-acid substitution in position 6 of the alpha-helix of the 6th ZNF, where Ca1 possesses Glutamine, instead of Lysine. Alleles msc7 and msc10 appear similar and share similarities to MUS 27mus1 trapped in Bulgaria (BUARD et al. 2014) and CAS Cc4 trapped in Grozny, Russia (KONO et al. 2014). The observations of CAS-like alleles in MUS populations are consistent with observations of many “MUS-like CAS” and “CAS-like MUS” samples in genomic datasets that show admixture patterns (FUJIWARA et al. 2022).

Testing Prdm9 alleles for sterility phenotypes in intersubspecific hybrids

To investigate whether newly identified wild Prdm9 alleles induce hybrid sterility in intersubspecific crosses, we adapted the crosses used in the laboratory models of hybrid sterility to eliminate possible variation due to the Hstx2 modifier. PWD (MUS) females were crossed with wild DOM males to test DOM alleles, emulating the PWD × B6 laboratory model. To test MUS alleles, we emulated the B6.DX1s × PWD laboratory model (MUKAJ et
al. 2020) and crossed wild MUS males to C57BL/6J-Chr X.1s<sup>PWDPh</sup>/ForeJ females (abbreviated B6.DX1s). B6.DX1s is a consomic DOM strain of C57BL/6/J background that carries the Hstx2 locus within a 69.6 Mb PWD sequence of the proximal end of the X-chromosome, essential for F<sub>1</sub>-hybrid sterility (Bhattacharyya et al. 2014; Balcova et al. 2016; Lustyk et al. 2019; Forejt et al. 2021).

All nine DOM sires and five of the sixteen MUS sires possessed t-haplotypes. The t-haplotype is a known meiotic driver, skewing transmission against wildtype Chr17 in the male germline. We, therefore, experienced a severe reduction of testable Prdm9 alleles on wildtype Chr17 in both MUS and DOM mice because of the over-transmission of the Prdm9<sup>mmt1</sup> allele within the t-haplotype. Indeed, more than 87% of offspring from fathers with t-haplotypes inherited the mmt1 allele, a significant deviation from the Mendelian 50:50 transmission ratio (two-tailed binomial probability P< 0.000001, approximated via normal). In contrast, all other Prdm9 alleles were transmitted at 50:50 in mice without t-haplotypes. To estimate the effect of Prdm9 on the fertility of hybrids, we used the weight of paired testes and the number of spermatozoa in paired epididymis as a proxy. Since these fertility parameters differ with the genetic background (Widmayer et al. 2020), we first looked into the physiological variation of the outcrossed source populations of wild mice as a control, which we compared with inbred mice (Figure S6). Wild mice from all three DOM populations had comparable sperm counts, while testes weights of CBG and MCF populations were significantly higher than in B6 mice. Similarly, wild MUS had elevated testes weights compared to PWD. Outcrossed mice also possess large variability in body weight, and we noticed that testes weight correlated with sperm count more robustly than testis weight normalized to animal body weight (Figure S7). Because wild mouse sires came from outcrossed populations, they were almost always heterozygous for different Prdm9 alleles. Several males from the same population shared one of the Prdm9 alleles. As a result, offspring from different crosses with such fathers acquired the same combination of Prdm9 alleles. We used them to determine whether differing genetic
backgrounds affect fertility parameters in F₁ hybrid offspring with the same Prdm9 allelic combination inherited from different fathers. We did not observe a significant background effect on sperm counts (Figure S8A), except in offspring that inherited t-haplotypes (Figure S8B). Paired testes weights, however, were significantly different in MUS males with msc6 and msc10 alleles and hybrid males carrying t-haplotypes (Figure S8). We analyzed the fertility parameters for all mice aged 60-100 (±2) days, as hybrid fertility parameters can vary with age, with a marked decline after 20 weeks (WIDMAYER et al. 2020). We performed regression analyses of age and fertility for each Prdm9 allele separately and for all alleles combined and detected no apparent effect of age on fertility in the tested age ranges, with only a weak positive correlation of age on testis weights in offspring with the mmt1KH and msc10 alleles (Figure S9). When we pooled fertility parameters of F₁-hybrid males by Prdm9 genotype, almost all had significantly higher TW and significantly elevated SC compared to control hybrid sterility crosses PWD × B6 and B6.DX1s x PWD. Exceptions are intersubspecific hybrids of PWD x DOM crosses that inherited the paternal mmt1MC allele and intersubspecific hybrids of B6.DX1s x MUS crosses that inherited the paternal msc11 and msc12 alleles.
Table S 3 and
Table S 4), whose SC was not significantly elevated compared to msc1. The hybrids inheriting paternal msc12 showed significantly reduced fertility compared to other tested alleles (Figure 1). Given that all F₁-hybrid males were either completely fertile or showed only reduced fertility, we wanted to test whether they would display chromosomal asynapsis, a hallmark characteristic of Prdm9-dependent hybrid sterility (FOREJT AND JANSA 2023). We performed immunofluorescence analyses on spermatocyte spreads of F₁-hybrids that inherited alleles msc10, msc11, and msc12 and hybrids that inherited the mmt1 allele in combination with both DOM and MUS t-haplotypes.

To determine if cells are in the pachytene stage when the synaptonemal complex is fully formed, we immunostained synaptonemal-complex protein 3 (SYCP3). In contrast to continuous SYCP3 staining in pachytene, SYCP3 staining is disorganized and patchy in the preceding late zygotene stage (when the synaptonemal complex is still forming) and the succeeding early diplotene stage (when the synaptonemal complex begins to disassemble as autosomes de-synapse and lateral elements separate), when the SYCP3 signal becoming visible as pair of thinner threads (DE LA FUENTE et al. 2007). Additionally, we assessed H2AX, which localizes to the X and Y chromosomes in males only at mid-to-late pachytene, forming a punctate “sex body” of transcriptional silencing (FERNANDEZ-CAPETILLO et al. 2003). We evaluated 48-113 of such pachynemas for asynapsis in each individual, scoring each HORMAD2 stained element (excluding sex chromosomes) as one asynapsis event. We determined the percentage of cells with asynapsis (and collected all data in the linked Dryad Repository). The F₁ hybrid males with wild MUS alleles msc11 and msc12 had an elevated proportion of asynaptic pachynemas (msc11, 42.4 ± 8.0 % and msc12 57.2 ± 12.5%), these frequencies are significantly elevated, compared to control mice of pure parental subspecies (B6.DX1s and PWD), with asynapsis levels of msc12 statistically indistinguishable from those of msc1 (Figure 1C). In contrast, hybrids with wild mouse allele msc10 had lower asynapsis, averaging 16.6 ± 8.6% of cells with asynapsis, not significantly different from fertile controls and significantly lower than sterile.
controls (Figure 1C). Similarly, mice with t-haplotypes also showed low asynapsis rates (DOM mmt1 11.5 ± 12%, MUS mmt1 17.4 ± 18.0%). We also tested the relationship between fertility parameters correlated and the meiotic asynapsis rate across all genotypes. We observed a weak but significant correlation between high asynapsis with both low sperm count (Pearson R² = 0.46, P < 0.0001) (Figure 1E) and paired testes weights (Pearson R² = 0.23, P= 0.0052) (Figure 1F). In summary, the tested Prdm9 allelic combinations were either completely fertile or showed a reduction of fertility. As hybrids with msc11 and msc12 alleles had the lowest sperm counts and paired testes weights and also showed the highest levels of chromosomal asynapsis, we can conclude that Prdm9 likely drives this effect, but any Prdm9-dependent effect is either not binomial, affected by overall genomic heterogeneity, or by potential genetic modifiers on the wild genetic backgrounds.

The relationship between Prdm9 genotype and wild-derived outbred genetic background

According to the hypothesis linking the Prdm9-driven hybrid sterility to the asymmetric erosion of PRDM9 binding motifs, the sterility of F₁ hybrids results from the erosion of MUS PRDM9	⁰⁻⁸⁻ binding sites in the PWD genome and DOM PRDM9⁷⁻⁸⁻ sites in the B6 genome (Davis et al. 2016). If a given MUS population with an unrelated “fertility” Prdm9 allele carries the same msc1 binding motif as PWD mice, then the fertility of (PWD x MUS) x B6 hybrids would segregate according to Prdm9; Prdm9×msc1 males would be sterile, and their Prdm9×MUS siblings would be fertile. However, if the given wild MUS genome does not carry any PRDM9 binding site erasure or the erased motifs match an unrelated Prdm9 allele, then Prdm9×msc1 and Prdm9×MUS male progeny should be fertile. The same assumption can be made for the outcome of PWD x (B6 x DOM) testcross to test the coupling of eroded PRDM9 binding motif with corresponding PRDM9 allelic zinc finger domains. To test the decoupling of eroded PRDM9 binding sites from the corresponding allelic form of Prdm9 in wild populations, we generated intraspecific (DOM x B6) hybrid males and crossed them to PWD females to
test the DOM-derived mmt1 and dom2 alleles against the mixed DOM background. To test MUS alleles and their genetic background, the B6.DX1 females were crossed with intraspecific (PWD × MUS) F1 hybrid males. We also performed an analogous cross in reciprocal orientation for MUS alleles, where (PWD × wild MUS) F1 hybrid females were crossed with B6 males. In the female germline, the mmt1 allele showed even transmission (two-tailed exact binomial probability, \( P = 0.136 \)), confirming male-specific \( t \)-haplotype transmission distortion (LYON 2003). Since not only \( Prdm9 \) but also X-chromosomes segregate in this cross, we only included males with the PWD haplotype containing the refined \( Hstx2^{PWD} \) locus, as defined by alleles of the X chromosomal microsatellite markers SR51, SX69084, and SX65100 (LUSTYK et al. 2019)(see also Figure S11). Siblings that inherited the \( Prdm9 \) alleles msc6, msc7, msc8, or the mmt1 of MUS or DOM all displayed fertility phenotypes within the physiologically normal range. In contrast, their brothers with the allelic combination msc1/dom2 were sterile (Figure 2), with testes' weight and sperm count comparable to (PWD x B6) F1 males with the same allelic combination. These results suggest that mice from MUS populations with different \( Prdm9 \) alleles may share the same pattern of erased \( PRDM9^{msc1} \) (PWD) binding motifs, implicating uncoupling of the erased \( PRDM9 \) binding motifs and \( PRDM9 \) zinc finger arrays at the population level. Furthermore, the partially wild-derived outbred background does not appear to carry additional major genetic modifiers that would prevent sterility per-se.

Variation of X-chromosomal haplotypes in \( Prdm9 \)-mediated sterility

Until now, information on the possible intrasubspecific variation of the Hstx2 locus has been lacking. Since the interval 65-69 Mb on Chr X containing the Hstx2 locus appears as a recombination cold spot (BRICK et al. 2018; LUSTYK et al. 2019), we focused on this region in the Kazakh (KH) MUS population and identified four different MUS X-chromosomal haplotypes Kha, KHB, KHC, and KHD using only three microsatellite markers (see Figure S11) indicating that recombination events within the Hstx2 locus could have occurred. Next, we tested whether these wild X-chromosomal haplotypes differed in the modulation of
Prdm9-driven hybrid sterility in crosses where Prdm9 and Hstx2 segregated. As previously observed with the Hstx2\textsuperscript{PWD} sterility allele, fertility co-segregated predominantly with the Prdm9 genotype, irrespective of which X-chromosomal haplotype the offspring possessed (Figure S12A-C). Regardless of these X-chromosomal haplotypes, all F\textsubscript{1}-hybrids inheriting any wild MUS Prdm9 alleles from their mothers were fertile, whereas siblings inheriting the msc1 alleles were sterile. To conclude, the Hstx2 did not show functionally defined intraspecific polymorphism within the studied MUS population (Figure S12D) and the results did not reveal other genetic modifiers on any individuals' wild genetic background along the maternal germline. An exception to this rule was the Prdm9\textsuperscript{msc1/dom2} male offspring, whose mothers were t-haplotype carriers, which produced an average of 2.7 M spermatozoa. These males differed significantly in both sperm count and testes weight from mice with identical genotypes whose mothers did not carry t-haplotypes (Mann-Whitney test; SC P < 0.0001, TW P = 0.0094) (Figure S12E), suggesting the presence of Prdm9 fertility modifier(s) in t-carrying populations. The apparent trans-effect of Prdm9, located on the t-haplotype, cannot be disentangled from that of other fertility modifiers within a single haplotype block. However, while testis Prdm9 expression levels in mice with t-haplotypes were not significantly different from those in mice without t-haplotypes, several other genes on t-haplotypes are enriched for copy gain events (Kelemen and Vicoso 2018) and show overexpression in the testis of \textsuperscript{t\textsuperscript{wt}} heterozygous mice (Kelemen et al. 2022). However, since the msc1/dom2 allelic combination leads to sterility even on variable genomic backgrounds, we can conclude that hybrid sterility is indeed under oligogenic control, with Prdm9 as the leading player.

**Phylogenetic analyses of Prdm9 alleles in mice**

It has been proposed that the role of Prdm9 in hybrid sterility could be related to the evolutionary divergence of homologous genomic sequences in DOM and MUS subspecies (Davies et al. 2016) and, more specifically, to the phenomenon of historical erosion of
genomic binding sites of PRDM9 ZNF domains (BAKER et al. 2015; SMAGULOVA et al. 2016) caused by repeated biased gene conversion. Consequently, only the Prdm9 alleles that have been present for longer evolutionary timescales should generate such partial erosion of their ZNFs binding motifs. To enquire into the evolutionary history of Prdm9 alleles, we analyzed the phylogenetic relationship of alleles present in our wild mice populations and other alleles for which Prdm9-mediated hybrid sterility had been studied (PARVANOV et al. 2010; MUKAJ et al. 2020). As an outgroup, the humanized Prdm9 “B-allele” was added (DAVIES et al. 2016). Since handling sequence repeats is challenging for multiple-sequence alignment algorithms, particularly when the number of repeat units differs, the allelic divergence of minisatellite sequences could not be assessed by standard assembly programs. In addition, genetic distance models (i.e., Tamura-Nei (TAMURA AND NEI 1993)) do not accurately reflect minisatellite evolution driven by de-novo recombination between repeats (JEFFREYS et al. 2013). Therefore, to reflect Prdm9 evolution more accurately, we applied an algorithm that computes Hamming Distances between minisatellite repeat, which takes not only point mutations and small indels but also within-repeat-unit processes ($w_{nu}=1$), as well as repeat-unit insertions and deletions ($w_{indel}=3.5$) and even repeat-unit duplication and slippage ($w_{slippage}=1.75$) into account (VARA et al. 2019; Damm et al. 2022). For a more conservative phylogenetic analysis, we only included nucleotide repeats that, when translated into amino acids, had a Hidden Markov Model (HMMER) bit score above 17.7, determined using (PERSIKOV AND SINGH 2014). This removed the nucleotide repeat coding for the first zinc finger in the ZNF array, which we found to be conserved in all Prdm9 alleles except mmt1.

In the neighbor-joining tree of Hamming distances rooted on the “humanized” Prdm9 allele, alleles mostly cluster according to mouse subspecies (as shown in Figure 3). However, not all alleles follow the MUS/DOM subspecies divide. The mmt1 Prdm9 allele found in all mice with t-haplotypes formed a separate branch irrespective of subspecies and mouse origin, a pattern typical of introgression. The large degree of conservation of Prdm9
The dom12 allele, neighboring a branch of MUS alleles, displays low divergence to the last common ancestor of MUS and DOM alleles. Except for mmt1 and dom12, all alleles are separated by subspecies origin (as seen in Figure 3 and Figure S5). The first node separates the dom8 allele from the Iranian population from all other DOM alleles clustering by subspecies, and a single node leads exclusively to all tested MUS alleles (Figure 3), which is broadly consistent with the evolutionary history of mice, with the DOM subspecies splitting first with estimated divergence time of 0.130-0.500 MYA, followed by the CAS and MUS subspecies around 0.110 - 0.320 MYA years ago (PHIFER-RIXEY et al. 2020). However, according to our phylogenetic reconstruction, all previously identified hybrid sterility alleles are subspecies-specific alleles of considerable divergence from a common ancestor.

However, as loci under positive selection can influence divergence times, we calculated a second distance matrix of Hamming distances after removing the hypervariable amino acids at -1, +3, and +6 of the alpha-helix, which are responsible for DNA binding (Figure S4A) before computing a second phylogeny. Indeed, larger divergence times are not driven by positive selection alone. The topology of the tree changes dramatically (compare Figure 3 with Figure S5) for alleles that differ outside of the positively selected sites. Some repeats
possess a Tryptophan (W) residue in position -5. Tryptophan’s nonpolar, aromatic, and neutral chemical properties differ from the Arginine (R) typically found in this position, which is polar and strongly basic. Secondly, there are two types of last repeats; the rarer one possesses Arginine (R) in position 13, while the more common type contains aliphatic and nonpolar Glycine (G) (Figure S2). The nucleotides coding for Glycine in this position appear to be the ancestral alleles, as the same amino acid is also encoded at this position in human PRDM9, included in the genetically engineered “humanized” allele B in mice, where it can rescue sterility (DAVIES et al. 2016)(Figure S3). While the mmt1 allele remained separated, some subspecies-specific nodes disappeared. The closely related alleles to the common ancestor of Mus musculus Prdm9 alleles are msc11, msc5, and msc10, which are now neighboring DOM alleles, possibly placing their origin before a clear separation into MUS and DOM subspecies. Curiously, while the full-length sequence of the msc7 allele had previously appeared most closely related to the msc10 allele (Figure 3), it is now found neighboring the msc2 allele from SKE/JPia, within a tree of subspecies-specific alleles (Figure S5). While only a few alleles with substitutions outside the hypervariable sites remained separated, pointing to longer divergence times between alleles, most alleles that differed only at hypervariable sites are found on the same branch once loci under positive selection are removed. They include, on the one hand, msc1, msc4, and msc9 and, on the other hand, dom6 and dom9. The divergence between Prdm9 alleles thus appears predominantly driven by positive selection on the hypervariable sites.

In conclusion, the complementary phylogenetic analyses support the accelerated evolution of the hypervariable DNA binding sites of PRDM9 protein and reiterate an evolutionary history in which Mus musculus originated in Asia and the Middle East before dispersing across Europe. The phylogenetic analyses further support a scenario in which MUS and DOM have split recently and are still speciating but do not reveal any apparent clustering of alleles co-inducing hybrid sterility by subspecies. Admittedly, no evidence was found to
support the idea that the hybrid sterility susceptible alleles (\textit{msc1, msc1, msc5, dom2, dom3, and dom5}) belong to the evolutionary oldest ones closest to the common ancestor. On the contrary, the \textit{msc1, msc2, and dom3} alleles are the most distal, sitting on the most distant branch of the phylogenetic tree (Figure 3 and Figure S5).

Comparative analyses of DNA binding motifs

The phylogenetic approaches are based on coding-nucleotide sequences; however, PRDM9 was identified as a candidate meiotic regulator based on a sequence motif enriched in human recombination hotspots but not in primate recombination hotspots (MYERS et al. 2010). To enquire into the DNA-binding motif, we predicted DNA binding motifs for each conceptually translated PRDM9-ZNF using the polynomial SVM prediction method (PERSIKOV AND SINGH 2014) which is shown in (Figure 4A). To enquire into whether similar coding sequences of the PRDM9 ZNF-array would predict binding to the same or highly similar motifs, we used TomTom (GUPTA et al. 2007). Indeed, many DNA binding motifs are highly similar to each other (Figure 4B). These include predicted DNA binding motifs of ZNF domains encoded by \textit{msc1, msc2, msc3, msc6, msc9, and msc12} in MUS, or \textit{dom6} and \textit{dom9, dom3, dom4} and \textit{dom11} in DOM. Highly similar DNA binding motifs of differing ZNF domains may be able to activate the same hotspots. Cross-activation of the same hotspot by PRDM9 variants encoded by several highly similar alleles has been observed in human hotspots (BERG et al. 2010; BERG et al. 2011). Likewise, highly similar predicted DNA-binding motifs were also enriched in contemporary mouse meiotic recombination hotspots, even in other mouse strains (SMAGULOVA et al. 2016). To investigate putative genome-wide targets of each predicted DNA binding motif, we used PWMScan (AMBROSINI et al. 2018). We then quantified how many genomic targets of predicted binding sites would be shared between PRDM9 ZNF domains of different alleles. Considerable overlap in the distribution of genome-wide putative binding sites is seen particularly across highly similar alleles (Figure 4C). Exceptionally high overlap of
putative genomic binding sites is observed for ZNF domains encoded by MUS HS alleles

\textit{msc1} and \textit{msc2}, \textit{msc3}, as well as between \textit{msc4} and \textit{msc12}, and to a lesser extent

between \textit{msc6} and \textit{msc1}.

\textbf{Conclusion}

In summary, none of the seven novel allelic combinations of wild MUS \textit{Prdm9} alleles with

DOM \textit{dom2} allele produced completely sterile F$_1$-hybrid male offspring, consistent with the

low incidence of sterile hybrids reported in the wild (TURNER \textit{et al.} 2012). Instead, we saw

either completely fertile intersubspecific hybrids (\textit{dom2} in combination with \textit{msc6}, \textit{msc7},

\textit{msc8}, \textit{msc9}, or \textit{msc10}) or a significant reduction of \textit{Prdm9}-dependent fertility and

increased levels of meiotic asynapsis (\textit{dom2} in combination with \textit{msc11}, or \textit{msc12}). Thus,

combined with the previous data from wild-derived inbred lines (MUKAJ \textit{et al.} 2020) it appears

that sterility alleles of \textit{Prdm9} may be rare. While the data on \textit{Prdm9} polymorphism in the

wild house mouse populations are accumulating, and indeed, the \textit{Prdm9} genes show

remarkable natural allelic divergence, with more than 150 alleles having been found in

mouse populations to date (BUARD \textit{et al.} 2014; KONO \textit{et al.} 2014; VARA \textit{et al.} 2019; MUKAJ

\textit{et al.} 2020), little is known about their DNA binding motifs and their degree of erosion. In

this context, the finding that five populations with fertile \textit{Prdm9} alleles carried evidence of

eroded \textit{msc1} hotspots was surprising, suggesting a decoupling of the evolutionary dynamics

of the PRDM9 zinc-finger domains and their binding sites. In other words, the erosion of the

\textit{msc1} hotspots may be much more common in natural populations than the \textit{msc1} allele itself.

This would align with the observation that recombination maps based on linkage
disequilibrium (LD) analyses revealed a significant overlap of historical hotspots of the

AKH population with contemporary \textit{msc1-activated} hotspots of the PWD strain

(WOOLDRIDGE AND DUMONT 2022). At the same time, there is weak conservation of

recombination maps at the broad and the fine scale, and most hotspots are unique to the

AHI, MCF, CGB, and AKH populations (WOOLDRIDGE AND DUMONT 2022). The other

unanswered question relates to the evolutionary age of the \textit{Prdm9} sterility alleles. In a simple
scenario, the sterility-inducing alleles would be expected to be the oldest, situated closest to
the common ancestor on the phylogenetic tree. Still, the analysis revealed the opposite:
msc1, msc2, and dom3 are the most distal, and therefore most likely the youngest, alleles.
Another question concerns the enigmatic t-haplotypes present in all three major mouse
subspecies and carrying, in all examined cases, the same Prdm9 allele coding for the same
zinc-finger domain. Is it so old because it arose before the ancestor split into the three
subspecies? If so, why does it not behave as a sterile allele? Where did it come from if it
is a recent introgression due to extremely high transmission distortion? The structure of
ZNF and other sequences of t haplotypes shows no similarity to any extant subspecies.
Clearly, more experimental evidence is needed to understand the evolutionary dynamics of
PRDM9-driven hybrid sterility.

**Availability of data and materials**

Nucleotide sequences of Prdm9 alleles were deposited to Genbank under accession
numbers OQ055171 - OQ055188. The fertility datasets generated and analyzed during the
current study, DNA binding predictions of C2H2 zinc-finger domains encoded by each
allele, as well as their genome-wide DNA binding predictions, are available in a Dryad
repository under DOI 10.5061/dryad.bzh89cm. The R package to calculate the genetic
distance between complex repeats is available at https://gitlab.gwdg.de/mpi-evolbio-it/repeatr.

**Acknowledgments**

We are grateful to Peter Donnelly for the transgene Prdm9<sup>fr1.t</sup>(PRDM9)<sub>Wtg</sub> (“humanized”) strain and Attila Toth for the HORMAD2 antibody. We thank Christine Pfeifle and the
total mouse house team at the MPI in Plön for their help with mouse breeding and
maintenance. We are grateful to Heike Harre for help with fertility phenotyping, Vladana
Fotopulosova for help with cytology, and Nicole Thomsen for help with DNA extraction
and genotyping.
Funding

Firstly, we would like to thank the Max Planck Society, the DFG (grant No. OD112/1-1 to LOH), the DAAD (57334341 to JF and LOH), and the Czech Science Foundation (grant No. 22-299-28S to JF) for funding.

Conflict of Interest

None declared

Author contributions

LOH, KFNA, ED, KKU, and AM acquired, analyzed, and interpreted data. JF, EP, and LOH conceived the project, designed the work, supervised data acquisition, and analyzed and interpreted the data. LOH wrote the manuscript. All authors have read and approved the final manuscript.

References


Altemose, N., N. Noor, E. Bitoun, A. Turnian, M. Imbeault et al., 2017a Human PRDM9 can bind and activate promoters, and other zinc-finger proteins associate with reduced recombination in cis.


Bateson, W., 1909 Heredity and Variation in Modern Lights.


Lawson, C., M. Gieske, B. Murdoch, P. Ye, Y. Li et al., 2011 Gene expression in the fetal mouse ovary is altered by exposure to low doses of bisphenol A. Biol Reprod 84: 79-86.


Myers, S., R. Bowden, A. Tumian, R. E. Bontrop, C. Freeman et al., 2010 Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. Science 327: 876-879.


Ullrich, K. K., M. Linnenbrink and D. Tautz, 2017 Introgression patterns between house mouse subspecies and species reveal genomic windows of frequent exchange.

Valiskova, B., S. Gregorova, D. Lustyk, P. Simecek, P. Jansa et al., 2022 Genic and chromosomal components of Prdm9-driven hybrid male sterility in mice (Mus musculus). Genetics 222.


Epigenetics Chromatin 8: 31.


Figures and Tables

A

B

D

C

E

F

997
Sperm Count (SC) [10^6/ml]

Testes Weight (TW) [g]

B6.DX1s x (PWDxMUS) (PWDxMUS)xB6 PWDx(B6xDOM)

<0.0001

<0.0001 0.0008 <0.0001

<0.0001

<0.0001 0.0006 0.0368 <0.0001

B6 x PWD

B6.DX1s x PWDxMUS

(PWDxMUS)xB6

PWDx(B6xDOM)
**Figure Legends**

**Figure 1.** Fertility parameters of intersubspecific hybrids were grouped by *Prdm9* genotype (A) Sperm count, and (B) paired testes weights for intersubspecific offspring of (B6.DX1s × wild MUS) or (PWD × wild DOM) crosses grouped by *Prdm9* genotype and compared to offspring of known hybrid sterility crosses (B6.DX1s × PWD) and (PWD × B6), using pairwise ANOVA with Bonferroni correction. All hybrid males carry the *Hstx1PWD* allele on Chr X. (C) The panels show spermatocyte spreads of two intersubspecific B6.DX1s × wild MUS hybrids, with differing *Prdm9* genotypes. The defects in chromosome asynapsis were assessed by antibody staining for HORMAD2 protein (green), which marks asynapsed autosomal chromosomes in addition to the nonhomologous parts of XY sex chromosomes that are physiologically observed in normally progressing meiocytes. DNA is counterstained with DAPI (blue). Synaptonemal complex assembly was evaluated by SYCP3 protein immunostaining (red) and the presence of yH2AX (grey). At the zygotene/pachytene transition, clouds of yH2AX mark chromatin associated with asynapsed axes. In addition, localized grey dots represent CEN-labeled centromeres. (D) The percentage of asynaptic cells on the Y-axis were grouped by *Prdm9* genotype. The percentage of asynaptic cells correlated with fertility parameters of intersubspecific F1-hybrids, namely (E) sperm count and (F) paired testes weights, with red dotted lines representing 95% confidence intervals, and *P* values, and Pearson values *r*, given on the bottom right.

**Figure 2** Fertility phenotypes segregate with parental *Prdm9* alleles in reciprocal intersubspecific hybrids. (A) Fertility parameters of control cross with the *Prdm9* allelic combination dom2/msc1 and Hstx2 allele from PWD or B6. (B) Intersubspecific F1 male offspring of (left) B6.DX1s females crossed to intrasubspecific MUS males (middle), Intraspecific MUS hybrid females crossed to B6 males (right) PWD females crossed to intrasubspecific DOM males. Data were pooled from parents with the same *Prdm9* genotype. Color and statistics, as in Figure 1.

**Figure 3** Neighbor-joining tree of the *Prdm9* exon12 minisatellite which encodes the DNA binding domain of the PRDM9 protein (A) cartoon depicting the amino-acids in positions -1, 3, and 6 of the alpha-helix of each C2H2 ZNFs that are responsible for DNA binding (B) Neighbor-joining tree calculated on the nucleotide sequences of all PRDM9 alleles in this study and (MUKAJ et al. 2020), that code for the C2H2 ZNFs array, with red nodes for DOM and blue nodes for MUS alleles, and with purple nodes depicting the t-haplotype allele found in MUS and DOM (C) Table depicting the C2H2 ZNF array encoded by each allele, with boxes (colored as in Figure S2 and Figure S3) representing only the amino acids responsible contacting DNA of each ZNF.

**Figure 4.** In-silico predicted PRDM9 DNA binding (A) PRDM9 DNA binding motifs are represented as sequence logos of the underlying positional weight matrices, which were predicted using the Polynomial Kernel method by (PERSIKOV et al. 2009; PERSIKOV AND SINGH 2014) on translated nucleotide sequences of alleles in this study and in (MUKAJ et al. 2020) (B) Motifs were compared using TomTom, within the MEME suite, which computes the probabilities that a random motif would be better matched than the input motif. TomTom output P-values were log-transformed and are shown in a heat matrix, such that darker colors represent better matching of sequence motifs, and lighter colors represent weaker similarities of motifs, with crossed out values representing incidences where not similarity was found. Black boxes represent incidences when the probability of another motif binding better than the motif itself is zero, as values of zero cannot be log -transformed. (C) Overlap of genomic binding sites between all ZNF domain encoding alleles, which we predicted all putative binding sites for each PRDM9 binding motif genome-wide, with brighter color showing higher genomic binding site overlap between different DNA binding motifs.
Figure S1: Distribution of original sampling sites of MUS (blue) and DOM (red), with framed circles representing mice evaluated for hybrid sterility in this study, and plain circles from Muka et al. 2020.
hypervariable positions of each C2H2 zinc finger in Figure S2 and in the phylogenetic analyses in Figure 3.

Figure S2: Types of Cysteine-2-Histidine-2 (C2H2) ZNFs found in PRDM9 MUS and DOM in this study (MUKAJ et al. 2020). We highlighted the Cysteines and Histidine residues in yellow and shaded all variable amino acids in gray, distinguishing between amino acids by color. As amino acids in positions -1, 3, and 6 of the alpha-helix are those responsible for the DNA-binding specificity of a given ZNF (as shown in the cartoon of Figure 3A), we used them as acronyms on the right. Boxes of the same color characterize the hypervariable positions of each C2H2 zinc finger in Figure S4 and in the phylogenetic analyses in Figure 3.
Figure S3: Types of C2H2 zinc fingers found in “humanized” PRDM9 variant “B.” We highlight the Cysteines and Histidine residues in yellow and shaded variable amino acids in gray. Those in positions -1, 3, and 6 of the alpha-helix are responsible for DNA binding (as shown in Figure 3A), and we used them as acronyms on the right. Boxes of the same color characterize the hypervariable positions of each C2H2 zinc finger in Figure S4 and in the phylogenetic analyses in Figure 3.
**Figure S4:** Types of C2H2 zinc finger arrays studies for hybrid sterility phenotypes

Representation of all C2H2 zinc finger arrays using only acronyms of the amino-acid positions responsible for DNA binding (as in Mus musculus domesticus and transgene (on C57Bl6/J Background))

The hypervariable positions of each C2H2 zinc finger in the phylogenetic analyses in Figure 3.
Figure S5: Neighbor-joining tree based on Hamming-Distances of nucleotides in the Prdm9 minisatellite calculated after removing hypervariable nucleotides that code for the amino acids responsible for DNA binding specificity of PRDM9 and that are known to be under strong positive selection.
Figure S6: Fertility Parameters of wild mice populations in Ploen, Germany. MUS; AKH Almaty, Kazakhstan; DOM; AHV Khazir, Iran; CBG Cologne-Bonn, Germany; MCF Massif-Central, France; twt 2nd haplotype genotype. Statistics as described previously.
Figure S7: Correlation analyses of fertility parameters in all types of performed crosses. The type of crossing scheme is (A) Intrasubspecific cross and different (B) Intersubspecific, with we tested the relationship of paired testes weight to body weight, as well as the sperm count to TW/BW ratio, and testes weights (from left to right),
Figure S8: Fertility phenotypes of hybrid offspring, grouped by Prdm9 genotype and sire ID (A) MUS from Kazakhstan without t-haplotypes, all offspring are sorted by Prdm9 genotype, with the father (sire) ID on the X-axis. One sire (50054290) was homozygous for Prdm9, all others heterozygous for different Prdm9 alleles (B). Offspring with t-haplotypes, grouped by source population of the sires, and sire IDs on the X-axis. (left) MUS from Kazakhstan (right), DOM from Cologne-Bonn, Germany (CBG), Massif-Central, France (MCF) and Ahvaz, Iran (AHI). Data pairs were compared using Welch’s t-test, and multiple comparisons were performed using ANOVA with Kruskal-Wallis’ test, corrected for multiple comparisons using Dunn’s test. Only significant values are shown on the graph with *P<0.0332*, **P<0.0021**, ***P<0.0002***, ****P<0.0001***.
Figure S9: Fertility parameters grouped by paternal Prdm9 allele from Figure 1, analyzed by age.
Figure S10: Fertility parameters of intra- and intersubspecific F1-hybrids (A) sperm count (B) paired testes weight of interspecific and Intrasubspecific mice. Paternal Prdm9 alleles are distinguished by colors, while the shape of the data points distinguishes crossing schemes. (Square data points) Intrasubspecific male F1 offspring of PWD females crossed to MUS males, or Intrasubspecific B6 females crossed to DOM males. Two types of interspecific crosses were performed; MUS males were crossed to B6.DX1s females and wild DOM males were crossed to PWD females (circles as data points, also shown in Figure 1); for the MUS males, a third cross was performed where they were mated with B6.DX1s Prdm9hu females (triangular data points). Since most males used as sires were heterozygous for different allelic combinations of Prdm9, data from several crosses pooled by Prdm9 genotype of the F1-hybrid offspring, and the significance of differences in fertility phenotypes was evaluated using ordinary one-way ANOVA with Bonferroni correction as before.
Figure S11: X-chromosomal haplotypes across the refined Hstx2 locus from (LUSTYK et al. 2019) based on allelic variation at microsatellite markers SX65100, SR51, and SX69084 in the laboratory and wild mice. All genes and microRNAs annotated within this locus are shown in the mm38 reference genome.
Figure S12: Segregation analyses of maternal Prdm9 genotype together with the variation of wild X-chromosomal haplotypes. Intersubspecific F1 male offspring of Intraspecific MUS hybrid females crossed to B6 males (as in Figure 2B) that inherited wild MUS X-chromosomal haplotypes. X-axes depict both the maternal Prdm9 allele and wild MUS X-chromosomal haplotype, defined as shown in (Figure S11). Fertility parameters, sperm count (top), and paired testes weights (bottom) are shown, and the significance of differences in fertility phenotypes was evaluated as before. (A-C) Segregation analyses grouped by maternal Prdm9 allele combination (A) maternal msc1 and msc6 alleles (B) maternal msc1 and msc8 alleles (C) msc1 and mmt1KH alleles (D-E) Pooled data of hybrids from different mothers that all inherited the same msc1/dom2 allelic combination (D) pooled by X-chromosomal haplotype (E) pooled by whether mothers possessed t-haplotypes.
Table S1: Variation, Introggression, and Recombination rate of outcrossed populations with data from (a) (LAWAL et al. 2021) (b) (WOOLDRIDGE AND DUMONT 2022) (c) (BANKER et al. 2022) (d) (STAUBACH et al. 2012)

<table>
<thead>
<tr>
<th>outcrossed populations</th>
<th># SNPS</th>
<th>SNP density (bp/SNP)</th>
<th># population private variants</th>
<th>Average introgression per individual (%)</th>
<th>% of the genome affected by introgression</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKH</td>
<td>10,937,288 a</td>
<td>225 b</td>
<td>1,872,782 a</td>
<td>NA</td>
<td>4.90 c</td>
</tr>
<tr>
<td>MCF</td>
<td>11,108,085 a</td>
<td>222 b</td>
<td>2,208,483 a</td>
<td>0.069 c</td>
<td>2.80 c</td>
</tr>
<tr>
<td>CBG</td>
<td>11,930,888 a</td>
<td>206 b</td>
<td>545,881 a</td>
<td>0.068 c</td>
<td>2.80 c</td>
</tr>
<tr>
<td>AHI</td>
<td>17,877,283 a</td>
<td>138 b</td>
<td>3,333,440 a</td>
<td>0.036 c</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table S2 Primers used to genotype the Hstx2 interval from (a)(Lustyk et al. 2019) and (b) P. Jansa personal communication

<table>
<thead>
<tr>
<th>STR marker</th>
<th>Forward Primer: (5’ to 3’)</th>
<th>Label</th>
<th>Reverse Primer: (5’ to 3’)</th>
<th>position in mm10</th>
</tr>
</thead>
<tbody>
<tr>
<td>SX65100 a</td>
<td>AAAAAAGGCTGCTGGAAGTCA 6-FAM</td>
<td></td>
<td>ATGAGGCTGGGATTCTTCCT</td>
<td>ChrX:65100392-65100563</td>
</tr>
<tr>
<td>SR51 b</td>
<td>AAGCAGGAGTGGGTGAGTTG 6-FAM</td>
<td></td>
<td>CTCCCTTTTGCCAAGGTCCA</td>
<td>ChrX:67841601-67841743</td>
</tr>
<tr>
<td>SX69084 a</td>
<td>AGGCCTTCTGGGCTTATCTC HEX</td>
<td></td>
<td>AAAGCTCATGGATGAGAAAACA</td>
<td>ChrX:69084171-69084417</td>
</tr>
</tbody>
</table>
Table S3: Difference in Sperm count (SC) grouped by Prdm9 Genotype of F1 hybrid male offspring, evaluated using Ordinary one-way ANOVA with Bonferroni’s multiple comparisons test.

<table>
<thead>
<tr>
<th>Ordinary one-way ANOVA</th>
<th>Summary</th>
<th>Bonferroni Adjusted -P- Value</th>
<th>Mean 1</th>
<th>Mean 2</th>
<th>Mean Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>msc6 vs. msc1</td>
<td>***</td>
<td>0.0001</td>
<td>22.1</td>
<td>0</td>
<td>22.1</td>
</tr>
<tr>
<td>msc7 vs. msc1</td>
<td>***</td>
<td>0.0007</td>
<td>20.8</td>
<td>0</td>
<td>20.8</td>
</tr>
<tr>
<td>msc8 vs. msc1</td>
<td>**</td>
<td>0.0043</td>
<td>18.2</td>
<td>0</td>
<td>18.2</td>
</tr>
<tr>
<td>msc9 vs. msc1</td>
<td>*</td>
<td>0.0246</td>
<td>20.3</td>
<td>0</td>
<td>20.3</td>
</tr>
<tr>
<td>msc10 vs. msc1</td>
<td>**</td>
<td>0.0088</td>
<td>18</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>msc11 vs. msc1</td>
<td>ns</td>
<td>&gt;0.9999</td>
<td>11.4</td>
<td>0</td>
<td>11.4</td>
</tr>
<tr>
<td>msc12 vs. msc1</td>
<td>ns</td>
<td>&gt;0.9999</td>
<td>9.97</td>
<td>0</td>
<td>9.97</td>
</tr>
<tr>
<td>mmt1** vs. msc1</td>
<td>**</td>
<td>0.0029</td>
<td>18.2</td>
<td>0</td>
<td>18.2</td>
</tr>
<tr>
<td>msc6 vs. msc12</td>
<td>***</td>
<td>0.0005</td>
<td>22.1</td>
<td>9.97</td>
<td>12.1</td>
</tr>
<tr>
<td>msc7 vs. msc12</td>
<td>**</td>
<td>0.0099</td>
<td>20.8</td>
<td>9.97</td>
<td>10.8</td>
</tr>
<tr>
<td>msc8 vs. msc12</td>
<td>ns</td>
<td>0.1046</td>
<td>18.2</td>
<td>9.97</td>
<td>8.26</td>
</tr>
<tr>
<td>msc9 vs. msc12</td>
<td>ns</td>
<td>&gt;0.9999</td>
<td>20.3</td>
<td>9.97</td>
<td>10.3</td>
</tr>
<tr>
<td>msc10 vs. msc12</td>
<td>ns</td>
<td>&gt;0.9999</td>
<td>11.4</td>
<td>9.97</td>
<td>1.39</td>
</tr>
<tr>
<td>mmt1** vs. msc12</td>
<td>*</td>
<td>0.0464</td>
<td>18.2</td>
<td>9.97</td>
<td>8.28</td>
</tr>
<tr>
<td>mmt1** vs. dom2</td>
<td>****</td>
<td>&lt;0.0001</td>
<td>20.3</td>
<td>0</td>
<td>20.3</td>
</tr>
<tr>
<td>mmt1** vs. dom2</td>
<td>ns</td>
<td>0.1394</td>
<td>14.3</td>
<td>0</td>
<td>14.3</td>
</tr>
<tr>
<td>mmt1** vs. dom2</td>
<td>****</td>
<td>&lt;0.0001</td>
<td>24.6</td>
<td>0</td>
<td>24.6</td>
</tr>
</tbody>
</table>
Table S 4 Difference in testis weight (TW) grouped by Prdm9 Genotype of F1 hybrid male offspring, evaluated using Ordinary one-way ANOVA with Bonferroni's multiple comparisons test.

<table>
<thead>
<tr>
<th>Ordinary one-way ANOVA</th>
<th>Summary</th>
<th>Bonferroni Adjusted -P- Value</th>
<th>Mean 1</th>
<th>Mean 2</th>
<th>Mean Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>msc6 vs. msc1</td>
<td>****</td>
<td>&lt;0.0001</td>
<td>0.206</td>
<td>0.0623</td>
<td>0.144</td>
</tr>
<tr>
<td>msc7 vs. msc1</td>
<td>****</td>
<td>&lt;0.0001</td>
<td>0.208</td>
<td>0.0623</td>
<td>0.145</td>
</tr>
<tr>
<td>msc8 vs. msc1</td>
<td>****</td>
<td>&lt;0.0001</td>
<td>0.213</td>
<td>0.0623</td>
<td>0.151</td>
</tr>
<tr>
<td>msc9 vs. msc1</td>
<td>****</td>
<td>&lt;0.0001</td>
<td>0.206</td>
<td>0.0623</td>
<td>0.143</td>
</tr>
<tr>
<td>msc10 vs. msc1</td>
<td>****</td>
<td>&lt;0.0001</td>
<td>0.202</td>
<td>0.0623</td>
<td>0.139</td>
</tr>
<tr>
<td>msc11 vs. msc1</td>
<td>****</td>
<td>&lt;0.0001</td>
<td>0.18</td>
<td>0.0623</td>
<td>0.117</td>
</tr>
<tr>
<td>msc12 vs. msc1</td>
<td>****</td>
<td>&lt;0.0001</td>
<td>0.169</td>
<td>0.0623</td>
<td>0.107</td>
</tr>
<tr>
<td>mmt1*KH vs. msc1</td>
<td>****</td>
<td>&lt;0.0001</td>
<td>0.203</td>
<td>0.0623</td>
<td>0.141</td>
</tr>
<tr>
<td>msc6 vs. msc12</td>
<td>*</td>
<td>0.0115</td>
<td>0.206</td>
<td>0.169</td>
<td>0.0371</td>
</tr>
<tr>
<td>msc7 vs. msc12</td>
<td>*</td>
<td>0.0107</td>
<td>0.208</td>
<td>0.169</td>
<td>0.0386</td>
</tr>
<tr>
<td>msc8 vs. msc12</td>
<td>***</td>
<td>0.0002</td>
<td>0.213</td>
<td>0.169</td>
<td>0.0442</td>
</tr>
<tr>
<td>msc9 vs. msc12</td>
<td>ns</td>
<td>&gt;0.9999</td>
<td>0.206</td>
<td>0.169</td>
<td>0.0365</td>
</tr>
<tr>
<td>msc10 vs. msc12</td>
<td>ns</td>
<td>0.1165</td>
<td>0.202</td>
<td>0.169</td>
<td>0.0324</td>
</tr>
<tr>
<td>msc11 vs. msc12</td>
<td>ns</td>
<td>&gt;0.9999</td>
<td>0.18</td>
<td>0.169</td>
<td>0.0103</td>
</tr>
<tr>
<td>mmt1^*KH vs. dom2</td>
<td>****</td>
<td>&lt;0.0001</td>
<td>0.189</td>
<td>0.0705</td>
<td>0.118</td>
</tr>
<tr>
<td>mmt1*K vs. dom2</td>
<td>****</td>
<td>&lt;0.0001</td>
<td>0.188</td>
<td>0.0705</td>
<td>0.118</td>
</tr>
<tr>
<td>mmt1^*KH vs. dom2</td>
<td>****</td>
<td>&lt;0.0001</td>
<td>0.171</td>
<td>0.0705</td>
<td>0.101</td>
</tr>
</tbody>
</table>