



HAL
open science

High prevalence of PRDM9-independent recombination hotspots in placental mammals

Julien Joseph, Djivan Prentout, Alexandre Laverré, Théo Tricou, Laurent Duret

► **To cite this version:**

Julien Joseph, Djivan Prentout, Alexandre Laverré, Théo Tricou, Laurent Duret. High prevalence of PRDM9-independent recombination hotspots in placental mammals. Proceedings of the National Academy of Sciences of the United States of America, 2024, 121 (23), pp.e2401973121. 10.1073/pnas.2401973121 . hal-04604269

HAL Id: hal-04604269

<https://hal.science/hal-04604269>

Submitted on 7 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HIGH PREVALENCE OF PRDM9-INDEPENDENT RECOMBINATION HOTSPOTS IN PLACENTAL MAMMALS

J. Joseph¹, D. Prentout², A. Laverre^{3,4}, T. Tricou¹, L. Duret¹

¹Laboratoire de Biométrie et Biologie Evolutive, Université Lyon 1, CNRS, UMR 5558, Villeurbanne, France

²Department of Biological Sciences, Columbia University, New York, NY 10027, USA

³Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

⁴Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland

julien.joseph@ens-lyon.fr

June 6, 2024

Abstract

1 In many mammals, recombination events are concentrated into hotspots directed by a se-
2 quence specific DNA-binding protein named PRDM9. Intriguingly, PRDM9 has been lost
3 several times in vertebrates, and notably among mammals, it has been pseudogenized in
4 the ancestor of canids. In absence of PRDM9, recombination hotspots tend to occur in
5 promoter-like features such as CpG islands. It has thus been proposed that one role of
6 PRDM9 could be to direct recombination away from those PRDM9-independent hotspots.
7 However, the ability of PRDM9 to direct recombination hotspots has been assessed only in a
8 handful of species, and a clear picture of how much recombination occurs outside of PRDM9-
9 directed hotspots in mammals is still lacking. In this study, we derived an estimator of past
10 recombination activity based on signatures of GC-biased gene conversion in substitution pat-
11 terns. We applied it to quantify recombination activity in PRDM9 independent hotspots in
12 52 species of boreoeutherian mammals. We observed a wide range of recombination rate at
13 these loci: several species (such as mice, humans, some felids or cetaceans) show a deficit
14 of recombination, while a majority of mammals display a clear peak of recombination. Our
15 results demonstrate that PRDM9-directed and PRDM9-independent hotspots can co-exist
16 in mammals, and that their co-existence appears to be the rule rather than the exception.
17 Additionally, we show that the location of PRDM9-independent hotspots is relatively stable
18 compared to that of PRDM9-directed hotspots, but that it nevertheless evolves slowly in
19 concert with DNA hypomethylation.

20 **Keywords** PRDM9 · Recombination landscape · Hotspots · Mammals · gBGC

21 **Significance statement**

22 Meiotic recombination is a key process both to produce gametes and to generate genetic diversity. In many
23 Eukaryotes, recombination tends to occur at specific loci (recombination hotspots). Two types of hotspots
24 have been described so far: in humans and mice, hotspots are directed by the protein PRDM9, whereas in
25 species that lack this gene, hotspots tend to occur in promoter-like features (“default hotspots”). In this
26 study, we demonstrate that despite possessing a functional PRDM9, most mammals also use default hotspots,
27 sometimes as much as in PRDM9-lacking species. Interestingly, the widespread coexistence of redundant
28 recombination-directing systems might explain why Prdm9 has been recurrently lost across animals.

29 **Introduction**

30 Meiotic recombination is a crucial step in the production of gametes for a large majority of eukaryotes. It is
31 initiated by programmed double-strand breaks (DSBs), that can be resolved in two types of recombination
32 events, using the homolog as a template: crossovers (COs), where there is a reciprocal exchange of chromo-
33 some arms, and non crossovers (NCOs), where DSB repair leads to gene conversion around the DSB site
34 without reciprocal exchange. In most eukaryotes, at least one CO per chromosome is necessary to ensure
35 correct segregation between homologs and is therefore mandatory for meiosis success (Page and Hawley,
36 2003; Gerton and Hawley, 2005). In many vertebrates, recombination events are not uniformly distributed
37 along the genome (reviewed in Stapley et al. (2017); Zelkowski et al. (2019)). Instead, they tend to be
38 concentrated in so-called recombination hotspots (Lichten and Goldman, 1995; Tock and Henderson, 2018).

39 In many mammals, the position of recombination hotspots is determined by the zinc-finger protein
40 PRDM9, which binds specific DNA motifs and recruits the DSB machinery through histone methylation
41 (Baudat et al., 2010; Myers et al., 2010; Parvanov et al., 2010; Diagouraga et al., 2018). This gene is highly
42 polymorphic and hundreds of alleles have been reported in mice and humans (Buard et al., 2014; Kono et al.,
43 2014; Alleva et al., 2021). Most allelic diversity is concentrated on residues of the zinc fingers that interact
44 with the DNA, leading to changes in DNA sequence specificity. Therefore, the position of recombination
45 hotspots varies within a population and between species (Auton et al., 2012; Smagulova et al., 2016; Alleva
46 et al., 2021). Additionally, PRDM9 tends to erode its targets through gene conversion (Baker et al., 2015;
47 Smagulova et al., 2016). As the available targets for a given PRDM9 allele become scarce, its ability to
48 generate enough COs for meiosis to succeed is compromised. A new allele with more targets will then
49 be positively selected leading to a red-queen dynamic accelerating the turnover of recombination hotspots
50 (Úbeda and Wilkins, 2011; Latrille et al., 2017; Baker et al., 2022; Genestier et al., 2023). It has also been
51 proposed that when PRDM9 binds symmetrically both homologs, it facilitates the repair of DSBs as COs
52 and thereby contributes to the success of meiosis (Davies et al., 2016; Li et al., 2019; Hinch et al., 2019).
53 Indeed, experiments in mice and rats showed that an inactivation of this protein drastically reduces fertility
54 (Mihola et al., 2019, 2021; Brick et al., 2012).

55 Despite its central role in recombination, PRDM9 has been repeatedly lost in vertebrates. (Baker et al.,
56 2017; Cavassim et al., 2022). Among amniotes, one loss occurred in the ancestor of archosaurs (crocodiles and

57 birds) (Singhal et al., 2015; Cavassim et al., 2022), and another one in the ancestor of canids (dogs, wolves
58 and foxes) (Oliver et al., 2009; Axelsson et al., 2012; Auton et al., 2013). In dogs and several passerines,
59 studies based on linkage disequilibrium (LD) showed that recombination hotspots tend to occur in CpG
60 islands, and more specifically in hypomethylated ones in dogs (Auton et al., 2013; Berglund et al., 2015;
61 Singhal et al., 2015; Kawakami et al., 2017). Likewise, the DSB hotspots of a mouse whose PRDM9 has
62 been inactivated through a knock-out (PRDM9^{-/-}) also occur in promoter-like features (Brick et al., 2012).
63 Interestingly, an increase of recombination rate near promoters has also been observed in dogs, birds, plants
64 and yeasts, which lack PRDM9 (Petes, 2001; Marand et al., 2017).

65 Those observations led to the conclusion that there exist two types of recombination landscapes in
66 vertebrates. The first one is PRDM9-dependent, fast evolving with recombination targeted away from
67 promoter-like features, while the second is relatively stable, with recombination occurring in promoter-like
68 feature such as CpG islands. However, a recent finding in snakes challenged this binary view. In rattle snakes
69 and corn snakes, which still possess a functional PRDM9, some hotspots are directed by PRDM9, but others
70 are directed toward CpG islands (Schield et al., 2020; Hoge et al., 2023). It was first proposed that PRDM9
71 could have a different role in rattle snake and direct recombination towards CpG islands (Schield et al.,
72 2020), but other observations suggest that the recombination landscapes in these snakes rather reflects the
73 inefficiency of their PRDM9-dependent pathway to direct DSBs away from CpG islands (Hoge et al., 2023).
74 Altogether, the two studies concur on the fact that the recombination landscape in snakes differs from what
75 is observed in mammals, despite the presence of PRDM9 (Schield et al., 2020; Hoge et al., 2023).

76 Most of our knowledge about PRDM9 function and evolution has been acquired in a handful of mammals
77 (mostly human and mice). In mice, Smagulova and colleagues analyzed the position of DSB hotspots
78 obtained by ChIP-seq of DMC1 in strains carrying different PRDM9 alleles. They measured the overlap of
79 these hotspots with the ones of a strain knocked out for Prdm9 (PRDM9-independent hotspots) (Smagulova
80 et al., 2016). Those strains showed differences in their capacity to target DSBs away from the hotspots of
81 the Prdm9^{-/-} mouse (hereafter referred to as 'MDH', for 'Mouse Default Hotspots'). Some strains show a
82 significant deficit of DSB hotspots at MDH loci, while others, carrying less dominant PRDM9 alleles, show
83 up to a 6-fold DSB hotspot enrichment at these loci, even though they represent a small proportion of all
84 recombination hotspots (~7%) (Smagulova et al., 2016). This shows that even when PRDM9 is present,
85 PRDM9-independent hotspots can be active in mammals. However, this activity could just be the reflection
86 of a specific PRDM9 deficiency in some mice strains and overall, we still have no clear idea on how prevalent
87 the usage of PRDM9-independent hotspots is in mammals. A comprehensive understanding would require
88 measures of fine-scale variations in recombination rate for a wide range of mammalian species, and ideally
89 across long periods of time.

90 Recombination hotspots can be mapped directly in meiotic cells (e.g. by chromatin immunoprecipitation
91 with antibodies to DMC1 in spermatocytes (Brick et al., 2012; Pratto et al., 2014; Smagulova et al., 2016;
92 Alleva et al., 2021)), but these molecular approaches are tedious and only amenable for a few model organ-
93 isms. High-resolution recombination maps can also be inferred from patterns of linkage disequilibrium (LD).
94 This approach is more scalable and provides information on sex-averaged historical recombination activity

95 at the population scale. However, this approach remains laborious and expensive (it requires the sequencing
96 of at least 10 individuals per species (Auton and McVean, 2007; Chan et al., 2012)), and it is sensitive to
97 various sources of errors (Spence and Song, 2019; Samuk and Noor, 2022; Raynaud et al., 2023). Hence, for
98 now, such LD-based recombination maps are available only for a very limited number of species.

99 Alternatively, substitution patterns have been found to be informative about past recombination rates.
100 In particular, it has been shown in mammals that recombination induces a transmission bias of GC alleles
101 through the process of GC-biased gene conversion (gBGC). This eventually leads to an elevation of the WS
102 (AT to GC) substitution rate, and a decrease of the SW (GC to AT) substitution rate (Nagylyaki, 1983;
103 Duret and Arndt, 2008; Glémin, 2010). The substitution rate matrix can be conveniently summarized by a
104 single parameter, the equilibrium GC-content (hereafter noted GC^*), which corresponds to the GC-content
105 that sequences would reach if the pattern of substitution observed in that branch remained constant over
106 time (Duret and Arndt, 2008). GC^* correlates well with the strength of DSB hotspots in mice (Clément
107 and Arndt, 2013), with the LD-based recombination rates in humans (Munch et al., 2014; Glémin et al.,
108 2015), and with the LD-based strength of recombination hotspots in dogs (Axelsson et al., 2012; Auton
109 et al., 2013). Moreover, GC^* reflects the recombination activity along the entire branch where substitution
110 patterns are analyzed, and hence can inform about past recombination events that are no longer detectable
111 with methods measuring recombination in individuals or populations (Lesecque et al., 2014; Munch et al.,
112 2014). This provides insights on the long term use of PRDM9-independent hotspots, integrated over long
113 periods of time, probably encompassing the rise and fall of several PRDM9 alleles. However, variation in
114 GC^* between species cannot be directly interpreted as variations in recombination rates alone, since GC^*
115 also depends on the mutation bias towards AT, the repair bias towards GC, the effective population size and
116 the mean length of the conversion tracts (Eyre-Walker, 1999; Glémin, 2010).

117 In this study, we present an estimator of relative recombination rates based on substitution patterns that
118 allows us to directly compare fine-scale recombination rate variations in a wide range of species using only
119 3 genomes (one focal genome, a sister species and an outgroup). We then use it to assess the recombination
120 activity at MDH loci in 52 species spanning the diversity of boreoeutherians. We reveal a high heterogeneity
121 in the use of these PRDM9-independent hotspots. We show that PRDM9 alleles in humans and mice
122 have been particularly efficient at directing DSBs away from PRDM9-independent hotspots but that these
123 two species are not representative of all mammals. Finally, we show that three species, namely the southern
124 elephant seal, the ring-tailed lemur and the daurian ground squirrel, have used PRDM9-independent hotspots
125 as much as PRDM9-deficient canids. This shows that the two kinds of hotspots-regulation mechanisms that
126 have been described so far in vertebrates are not mutually exclusive and that the fine-scale recombination
127 landscapes of many mammals are much closer to those of birds and other PRDM9-lacking amniotes than
128 previously thought. We further show that the recombination activity observed at MDH loci in PRDM9-
129 containing mammals depends on the conservation of their DNA methylation pattern, which suggests a link
130 between the evolution of DNA methylation and of PRDM9-independent recombination landscapes.

131 **Results**132 **Conservation of recombination hotspots between PRDM9-deficient mammals**

133 In finches and flycatchers, the fine-scale recombination landscape has been shown to be stable through time,
134 as a large proportion of hotspots are shared between closely related species (Singhal et al., 2015; Kawakami
135 et al., 2017). To test whether loci corresponding to PRDM9-independent hotspots are also evolutionary stable
136 in mammals, we analyzed the overlap between recombination hotspots detected in dogs, which naturally lack
137 PRDM9, and those identified in the *Prdm9*^{-/-} mutant mouse (MDH) (Smagulova et al., 2016). Among the
138 30,929 MDH, 15,009 (49%) could be assigned to one-to-one orthologous loci in the dog genome. Among the
139 7008 dog hotspots identified in the LD-based recombination map of dogs (Auton et al., 2013), 34% overlap
140 with MDH loci (Suppfig. S1)(compared to 0.06% expected by chance, given that dog hotspots and MDH
141 loci cover respectively 2% and 3% of the dog genome). Although this enrichment is very strong, it should be
142 noted that 42% of the dog hotspots that could be mapped on the mouse genome occur outside of MDH loci
143 (1809/3109) (Suppfig. S1). This number is difficult to interpret because it has been shown that LD-based
144 methods can produce a high number of false positives (Raynaud et al., 2023). Moreover, DSB hotspots are
145 obtained on males only, while LD maps are sex-averaged. It is also possible that some MDH loci have not
146 been identified as hotspots in dogs simply because they did not meet the threshold criteria to be defined as
147 such. To avoid the problem of the arbitrary threshold, we computed the LD-based recombination rate in dogs
148 (from Auton et al. (2013)) as a function of the distance to the closest MDH loci (Fig. 1A&B) (Auton et al.,
149 2013). We divided the 30,929 mouse *Prdm9*^{-/-} hotspots into three equally sized categories of strength, based
150 on DMC1 ChIP-seq read counts (Smagulova et al., 2016). Respectively 5,266 strong MDH, 4,961 medium
151 MDH and 4,781 weak MDH could be mapped on the dog genome. There is a sharp peak of recombination
152 centered on MDH loci in dogs. This peak is higher for strong MDH and weaker for medium and weak ones
153 (Fig. 1A). This confirms that many recombination hotspots are conserved between *Prdm9*^{-/-} mice and dogs,
154 but it does not rule out the existence of species-specific recombination hotspots. To explore factors that
155 might drive the evolution of PRDM9-independent hotspots, we analyzed their DNA methylation level in the
156 germline. Indeed, indirect evidences suggested that recombination hotspots in dogs are associated to germline
157 hypomethylated regions (HMRs) (Berghlund et al., 2015). Using HMRs identified by bi-sulfite sequencing in
158 dog sperm (Qu et al., 2018), we observed that 74% of dog hotspots are located inside HMRs, which represent
159 only 3.7% of the dog genome. In mice, the overlap is even stronger: using HMRs identified in mouse sperm
160 (Hammoud et al., 2014), we observed that out of the 30,929 hotspots found in the *Prdm9*^{-/-} mutant, 93%
161 are located within HMRs, which cover only 4.6% of the mouse genome (see methods for details). This
162 indicates that PRDM9-independent hotspots are associated with DNA hypomethylation both in *Prdm9*^{-/-}
163 mutant mice and in canids. Interestingly, 48% of MDH loci are methylated in dog sperm if we restrict the
164 definition of hotspots to their midpoint (7,186/15,009). This shows that many MDH loci are specifically
165 hypomethylated in mice but not in dogs. This is consistent with previous observations showing that murid
166 genomes have accumulated many new HMRs compared to other mammals (Qu et al., 2018). To test whether
167 these shifts in methylation levels are associated with changes in recombination activity, we computed the
168 LD-based recombination rate in dogs as a function of the distance to the closest MDH locus, separating

169 those whose midpoints overlap a HMR in dogs (7,186), and those that do not (7,283)(Fig. 1B). There is a
 170 high and pronounced recombination peak at MDH loci that are hypomethylated in dog sperm (Fig. 1B). In
 171 contrast there is almost no elevation of recombination at MDH loci that are methylated in dogs (Fig. 1B).
 172 This confirms that methylation is clearly associated to recombination hotspots in the absence of PRDM9,
 173 and that many PRDM9-independent recombination hotspots are species-specific.

174 **Equilibrium GC content is a good predictor of past recombination activity**

175 To assess whether equilibrium GC content could be used as a proxy of recombination rate in other species,
 176 we repeated the above analyses using GC^* instead of LD-based recombination rate (Fig. 1C&D). Strikingly,
 177 the profile of GC^* perfectly mirrors the LD-based recombination rate profile. It should however be noticed
 178 that the GC^* peaks are slightly sharper than the LD-based ones (Fig. 1C&D), as are the peaks of DMC1
 179 ChIP-seq read coverage in the *Prdm9*^{-/-} mouse (Suppfig. S2). This suggests that GC^* is able to capture
 180 signals of past recombination with a higher spatial resolution than LD.

181 **Estimation of the relative recombination rate in PRDM9-independent hotspots**

182 Following a large body of literature, we showed that GC^* can be very informative on intra-genomic recom-
 183 bination rate variations (Pessia et al., 2012; Auton et al., 2013; Clément and Arndt, 2013; Lartillot, 2013;
 184 Munch et al., 2014; Glémin et al., 2015; Singhal et al., 2015; Figuet et al., 2015; Bolívar et al., 2016; Galtier
 185 et al., 2018; Charlesworth et al., 2020). However, the height of the GC^* peak in hotspots is difficult to
 186 interpret in term of recombination rate because it is also affected by other parameters (the length of gene
 187 conversion tracts, the mutation bias towards AT, the mismatch repair bias towards GC and the effective
 188 population size), which can vary between species (Lartillot, 2013; Galtier et al., 2018; Galtier, 2021). We
 189 thus derived an estimator that can capture the relative recombination rate at PRDM9-independent hotspots
 190 that controls for those parameters and is therefore comparable between species. Using the probability of
 191 fixation of AT and GC alleles in presence of gBGC derived by Nagylaki (1983), we obtained an expression
 192 of the ratio of the recombination rate within hotspots relative to their flanking regions (see details in the
 193 Methods). This relative recombination rate only depends on GC^* inside hotspots and in flanking regions,
 194 and on the mutation bias (GC^μ).

$$\frac{r_{hot}}{r_{flank}} = \frac{\text{logit}(GC_{hot}^*) - \text{logit}(GC^\mu)}{\text{logit}(GC_{flank}^*) - \text{logit}(GC^\mu)} \quad (1)$$

195 where r_{hot} is the recombination rate in hotspots, r_{flank} the recombination rate in flanking regions and
 196 GC^μ the GC content expected under mutation only. It should be noted that r_{hot} and r_{flank} encompass all
 197 recombination events that can lead to gBGC (potentially COs and/or NCOs).

198 For the rest of the study, hotspots are defined as the 400 bp regions centered on their midpoint, and
 199 flanking regions as those spanning from 5 to 8 kb upstream and downstream of the hotspots (Fig. 2B).

200 Using the three values GC_{hot}^* , GC_{flank}^* and GC^μ , it is possible to compute a measure of the relative
 201 recombination rate within hotspots compared to their flanking regions, which can then be compared between

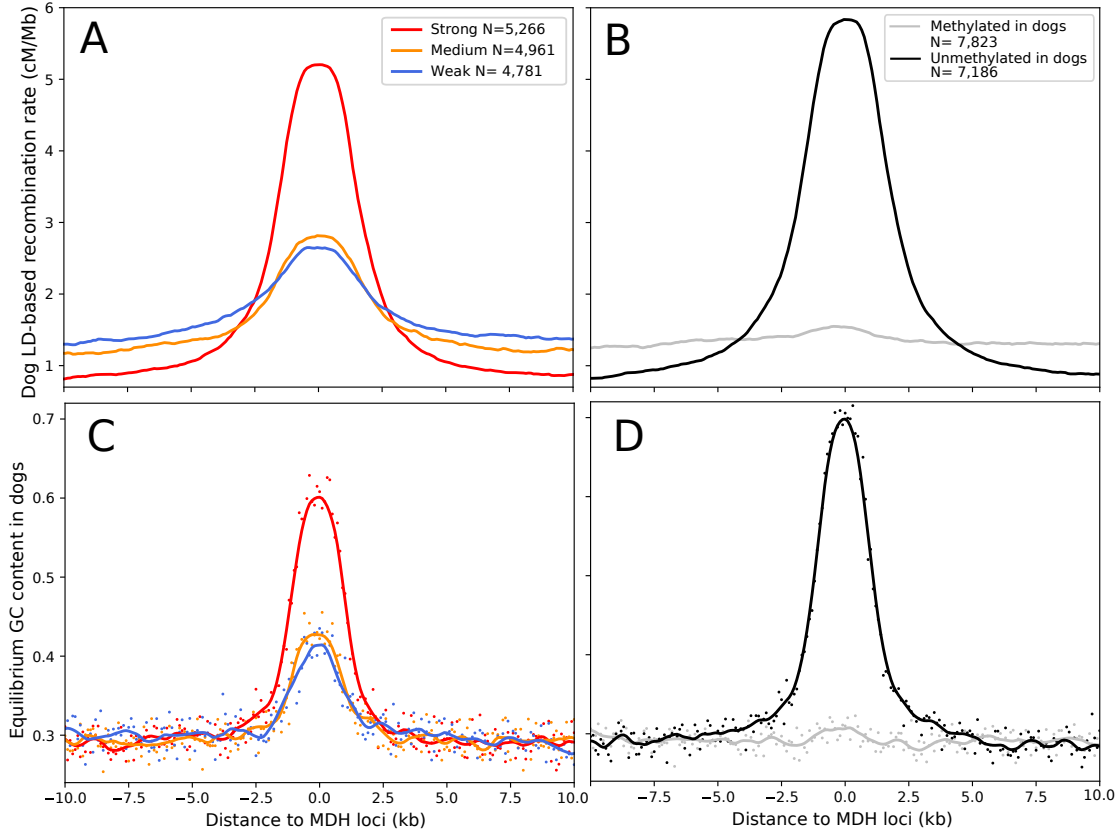


Figure 1: A) Dog LD-based recombination rate as a function of the distance to the closest MDH locus. MDH were divided in three equally sized categories of strength: strong hotspots in red (10-190 FPKM), medium hotspots in orange (5-10 FPKM) and weak hotspots in blue (0-5 FPKM). The line corresponds to the mean value of LD-based recombination rate in a 100 bp window. B) Same as A but MDH loci were divided in two categories depending on their methylation level in dog sperm. MDH loci that are found outside hypomethylated regions in dogs are in grey and those found inside hypomethylated regions are in black. C) & D) Equilibrium GC content in dogs as a function of the distance to the closest MDH locus. Using the same partitions of hotspots as A) and B). Points correspond to the mean value of GC^* in a 100 bp window. The line corresponds to a smoothing of the data with a loess function.

202 species. Equation 1 holds true under the assumption that the other four parameters affecting the strength
 203 of gBGC (effective population size, mismatch repair bias, length of the conversion tract and the mutation
 204 bias) do not differ between the hotspots and their flanking regions (see Methods). As shown in the previous
 205 section, PRDM9-independent hotspots are often hypomethylated. This implies that the mutation rate from
 206 CpG to TpG or CpA in hotspots is lower than in the flanking regions, which violates our assumption of a
 207 constant mutation bias. To avoid this problem, we excluded CpG sites from all the analyses.

208 GC^*_{hot} and GC^*_{flank} can easily be computed from the substitutions in PRDM9-independent hotspots and
 209 their flanking region but GC^μ is more difficult to estimate. Interestingly, it has been shown that genome-
 210 wide variations of GC^* are mainly the result of gBGC, and that GC^μ is quite constant along chromosomes,

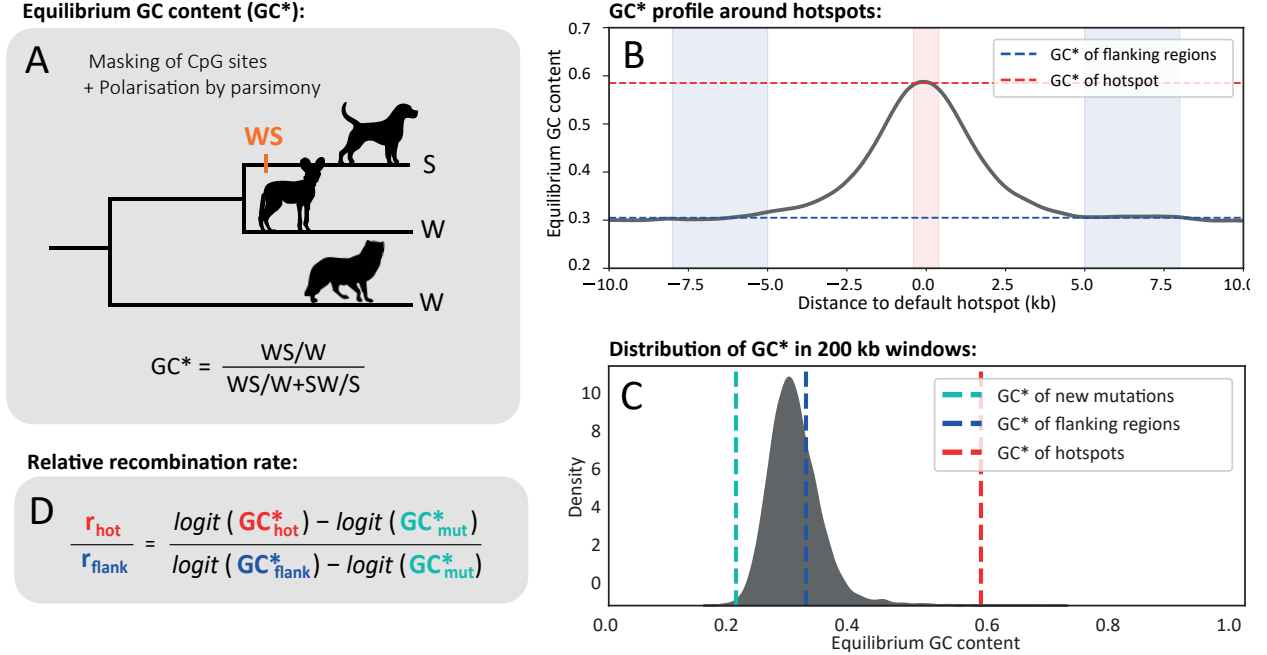


Figure 2: Overview of the method for inferring relative recombination rates in PRDM9-independent hotspots. A) We call substitutions using parsimony on trios of closely related species after having masked CpG dinucleotides. B) We compute GC^* in 400 bp windows centered on the midpoint of the PRDM9-independent hotspots, and GC^* in the flanking regions (from 5 to 8 kb upstream and downstream of the center of the PRDM9-independent hotspot). C) We compute the genome-wide distribution of GC^* in 200 kb windows and take the 1st percentile as the GC^* of new mutations. D) Using the probability of fixation of AT and GC alleles in presence of gBGC derived by Nagylaki (1983), we compute the relative recombination rate as a function of the three values of GC^* (see methods).

211 despite large-scale variation in mutation rates (Smith et al., 2018). An approach to estimate GC^μ consists
 212 in measuring variation in substitution patterns along the genome, and to consider the regions of the genome
 213 with the lowest GC^* as a proxy for GC^μ (Lartillot, 2013). Following this logic, we divided the genome of
 214 each species in windows of 200kb, and defined GC^μ as the value of the first percentile of the distribution of
 215 GC^* to avoid outliers (Fig. 2B) (see methods for detailed justifications). This method allows us to estimate
 216 GC^μ for a wide range of species, simply based on substitution patterns in the terminal branch.

217 PRDM9-independent recombination hotspots are active in most mammals

218 Using this estimator of relative recombination rate, we assessed whether MDH loci showed an enrichment of
 219 recombination in other mammals. We identified MDH orthologous loci in the genome of 51 other mammals
 220 and estimated the relative recombination rates at these loci using the method described above (gBGC-
 221 based relative recombination rates). Around 75% of the species (39/52) show a significant enrichment of
 222 recombination in MDH loci compared to flanking regions (Fig. 3).

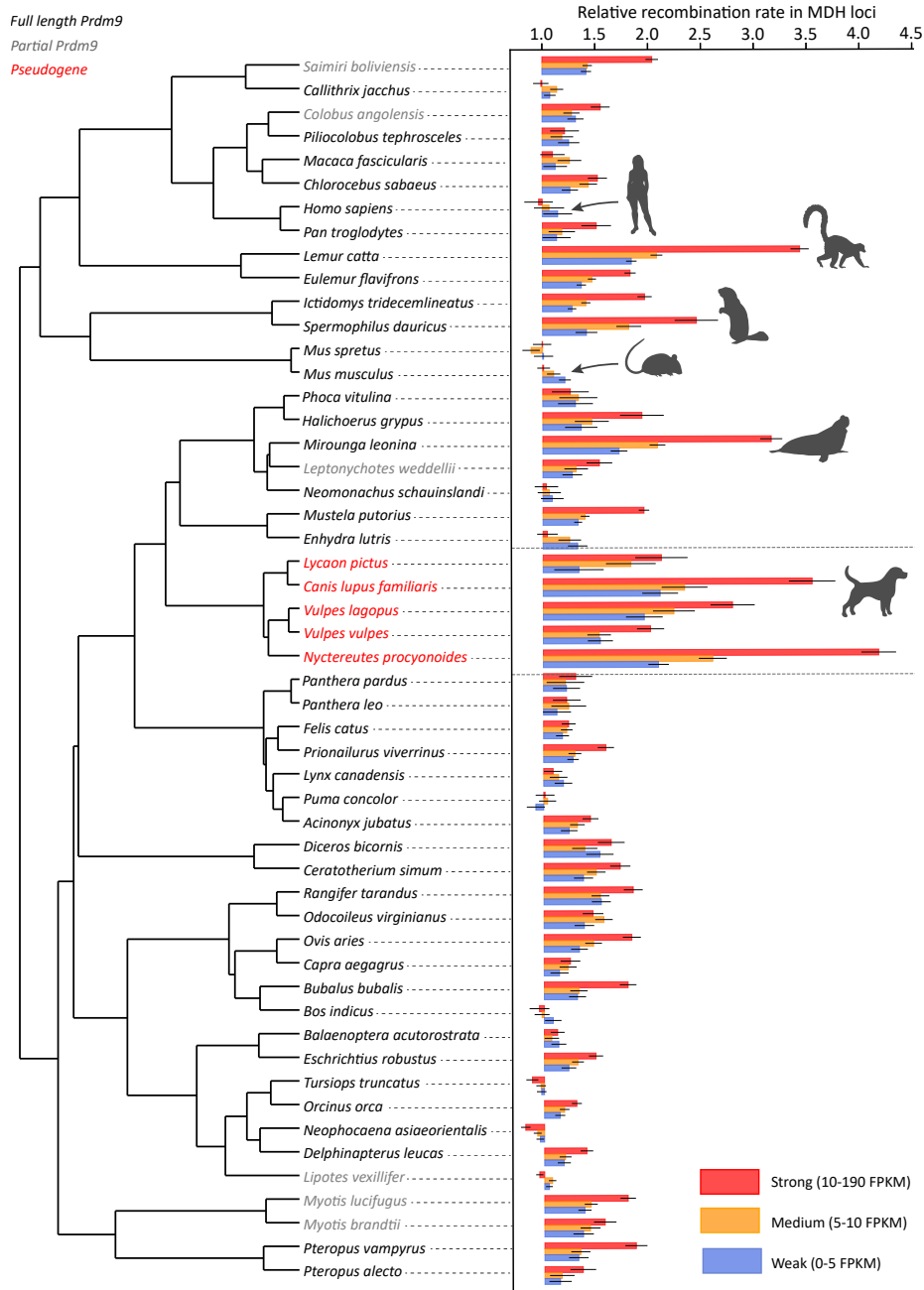


Figure 3: Relative recombination rates at loci orthologous to mouse $Prdm9^{-/-}$ DSB hotspots (MDH loci) in 52 mammals. MDH loci were binned in 3 equally sized categories of strength based on the number of DMC1 Chip-seq reads of $Prdm9^{-/-}$ DSB hotspots. The number of MDH loci for each category varied from $\sim 4,000$ in *Myotis brandtii* to $\sim 9,000$ in *Mus spretus* (see details in Supplementary Table 1). The tree has been retrieved from TimeTree5 (Kumar et al., 2022). Species with a complete PRDM9 are written in black, in grey species for which we failed to find a complete PRDM9 in the reference genome assembly, and in red the 5 canids (where PRDM9 is a pseudogene). Error bars correspond to a 95% confidence interval obtained by bootstrapping the substitutions for computing GC^*_{flank} and GC^*_{hot} . We considered the recombination enrichment to be significant if the confidence interval is above one.

223 In 77% of those species (30/39), the recombination activity at MDH loci is conserved, with strong MDH
224 loci showing a significantly higher recombination enrichment than weak ones (Fig. 3). The remaining 23% of
225 species (9/39) show a lower recombination activity in PRDM9-independent hotspots. Therefore, we might be
226 lacking statistical power to confirm a conservation of hotspot strength in those species (Fig. 3). Interestingly,
227 in 10% of species (5/52), including mice, there is less recombination in strong PRDM9-independent hotspots
228 compared to weak ones, which is consistent with the active deviation of recombination away from those sites
229 observed in those mice having the most dominant PRDM9 alleles (Smagulova et al., 2016). The causes for
230 this active deviation are however still not clear (Smagulova et al., 2016).

231 To test whether as in dogs, methylation plays a role in determining hotspots in other mammals, we
232 separated MDH loci in two subsets: loci for which the hypomethylation pattern is conserved between mice
233 and dogs (and thus likely to be consistently hypomethylated in other mammals as shown in Qu et al. (2018)),
234 and loci that are hypomethylated in mouse but not dogs (a majority of which corresponding to mouse-specific
235 HMRs (Qu et al., 2018)). For MDH loci whose hypomethylation pattern is not conserved, we observed a
236 very weak increase in recombination in all species (Suppfig. S4).

237 Conversely, MDH loci with a conserved hypomethylation pattern, show contrasting levels of recombina-
238 tion activity: 25% of species (13/52, including mice and humans) show a deficit of recombination at these
239 loci, 12% (6/52) show no elevation of recombination, whereas 63% (33/52) show a strong recombination
240 activity. This indicates that DNA hypomethylation is associated with a deficit of recombination in some
241 species (such as mice and humans), while it is associated with recombination hotspot in others. This latter
242 group includes the 5 canids (which all lack Prdm9), but also many other mammals with an intact Prdm9.
243 This shows that DNA hypomethylation can be associated with recombination hotspots even in the presence
244 of Prdm9.

245 To get further insight into the evolution of PRDM9-independent hotspots in mammals, we measured the
246 relative activity at loci orthologous to dog LD-based recombination hotspots (DRH for 'Dog Recombination
247 Hotspots') in the 51 other mammals. We observed a strong correlation between the recombination activity
248 at MDH loci and DRH loci (Suppfig S5), which is expected since they largely overlap. Nevertheless, species
249 that are phylogenetically closer to dogs show higher recombination activity in DRH loci whereas those
250 phylogenetically closer to mouse show higher recombination activity in MDH loci (Suppfig S5). This confirms
251 that despite a general conservation, PRDM9-independent hotspots are still evolving in mammals.

252 Canids, which all lack PRDM9, dominate the list of species that exhibit the highest recombination levels
253 at MDH loci, holding top ranks out of 52 (Fig. 4A). This finding both shows that the recombination landscape
254 is stable in canids as it is in passerines, and validates our relative recombination rate estimator (Fig. 4A).
255 Interestingly, there are several other mammals that show a similar enrichment of recombination activity at
256 MDH loci. Notably, three species show a recombination activity at MDH loci significantly higher than some
257 of the canids: ring-tailed lemurs (*Lemur catta*), southern elephant seals (*Mirounga leonina*) and daurian
258 ground squirrels (*Spermophilus dauricus*) (Fig. 4A).

259 It should be noted that these three species encode a full-length PRDM9, encompassing the four protein
260 domains (KRAB, SSXRD, SET and the zinc finger array). The PRDM9 allele represented in the reference

261 genome assembly contains 11 zinc fingers in daurian ground squirrels, and 14 zinc fingers in ring-tailed
 262 lemurs. For the southern elephant seal, there is only one complete zinc finger represented in the reference
 263 genome because of an assembly gap within the array. In ring-tailed lemurs and daurian ground squirrels,
 264 the comparison of zinc finger sequences showed an excess of amino-acid changes relative to the neutral
 265 expectation, particularly at positions -1, 3, and 6 that are involved in DNA binding, (SuppFig. S6).

266 This suggests that PRDM9 has been subject to positive selection in these two lineages, which is suggestive
 267 of the red-queen dynamic that is expected when PRDM9 determines the location of recombination hotspots.
 268 Thus, PRDM9 does not show any sign of pseudogenization or functional change in these three species where
 269 PRDM9-independent hotspots appear to have been particularly active.

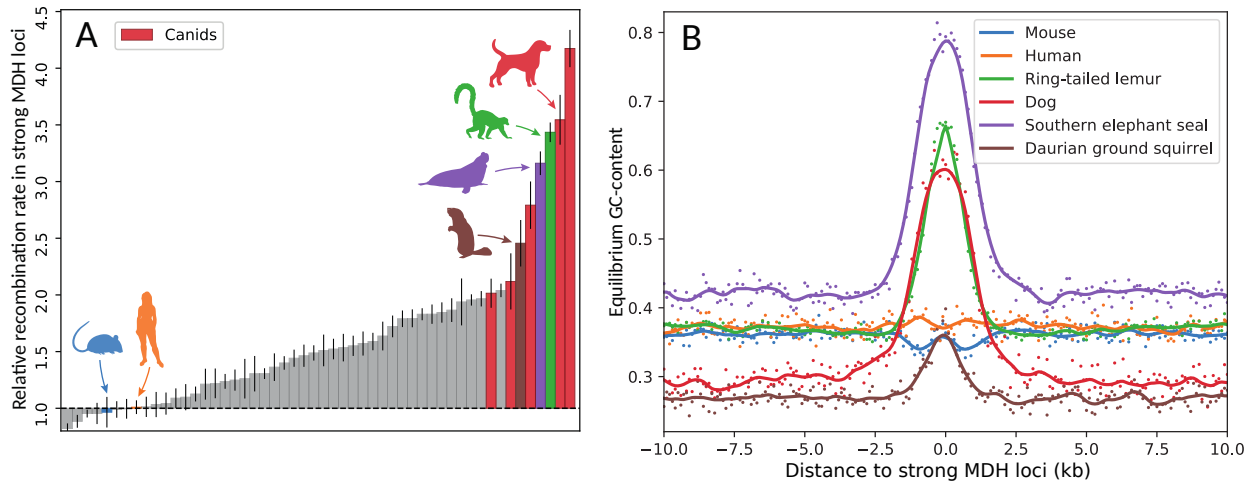


Figure 4: (A): Sorted relative recombination rate in strong DSB hotspots (>10 FPKM) of a PRDM9 $-/-$ mouse in 52 mammals. Error bars correspond to a 95% confidence interval obtained by bootstrapping the substitutions for computing GC_{flank}^* and GC_{hot}^* . (B): GC^* as a function of the distance to the center of the closest MDH locus for humans, mice, dogs, and the three outlier species. Each point corresponds to an estimation of GC^* in a 100 bp window.

270 As an illustration of our substitution-based approach, we plotted GC^* as a function of the distance to
 271 the closest strong MDH loci in dogs, humans, mice and the three outlier species (Fig. 4B). We observed no
 272 elevation of GC^* in humans and mice in strong MDH loci, confirming that very little recombination occurred
 273 in their lineage (Fig. 4B). In dogs, ring-tailed lemurs, southern elephant seals and daurian ground squirrels,
 274 there is a pronounced peak of GC^* at strong MDH loci (Fig. 4B).

275 Discussion

276 The current paradigm is that vertebrates either possess a full-length functional PRDM9, recombine away
 277 from promoter-like features, and display a fast evolving recombination landscape, or they lack a functional
 278 PRDM9 in which case they consistently recombine in CpG islands (e.g. in canids, birds or the swordtail fish)
 279 (Baker et al., 2017). Our results revealed that there is a continuum between these two types of recombination

280 landscapes and that despite the presence of PRDM9, some species use PRDM9-independent hotspots as much
281 as canids. We showed that the activity of those PRDM9-independent hotspots is highly dependent on DNA
282 hypomethylation. This implies that despite a general conservation, PRDM9-independent recombination
283 hotspots are evolving slowly, in concert with germline DNA hypomethylation (Qu et al., 2018; Berglund
284 et al., 2015).

285 From a methodological perspective, while signatures of gBGC have been commonly used in studies of
286 recombination landscapes (Axelsson et al., 2012; Auton et al., 2013; Munch et al., 2014; Lesecque et al.,
287 2014; Singhal et al., 2015; Charlesworth et al., 2020; Hoge et al., 2023), the approach presented here allows
288 to quantify gBGC-based relative recombination rates along a branch that are comparable between species.
289 The analysis of gBGC signatures only requires the genome of three closely related species, and is able to
290 detect recombination activity at a given set of loci with a very high spatial resolution, even better than LD-
291 based methods (compare Fig. 2D and 2B). It thus offers the possibility for large-scale comparative studies
292 of fine-scale recombination landscapes. However, this approach requires a large number of substitutions to
293 estimate GC^* precisely and is therefore not appropriate to measure recombination at a single locus.

294 Moreover, it should be noted that our estimation of recombination activity using gBGC does not allow one
295 to conclude on the nature of recombination events (COs or NCOs) that we detect at PRDM9-independent
296 hotspots. In humans, there is evidence that both COs and NCOs induce gBGC but in mice, only NCOs
297 appear to do so (Williams et al., 2015; Arbeithuber et al., 2015; Halldorsson et al., 2016; Li et al., 2019).
298 This suggests that the type of recombination events triggering gBGC can vary among mammals. Thus, it
299 is possible that while a large number of DSB occur at PRDM9-independent hotspots in our three outlier
300 species, COs still tend to be associated to PRDM9-directed hotspots. Altogether, while the enrichment of
301 recombination events in PRDM9-independent hotspots is very clear in numerous mammals, the way they
302 are repaired remains to be explored.

303 **On the maintenance of a default hotspot regulation mechanism**

304 It had been previously demonstrated that both PRDM9-dependent and PRDM9-independent pathways
305 coexist in two snakes genera (Hoge et al., 2023). The authors suggested that a change in the binding affinity
306 of a gene which operates downstream of PRDM9 could explain this coexistence and that selection might be
307 operating on these genes to fine-tune the usage of PRDM9-independent hotspots in vertebrates (Hoge et al.,
308 2023).

309 In placental mammals we revealed that the coexistence of both PRDM9-dependent and PRDM9-
310 independent pathways to direct DSBs is pervasive. We showed that those pathways determine recombination
311 hotspots with varying proportions across species. Interestingly, these variations do not show a strong phylo-
312 genetic structure, suggesting that this evolution can be very rapid. Furthermore, the species with the highest
313 levels of PRDM9-independent hotspot usage have quite contrasted life history traits, which they mostly share
314 with sister species with lower PRDM9-independent hotspot usage. Thus, it is difficult to imagine that dif-
315 ferences of PRDM9-independent usage depend on reproductive life history traits. Moreover, in mice, even if
316 the usage of PRDM9-independent hotspots is quite low overall, there exist substantial variations that seem

317 to depend only on PRDM9 alleles (Smagulova et al., 2016). A plausible explanation for the maintenance of
318 the PRDM9-independent pathway in placental mammals would be that variation in PRDM9-independent
319 hotspots usage depends on the efficacy of PRDM9 alleles to recruit the DSB machinery.

320 Overall, proper chromosome pairing can be achieved through the two different pathways mentioned
321 above (PRDM9-dependent or PRDM9-independent). When the efficiency of one pathway is altered, better
322 chromosome pairing can be restored either by a mutation restoring the efficiency of the altered pathway,
323 or by a mutation increasing the efficiency of the other pathway. For the PRDM9 pathway, we know that
324 alleles inevitably decrease in efficiency due to the erosion of their high affinity targets, which reduces the
325 probability of symmetrical binding, and thus impairs efficient chromosome pairing (Baker et al., 2015, 2022;
326 Latrille et al., 2017; Genestier et al., 2023). This efficiency can be restored either by a new PRDM9 allele
327 inducing a red-queen dynamic (Úbeda and Wilkins, 2011; Latrille et al., 2017; Baker et al., 2022; Genestier
328 et al., 2023), but also by a mutation increasing the efficiency of the PRDM9-independent pathway. Every
329 mutation increasing the efficiency of the PRDM9-independent pathway lessens the deleterious effect of a
330 mutation that reduces PRDM9 efficiency. Conversely, new efficient PRDM9 alleles will lessen the deleterious
331 effect of a mutation that decreases the efficiency of the PRDM9-independent pathway.

332 Of note, if this dynamic reaches a point where the PRDM9-independent pathway becomes sufficient for
333 correct chromosome pairing, PRDM9 can be lost without strong fitness consequences. Under this model,
334 PRDM9 is lost through the accumulation of small effect mutations which reduce its utility, rather than a
335 sudden loss that would imply very inefficient selection. The continuum in the use of PRDM9-independent
336 hotspots we observe in mammals could reflect different stages along this path, and despite still having a
337 fully functional PRDM9, the three outlier species could be on their way of losing it. It is also possible that
338 their PRDM9 have only been going through a temporary inefficient phase compensated by the PRDM9-
339 independent pathway, but has now been rescued by a new efficient PRDM9 allele.

340 **The recombination landscape of amniotes**

341 Overall, our results suggest that in addition to PRDM9-directed hotspots, many mammals share some
342 of their recombination hotspots with other amniotes, (Axelsson et al., 2012; Singhal et al., 2015; Schield
343 et al., 2020; Hoge et al., 2023), and therefore the fine-scale recombination landscapes of mammals, birds and
344 snakes is probably more similar than previously thought (Baker et al., 2017). However, the determinants of
345 PRDM9-independent hotspots usage remain unclear. Interestingly DNA methylation has also been found
346 to be a suppressor of recombination in a fungi hotspot (Maloisel and Rossignol, 1998), in plants (He et al.,
347 2017; Choi et al., 2018), and in honey bees (Wallberg et al., 2015), which suggests that local hypomethyla-
348 tion is a common determinant of recombination hotspots in eukaryotes. However, the association between
349 hypomethylation and recombination has not been formally established in non-mammalian amniotes and
350 remains to be tested. Moreover, this association need not be causal, as the potentially diverse molecular
351 mechanisms of PRDM9-independent recombination in amniotes remain largely unknown. In particular, the
352 results presented here suggest that despite having lost PRDM9, red foxes (*Vulpes vulpes*) and african wild
353 dogs (*Lycaon pictus*) have only a mild recombination enrichment in PRDM9-independent hotspots compared
354 to other canids. It has been previously noted that the number of recombination hotspots varies between

355 PRDM9-deficient amniotes. Notably in finches and flycatchers, only few LD-based hotspots have been re-
356 ported compared to dogs (Auton et al., 2013; Singhal et al., 2015; Kawakami et al., 2017). Altogether, it is
357 still not clear what drives the concentration of recombination events in absence of PRDM9, and why some
358 species have numerous hotspots and others less.

359 Finally, the widespread use of PRDM9-independent recombination hotspots demonstrated in the present
360 study is likely to have important consequences for genome evolution. In particular, the fact that gBGC is
361 stronger in hypomethylated regions in numerous mammals and in several passerines provides a convincing
362 explanation for the widespread GC-richness of CpG islands in amniotes.

363 Material & Methods

364 PRDM9-independent hotspots and DNA methylation datasets

365 Data on the location of DSB hotspots in *Prdm9*^{-/-} knock-out mice, detected by DMC1 ChIP-seq experiment,
366 were retrieved from the study by Smagulova et al. (2016). The dog LD-based recombination map and the
367 position of hotspots were retrieved from the study of Auton et al. (2013). We excluded hotspots that were
368 larger than 20 kb as they do not fit the definition of hotspots. Data sets of hypomethylated regions in mouse
369 and dog sperm (identified using Bi-sulfite sequencing) were retrieved from the literature (Hammoud et al.,
370 2014; Qu et al., 2018). Even though methylation data on spermatocytes would have been more fit for the
371 task at hand, only sperm is available in the literature for dogs.

372 Whole genome alignments

373 For most mammals except canids, felids and phocids, whole genome alignments (WGAs) were obtained from
374 Genereux et al. (2020). In order to get further phylogenetic resolution in canids, and in closely related
375 outgroups, we generated a WGA of high quality genomes for 19 carnivores downloaded from NCBI (Supple-
376 mentary Table 1), using the Progressive Cactus aligner (v1.3.0) (Armstrong et al., 2020). We first defined a
377 "guide" species tree using the topology obtained from TimeTree5 (Kumar et al., 2022). To streamline the
378 computational process, we ran Progressive Cactus separately for canids, felids and phocids species, using
379 different "-root" options on the same guide tree. We created a root alignment by running Progressive Cactus
380 with the inferred ancestral genome of each of the three clades. We obtained the final WGA using the "ha-
381 lAppendSubtree" command to iteratively include the three sub-alignments at the corresponding ancestral
382 nodes (Hickey et al., 2013).

383 Defining orthologous regions

384 To find the orthologous regions of the PRDM9-independent hotspots in the genomes of other mammals we
385 used *halLiftOver* (Hickey et al., 2013). We first made a liftover from the mouse/dog genome to the target
386 genome using the midpoint of each feature and removed multi-mapping features. Then we lifted back the
387 single-mapping features from the target genome to the dog/mouse genome and again removed multi-mapping
388 features. This approach ensures that all orthologous loci were one-to-one.

389 Hotspots overlap

390 We considered hotspots to be overlapping if their midpoint was at less than 5 kb one from another. This is
 391 equivalent to a strict overlap for hotspots defined as 5 kb windows centered on their midpoint. Using this
 392 approach, we calculated the percentage of the genome covered by the hotspots by multiplying the number
 393 of hotspots by 5,000, and dividing by the assembly size. For the overlap with HMRs we defined hotspots as
 394 the 5 kb windows centered on their midpoint and kept the size of HMRs defined in the study of Qu et al.
 395 2018 (Qu et al., 2018). The percentage of the genome covered by HMRs was computed as the sum of all
 396 HMR sizes divided by the assembly size.

397 Substitution mapping

398 We selected trios of closely related species such that the divergence between the three species was low enough
 399 to avoid double substitutions, but with an outgroup distant enough to avoid incomplete lineage sorting based
 400 on the guide tree used in the Zoonomia WGA (Genereux et al., 2020). We tried to take trios spanning the
 401 diversity of boreoeutherians, avoiding over-sampling of disproportionately represented groups (primates and
 402 artiodactyles). A complete list of the trios used are available in Supplementary table 3. *A posteriori*, it
 403 appeared that there were substantial variations between the branch lengths used to map substitutions, but
 404 we showed that the divergence was still low enough ($< 2.5\%$) and did not influence our result (Suppfig.
 405 S7). Genome quality was very variable. We thus controlled that genome quality (approximated by the
 406 N50 statistics) did not influence our results (Suppfig. S8). To call substitutions, we retrieved multispecies
 407 alignment using hal2maf (Hickey et al., 2013). We excluded alignment blocks which size was inferior to 50
 408 bp to avoid poorly aligned regions, and duplicated regions. We then excluded CpG sites (sites for which at
 409 least one of the three species has a CpG) to avoid convergent mutations. Finally, we called substitutions
 410 using parsimony as depicted in Fig. 2.

411 Measures of equilibrium GC content

412 We compute the equilibrium GC content as follows:

$$GC^* = \frac{WS/W}{WS/W + SW/S} \quad (2)$$

413 With W the AT content of the region (CpG masked), S the GC content of the region (CpG masked),
 414 WS the number of Weak to Strong substitutions and SW the number of Strong to Weak substitutions.

415 LD-based recombination rate and GC^* profiles around hotspots

416 We cut the genome in windows of 100 bp. We extracted the LD-based recombination rate for each window.
 417 We also computed the distance between the midpoint of the window and the closest midpoint of a hotspot
 418 using bedtools closest (Quinlan and Hall, 2010). We then made bins of distances to the closest hotspot
 419 every 100 bp. For GC^* , we repeated the same procedure, but we computed the counts of WS and SW

420 substitution for each window, and then computed GC^* using the total substitutions count for all windows
 421 in each distance bin as described in the previous section.

422 **Estimation of the mutation bias**

423 Germline mutations rates have been measured by sequencing parent–offspring trios in 36 mammalian species
 424 (Bergeron et al., 2023). However, for most species, the number of detected mutations is too limited (typically
 425 less than 100 de novo mutations) to estimate GC^μ accurately. Thus, measures of GC^μ based on empirical
 426 data are associated with very large confidence intervals (Wong et al., 2016; Milholland et al., 2017; Wang
 427 et al., 2020). Even when there is enough statistical power, the results vary substantially between different
 428 datasets (Wong et al., 2016; Milholland et al., 2017). In addition to the issue of reproducibility, it has been
 429 demonstrated that the mutation spectra can vary rapidly within human populations (Harris and Pritchard,
 430 2017). Therefore, GC^μ estimated from living individuals may not necessarily reflect the average GC^μ in
 431 the terminal branch. We therefore took a similar approach to Lartillot (2013). We divided the genomes in
 432 windows of 200 kb and took the value of the first percentile of GC^* as an estimate of GC^μ . To ensure that
 433 our results were not sensitive to our estimation of GC^μ , we used different thresholds to compute it (Suppfig.
 434 S3B&C), and recovered the same results. We also controlled that our gBGC-based relative recombination
 435 rates were not correlated to our estimations of GC^μ (Suppfig. S3A).

436 **Estimation of relative recombination rates from GC^***

437 Using a Wright-Fischer diffusion approximation and assuming that mutations are selectively neutral, the
 438 rate of Weak-to-Strong substitution in a given branch can be written as follows (Nagylaki, 1983)

$$q_{WS} = 2Ne\mu_{WS} \frac{2b}{1 - e^{-4Neb}} T \quad (3)$$

439 where μ_{SW} is the mutation rate per generation from W to S, b the gBGC coefficient, T the divergence
 440 time from the ancestral node in generations, and Ne the effective population size.

441 The gBGC coefficient is directly linked to the recombination rate, with $b = b_0 r l$ where r is the recombi-
 442 nation rate per base pair per meiosis, b_0 the repair bias, and l the length of the conversion tract in base pair.
 443 It should be noted that r encompasses all recombination events that can lead to gBGC (CO and/or NCO)

444 Similarly, the rate of Strong-to-Weak substitutions can be written as follows:

$$q_{SW} = 2Ne\mu_{SW} \frac{2b}{e^{4Neb} - 1} T \quad (4)$$

445 The equilibrium GC content can be written as follows:

$$GC^* = \frac{q_{WS}}{q_{WS} + q_{SW}} \quad (5)$$

446 Thus, we can write:

$$\frac{GC^*}{1 - GC^*} = \frac{q_{WS}}{q_{SW}} \quad (6)$$

447 Simplifying the previous equations we obtain:

$$\frac{GC^*}{1 - GC^*} = \frac{\mu_{WS}}{\mu_{SW}} e^{4Neb} \quad (7)$$

448 Thus, the population-scaled gBGC coefficient. ($B = 4Neb$) can be written as:

$$B = \log\left(\frac{GC^*}{1 - GC^*}\right) - \log\left(\frac{\mu_{WS}}{\mu_{SW}}\right) \quad (8)$$

449 or:

$$B = \text{logit}(GC^*) - \text{logit}(GC^\mu) \quad (9)$$

450 where GC^μ is the equilibrium GC content under the mutational bias only ($GC^\mu = \frac{\mu_{WS}}{\mu_{SW} + \mu_{WS}}$).

451 Let us note B_{hot} the population-scaled gBGC coefficient within hotspots:

$$B_{hot} = \text{logit}(GC_{hot}^*) - \text{logit}(GC_{hot}^\mu) \quad (10)$$

452 And B_{flank} the population-scaled gBGC coefficient in their flanking regions (defined here as 3kb-long
453 segments, located at 5-kb of the hotspot center, Fig. 2A)

$$B_{flank} = \text{logit}(GC_{flank}^*) - \text{logit}(GC_{flank}^\mu) \quad (11)$$

454 B depends on b_0 , N_e , l and r ($B = 4Nerb_0l$). The first three parameters (b_0 , N_e , l) are not expected to
455 differ between the hotspot and their flanking regions. Thus, the ratio between the recombination rate within
456 hotspot (r_{hot}) over the recombination rate in their flanking regions (r_{flank}) can be written as:

$$\frac{r_{hot}}{r_{flank}} = \frac{B_{hot}}{B_{flank}} \quad (12)$$

457 And thus, under the assumption that the mutational bias does not differ between hotspots and their
458 flanking regions (i.e. $GC_{hot}^\mu = GC_{flank}^\mu = GC^\mu$)

$$\frac{r_{hot}}{r_{flank}} = \frac{\text{logit}(GC_{hot}^*) - \text{logit}(GC^\mu)}{\text{logit}(GC_{flank}^*) - \text{logit}(GC^\mu)} \quad (13)$$

459 Annotation of PRDM9 in mammals

460 We investigated the presence of PRDM9 homologs in each of the 52 species analyzed. The full-length PRDM9
461 isoform encompasses four domains (KRAB, SSXRD, SET and the zinc finger array). It is encoded by 10
462 exons (corresponding to exons 2 to 11 of human PRDM9, exon 1 being within the 5'UTR): exons 3 and 4
463 encode the KRAB domain, exon 7 encodes the SSXRD domain, exons 8-10 encode the SET domain, and exon

464 11 encodes the zinc finger array. We first searched for PRDM9 homologs by sequence similarity (Camacho
 465 et al., 2009) against mammalian proteins annotated in RefSeq (<https://www.ncbi.nlm.nih.gov/blast/>), using
 466 the human protein (NP_001363829.1) as a query. We performed a multiple alignment of the strongest hits,
 467 to assess their completeness: homologs were considered as complete if they encompassed the 10 protein-
 468 coding exons, from the start codon, up to the beginning of the zinc finger array. By this approach, we
 469 identified complete PRDM9 homologs in 16 species (Supplementary Table 2). For the 36 other species, we
 470 further analyzed the corresponding reference genome to identify potential PRDM9 homologs. We performed
 471 a TBLASTN search against reference genomes, using the 16 previously identified PRDM9 as queries, and
 472 extracted loci presenting hits with the zinc finger domain and with at least 2 other PRDM9 exons, within
 473 less than 100 kb. Then, for each candidate genomic fragment, we used GeneWise (Birney et al., 2004) to
 474 annotate protein-coding regions by similarity with a representative complete PRDM9 protein taken from a
 475 closely related species. Of note, GeneWise does take frameshifts into account, and is therefore appropriate to
 476 annotate both genes and pseudogenes. In most species we identified one single candidate locus per genome.
 477 When several loci were found, we retained the one(s) encoding the most complete protein. By this approach,
 478 we identified PRDM9 loci in all 36 species, 25 of which encode a complete PRDM9 protein. Thus, in
 479 total, we identified complete PRDM9 proteins in 41 of the 52 species analyzed. In agreement with previous
 480 reports ((Axelsson et al., 2012; Auton et al., 2013)), we found PRDM9 to be pseudogenized in the 5 canids
 481 (*Lycaon pictus*, *Canis lupus familiaris*, *Vulpes lagopus*, *Vulpes vulpes*, *Nyctereutes procyonoides*). In the 6
 482 remaining cases, we failed to annotate a complete PRDM9 protein: either several exons were missing (*Myotis*
 483 *lucifugus*, *Myotis brandtii*, *Leptonychotes weddellii*, *Colobus angolensis*), or the gene contained one exon with
 484 a frameshifting mutation (respectively in exon 8 in *Saimiri boliviensis* and in exon 10 in *Lipotes vexillifer*).
 485 In absence of data from more individuals, it is difficult to state whether these cases result from sequencing
 486 errors or assembly artefacts or if they correspond to bona fide pseudogenes. We therefore tentatively annotate
 487 these 6 cases as ‘partial’ Prdm9. The number of zinc fingers in the annotated proteins vary from 0 to 14.
 488 This number should however be considered with caution because the zinc finger array is encoded by a highly
 489 polymorphic minisatellite repeat, which is prone to errors during genome assembly. The detailed list of
 490 PRDM9 sequences is given in Supplementary Table 2 and the corresponding protein multiple alignment can
 491 be found at <https://zenodo.org/records/10149667>.

492 Acknowledgments

493 We wish to thank Nicolas Lartillot, Carina Farah Mugal, and Bernard de Massy for very useful reviews on a
 494 previous version of this manuscript, and Anaïs Duhamel for help with the figures. This work was performed
 495 using the computing facilities of the CC LBBE/PRABI. **Funding:** Agence Nationale de la Recherche, Grant
 496 ANR-19-CE12-0019 / HotRec. **Author contributions:** Original idea: J.J.; Model conception: J.J.; Code:
 497 J.J., D.P, A.L, T.T and L.D; Data analyses: J.J., D.P, A.L, T.T and L.D; Interpretation: J.J., D.P, A.L, T.T
 498 and L.D; First draft: J.J.; Editing and revisions: J.J., D.P, A.L, T.T and L.D; Funding: L.D. **Competing**
 499 **interests:** The authors declare no conflicts of interest.

500 **Data availability**

501 Analysis scripts and documentation for whole genome alignments are available at
 502 <https://github.com/AlexandreLaverre/CANCGI-WGA/> The hal file of the Carnivore WGA, along
 503 with analysis scripts and documentation for PRDM9 annotation and the study of positive selection can
 504 be found at <https://zenodo.org/records/10149667>. The remaining analysis scripts and documentation are
 505 accessible at <https://gitlab.in2p3.fr/julien.joseph/defhot>.

506 **References**

- 507 Alleva, B., Brick, K., Pratto, F., Huang, M., and Camerini-Otero, R. D. (2021). Cataloging Human PRDM9
 508 Allelic Variation Using Long-Read Sequencing Reveals PRDM9 Population Specificity and Two Distinct
 509 Groupings of Related Alleles. *Frontiers in Cell and Developmental Biology*, 9.
- 510 Arbeithuber, B., Betancourt, A. J., Ebner, T., and Tiemann-Boege, I. (2015). Crossovers are associated with
 511 mutation and biased gene conversion at recombination hotspots. *Proceedings of the National Academy of
 512 Sciences*, 112(7):2109–2114.
- 513 Armstrong, J., Hickey, G., Diekhans, M., Fiddes, I. T., Novak, A. M., Deran, A., Fang, Q., Xie, D., Feng,
 514 S., Stiller, J., Genereux, D., Johnson, J., Marinescu, V. D., Alföldi, J., Harris, R. S., Lindblad-Toh,
 515 K., Haussler, D., Karlsson, E., Jarvis, E. D., Zhang, G., and Paten, B. (2020). Progressive Cactus is a
 516 multiple-genome aligner for the thousand-genome era. *Nature*, 587(7833):246–251.
- 517 Auton, A., Fledel-Alon, A., Pfeifer, S., Venn, O., Ségurel, L., Street, T., Leffler, E. M., Bowden, R., Aneas,
 518 I., Broxholme, J., Humburg, P., Iqbal, Z., Lunter, G., Maller, J., Hernandez, R. D., Melton, C., Venkat,
 519 A., Nobrega, M. A., Bontrop, R., Myers, S., Donnelly, P., Przeworski, M., and McVean, G. (2012). A
 520 Fine-Scale Chimpanzee Genetic Map from Population Sequencing. *Science*, 336:7.
- 521 Auton, A. and McVean, G. (2007). Recombination rate estimation in the presence of hotspots. *Genome
 522 Res.*, 17(8):1219–1227.
- 523 Auton, A., Rui Li, Y., Kidd, J., Oliveira, K., Nadel, J., Holloway, J. K., Hayward, J. J., Cohen, P. E., Grealley,
 524 J. M., Wang, J., Bustamante, C. D., and Boyko, A. R. (2013). Genetic Recombination Is Targeted towards
 525 Gene Promoter Regions in Dogs. *PLoS Genet*, 9(12):e1003984.
- 526 Axelsson, E., Webster, M. T., Ratnakumar, A., Consortium, T. L., Ponting, C. P., and Lindblad-Toh, K.
 527 (2012). Death of PRDM9 coincides with stabilization of the recombination landscape in the dog genome.
 528 *Genome Res.*, 22(1):51–63.
- 529 Baker, C. L., Kajita, S., Walker, M., Saxl, R. L., Raghupathy, N., Choi, K., Petkov, P. M., and Paigen, K.
 530 (2015). PRDM9 Drives Evolutionary Erosion of Hotspots in *Mus musculus* through Haplotype-Specific
 531 Initiation of Meiotic Recombination. *PLOS Genetics*, 11(1):e1004916.
- 532 Baker, Z., Przeworski, M., and Sella, G. (2022). Down the Penrose stairs: How selection for fewer recombi-
 533 nation hotspots maintains their existence.

- 534 Baker, Z., Schumer, M., Haba, Y., Bashkirova, L., Holland, C., Rosenthal, G. G., and Przeworski, M. (2017).
 535 Repeated losses of PRDM9-directed recombination despite the conservation of PRDM9 across vertebrates.
 536 *eLife*, 6:e24133.
- 537 Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G., and de Massy, B.
 538 (2010). PRDM9 Is a Major Determinant of Meiotic Recombination Hotspots in Humans and Mice. *Science*,
 539 327(5967):836–840.
- 540 Bergeron, L. A., Besenbacher, S., Zheng, J., Li, P., Bertelsen, M. F., Quintard, B., Hoffman, J. I., Li, Z.,
 541 St. Leger, J., Shao, C., Stiller, J., Gilbert, M. T. P., Schierup, M. H., and Zhang, G. (2023). Evolution of
 542 the germline mutation rate across vertebrates. *Nature*, pages 1–7.
- 543 Berglund, J., Quilez, J., Arndt, P. F., and Webster, M. T. (2015). Germline Methylation Patterns Determine
 544 the Distribution of Recombination Events in the Dog Genome. *Genome Biology and Evolution*, 7(2):522–
 545 530.
- 546 Birney, E., Clamp, M., and Durbin, R. (2004). GeneWise and Genomewise. *Genome Res.*, 14(5):988–995.
- 547 Bolívar, P., Mugal, C. F., Nater, A., and Ellegren, H. (2016). Recombination Rate Variation Modulates
 548 Gene Sequence Evolution Mainly via GC-Biased Gene Conversion, Not Hill–Robertson Interference, in an
 549 Avian System. *Mol Biol Evol*, 33(1):216–227.
- 550 Brick, K., Smagulova, F., Khil, P., Camerini-Otero, R. D., and Petukhova, G. V. (2012). Genetic recombi-
 551 nation is directed away from functional genomic elements in mice. *Nature*, 485(7400):642–645.
- 552 Buard, J., Rivals, E., Segonzac, D. D. d., Garres, C., Caminade, P., Massy, B. d., and Boursot, P. (2014).
 553 Diversity of Prdm9 Zinc Finger Array in Wild Mice Unravels New Facets of the Evolutionary Turnover of
 554 this Coding Minisatellite. *PLOS ONE*, 9(1):e85021.
- 555 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T. L. (2009).
 556 BLAST+: architecture and applications. *BMC Bioinformatics*, 10(1):421.
- 557 Cavassim, M. I. A., Baker, Z., Hoge, C., Schierup, M. H., Schumer, M., and Przeworski, M. (2022). PRDM9
 558 losses in vertebrates are coupled to those of paralogs ZCWPW1 and ZCWPW2. *Proceedings of the National
 559 Academy of Sciences*, 119(9):e2114401119.
- 560 Chan, A. H., Jenkins, P. A., and Song, Y. S. (2012). Genome-Wide Fine-Scale Recombination Rate Variation
 561 in *Drosophila melanogaster*. *PLOS Genetics*, 8(12):e1003090.
- 562 Charlesworth, D., Zhang, Y., Bergero, R., Graham, C., Gardner, J., and Yong, L. (2020). Using GC Content
 563 to Compare Recombination Patterns on the Sex Chromosomes and Autosomes of the Guppy, *Poecilia
 564 reticulata*, and Its Close Outgroup Species. *Molecular Biology and Evolution*, 37(12):3550–3562.
- 565 Choi, K., Zhao, X., Tock, A. J., Lambing, C., Underwood, C. J., Hardcastle, T. J., Serra, H., Kim, J.,
 566 Cho, H. S., Kim, J., Ziolkowski, P. A., Yelina, N. E., Hwang, I., Martienssen, R. A., and Henderson,
 567 I. R. (2018). Nucleosomes and DNA methylation shape meiotic DSB frequency in *Arabidopsis thaliana*
 568 transposons and gene regulatory regions. *Genome Res.*, 28(4):532–546.
- 569 Clément, Y. and Arndt, P. F. (2013). Meiotic Recombination Strongly Influences GC-Content Evolution in
 570 Short Regions in the Mouse Genome. *Molecular Biology and Evolution*, 30(12):2612–2618.

- 571 Davies, B., Hatton, E., Altemose, N., Hussin, J. G., Pratto, F., Zhang, G., Hinch, A. G., Moralli, D., Biggs,
572 D., Diaz, R., Preece, C., Li, R., Bitoun, E., Brick, K., Green, C. M., Camerini-Otero, R. D., Myers,
573 S. R., and Donnelly, P. (2016). Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice.
574 *Nature*, 530(7589):171–176.
- 575 Diagouraga, B., Clément, J. A. J., Duret, L., Kadlec, J., de Massy, B., and Baudat, F. (2018). PRDM9
576 Methyltransferase Activity Is Essential for Meiotic DNA Double-Strand Break Formation at Its Binding
577 Sites. *Molecular Cell*, 69(5):853–865.e6.
- 578 Duret, L. and Arndt, P. F. (2008). The Impact of Recombination on Nucleotide Substitutions in the Human
579 Genome. *PLOS Genetics*, 4(5):e1000071.
- 580 Eyre-Walker, A. (1999). Evidence of Selection on Silent Site Base Composition in Mammals: Potential
581 Implications for the Evolution of Isochores and Junk DNA. *Genetics*, 152(2):675–683.
- 582 Figuet, E., Ballenghien, M., Romiguier, J., and Galtier, N. (2015). Biased Gene Conversion and GC-Content
583 Evolution in the Coding Sequences of Reptiles and Vertebrates. *Genome Biology and Evolution*, 7(1):240–
584 250.
- 585 Galtier, N. (2021). Fine-scale quantification of GC-biased gene conversion intensity in mammals. *Peer*
586 *Community Journal*, 1.
- 587 Galtier, N., Roux, C., Rousselle, M., Romiguier, J., Figuet, E., Glémin, S., Bierne, N., and Duret, L. (2018).
588 Codon Usage Bias in Animals: Disentangling the Effects of Natural Selection, Effective Population Size,
589 and GC-Biased Gene Conversion. *Molecular Biology and Evolution*, 35(5):1092–1103.
- 590 Genereux, D. P., Serres, A., Armstrong, J., Johnson, J., Marinescu, V. D., Murén, E., Juan, D., Bejerano,
591 G., Casewell, N. R., Chemnick, L. G., Damas, J., Di Palma, F., Diekhans, M., Fiddes, I. T., Garber,
592 M., Gladyshev, V. N., Goodman, L., Haerty, W., Houck, M. L., Hubley, R., Kivioja, T., Koepfli, K.-P.,
593 Kuderna, L. F. K., Lander, E. S., Meadows, J. R. S., Murphy, W. J., Nash, W., Noh, H. J., Nweeia, M.,
594 Pfenning, A. R., Pollard, K. S., Ray, D. A., Shapiro, B., Smit, A. F. A., Springer, M. S., Steiner, C. C.,
595 Swofford, R., Taipale, J., Teeling, E. C., Turner-Maier, J., Alföldi, J., Birren, B., Ryder, O. A., Lewin,
596 H. A., Paten, B., Marques-Bonet, T., Lindblad-Toh, K., Karlsson, E. K., and Zoonomia Consortium (2020).
597 A comparative genomics multitool for scientific discovery and conservation. *Nature*, 587(7833):240–245.
- 598 Genestier, A., Duret, L., and Lartillot, N. (2023). Bridging the gap between the evolutionary dynamics and
599 the molecular mechanisms of meiosis: a model based exploration of the PRDM9 intra-genomic Red Queen.
- 600 Gerton, J. L. and Hawley, R. S. (2005). Homologous chromosome interactions in meiosis: diversity amidst
601 conservation. *Nat Rev Genet*, 6(6):477–487.
- 602 Glémin, S. (2010). Surprising Fitness Consequences of GC-Biased Gene Conversion: I. Mutation Load and
603 Inbreeding Depression. *Genetics*, 185(3):939–959.
- 604 Glémin, S., Arndt, P. F., Messer, P. W., Petrov, D., Galtier, N., and Duret, L. (2015). Quantification of
605 GC-biased gene conversion in the human genome. *Genome Res.*, 25(8):1215–1228.
- 606 Halldorsson, B. V., Hardarson, M. T., Kehr, B., Styrkarsdottir, U., Gylfason, A., Thorleifsson, G., Zink,
607 F., Jonasdottir, A., Jonasdottir, A., Sulem, P., Masson, G., Thorsteinsdottir, U., Helgason, A., Kong, A.,

- 608 Gudbjartsson, D. F., and Stefansson, K. (2016). The rate of meiotic gene conversion varies by sex and
609 age. *Nat Genet*, 48(11):1377–1384.
- 610 Hammoud, S. S., Low, D. H. P., Yi, C., Carrell, D. T., Guccione, E., and Cairns, B. R. (2014). Chromatin
611 and Transcription Transitions of Mammalian Adult Germline Stem Cells and Spermatogenesis. *Cell Stem*
612 *Cell*, 15(2):239–253.
- 613 Harris, K. and Pritchard, J. K. (2017). Rapid evolution of the human mutation spectrum. *eLife*, 6:e24284.
- 614 He, Y., Wang, M., Dukowic-Schulze, S., Zhou, A., Tiang, C.-L., Shilo, S., Sidhu, G. K., Eichten, S., Bradbury,
615 P., Springer, N. M., Buckler, E. S., Levy, A. A., Sun, Q., Pillardy, J., Kianian, P. M. A., Kianian, S. F.,
616 Chen, C., and Pawlowski, W. P. (2017). Genomic features shaping the landscape of meiotic double-strand-
617 break hotspots in maize. *Proceedings of the National Academy of Sciences*, 114(46):12231–12236.
- 618 Hickey, G., Paten, B., Earl, D., Zerbino, D., and Haussler, D. (2013). HAL: a hierarchical format for storing
619 and analyzing multiple genome alignments. *Bioinformatics*, 29(10):1341–1342.
- 620 Hinch, A. G., Zhang, G., Becker, P. W., Moralli, D., Hinch, R., Davies, B., Bowden, R., and Donnelly, P.
621 (2019). Factors influencing meiotic recombination revealed by whole-genome sequencing of single sperm.
622 *Science*, 363(6433):eaau8861.
- 623 Hoge, C. R., Manuel, M. d., Mahgoub, M., Okami, N., Fuller, Z. L., Banerjee, S., Baker, Z., McNulty, M.,
624 Andolfatto, P., Macfarlan, T. S., Schumer, M., Tzika, A. C., and Przeworski, M. (2023). Patterns of
625 recombination in snakes reveal a tug of war between PRDM9 and promoter-like features.
- 626 Kawakami, T., Mugal, C. F., Suh, A., Nater, A., Burri, R., Smeds, L., and Ellegren, H. (2017). Whole-
627 genome patterns of linkage disequilibrium across flycatcher populations clarify the causes and consequences
628 of fine-scale recombination rate variation in birds. *Mol Ecol*, 26(16):4158–4172.
- 629 Kono, H., Tamura, M., Osada, N., Suzuki, H., Abe, K., Moriwaki, K., Ohta, K., and Shiroishi, T. (2014).
630 Prdm9 Polymorphism Unveils Mouse Evolutionary Tracks. *DNA Research*, 21(3):315–326.
- 631 Kumar, S., Suleski, M., Craig, J. M., Kasprowitz, A. E., Sanderford, M., Li, M., Stecher, G., and Hedges,
632 S. B. (2022). TimeTree 5: An Expanded Resource for Species Divergence Times. *Molecular Biology and*
633 *Evolution*, 39(8):msac174.
- 634 Lartillot, N. (2013). Phylogenetic Patterns of GC-Biased Gene Conversion in Placental Mammals and the
635 Evolutionary Dynamics of Recombination Landscapes. *Molecular Biology and Evolution*, 30(3):489–502.
- 636 Latrille, T., Duret, L., and Lartillot, N. (2017). The Red Queen model of recombination hot-spot evolu-
637 tion: a theoretical investigation. *Philosophical Transactions of the Royal Society B: Biological Sciences*,
638 372(1736):20160463.
- 639 Lesecque, Y., Glémin, S., Lartillot, N., Mouchiroud, D., and Duret, L. (2014). The Red Queen Model of
640 Recombination Hotspots Evolution in the Light of Archaic and Modern Human Genomes. *PLOS Genetics*,
641 10(11):e1004790.
- 642 Li, R., Bitoun, E., Altemose, N., Davies, R. W., Davies, B., and Myers, S. R. (2019). A high-resolution map
643 of non-crossover events reveals impacts of genetic diversity on mammalian meiotic recombination. *Nat*
644 *Commun*, 10(1):3900.

- 645 Lichten, M. and Goldman, A. S. H. (1995). Meiotic recombination hotspots. *Annu. Rev. Genet.*, 29(1):423–
646 444.
- 647 Maloisel, L. and Rossignol, J.-L. (1998). Suppression of crossing-over by DNA methylation in *Ascobolus*.
648 *Genes & Development*, 12(9):1381–1389.
- 649 Marand, A. P., Jansky, S. H., Zhao, H., Leisner, C. P., Zhu, X., Zeng, Z., Crisovan, E., Newton, L.,
650 Hamernik, A. J., Veilleux, R. E., Buell, C. R., and Jiang, J. (2017). Meiotic crossovers are associated with
651 open chromatin and enriched with Stowaway transposons in potato. *Genome Biol*, 18(1):1–16.
- 652 Mihola, O., Landa, V., Pratto, F., Brick, K., Kobets, T., Kusari, F., Gasic, S., Smagulova, F., Grey, C.,
653 Flachs, P., Gergelits, V., Tresnak, K., Silhavy, J., Mlejnek, P., Camerini-Otero, R. D., Pravenec, M.,
654 Petukhova, G. V., and Trachtulec, Z. (2021). Rat PRDM9 shapes recombination landscapes, duration of
655 meiosis, gametogenesis, and age of fertility. *BMC Biol*, 19(1):1–20.
- 656 Mihola, O., Pratto, F., Brick, K., Linhartova, E., Kobets, T., Flachs, P., Baker, C. L., Sedlacek, R., Paigen,
657 K., Petkov, P. M., Camerini-Otero, R. D., and Trachtulec, Z. (2019). Histone methyltransferase PRDM9
658 is not essential for meiosis in male mice. *Genome Res.*, 29(7):1078–1086.
- 659 Milholland, B., Dong, X., Zhang, L., Hao, X., Suh, Y., and Vijg, J. (2017). Differences between germline
660 and somatic mutation rates in humans and mice. *Nat Commun*, 8(1):15183.
- 661 Munch, K., Mailund, T., Dutheil, J. Y., and Schierup, M. H. (2014). A fine-scale recombination map of the
662 human–chimpanzee ancestor reveals faster change in humans than in chimpanzees and a strong impact of
663 GC-biased gene conversion. *Genome Res.*, 24(3):467–474.
- 664 Myers, S., Bowden, R., Tumian, A., Bontrop, R. E., Freeman, C., MacFie, T. S., McVean, G., and Donnelly,
665 P. (2010). Drive Against Hotspot Motifs in Primates Implicates the PRDM9 Gene in Meiotic Recombi-
666 nation. *Science*, 327(5967):876–879.
- 667 Nagylaki, T. (1983). Evolution of a finite population under gene conversion. *Proceedings of the National*
668 *Academy of Sciences*, 80(20):6278–6281.
- 669 Oliver, P. L., Goodstadt, L., Bayes, J. J., Birtle, Z., Roach, K. C., Phadnis, N., Beatson, S. A., Lunter,
670 G., Malik, H. S., and Ponting, C. P. (2009). Accelerated Evolution of the Prdm9 Speciation Gene across
671 Diverse Metazoan Taxa. *PLOS Genetics*, 5(12):e1000753.
- 672 Page, S. L. and Hawley, R. S. (2003). Chromosome Choreography: The Meiotic Ballet. *Science*,
673 301(5634):785–789.
- 674 Parvanov, E. D., Petkov, P. M., and Paigen, K. (2010). Prdm9 Controls Activation of Mammalian Recom-
675 bination Hotspots. *Science*, 327(5967):835–835.
- 676 Pessia, E., Popa, A., Mousset, S., Rezvoy, C., Duret, L., and Marais, G. A. (2012). Evidence for widespread
677 GC-biased gene conversion in eukaryotes. *Genome biology and evolution*, 4(7):675–682.
- 678 Petes, T. D. (2001). Meiotic recombination hot spots and cold spots. *Nat Rev Genet*, 2(5):360–369.
- 679 Pratto, F., Brick, K., Khil, P., Smagulova, F., Petukhova, G. V., and Camerini-Otero, R. D. (2014). Recom-
680 bination initiation maps of individual human genomes. *Science*, 346(6211):1256442–1256442.

- 681 Qu, J., Hodges, E., Molaro, A., Gagneux, P., Dean, M. D., Hannon, G. J., and Smith, A. D. (2018).
682 Evolutionary expansion of DNA hypomethylation in the mammalian germline genome. *Genome Res.*,
683 28(2):145–158.
- 684 Quinlan, A. R. and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features.
685 *Bioinformatics*, 26(6):841–842.
- 686 Raynaud, M., Gagnaire, P.-A., and Galtier, N. (2023). Performance and limitations of linkage-disequilibrium-
687 based methods for inferring the genomic landscape of recombination and detecting hotspots: a simulation
688 study. *Peer Community Journal*, 3.
- 689 Samuk, K. and Noor, M. A. F. (2022). Gene flow biases population genetic inference of recombination rate.
690 *G3 Genes|Genomes|Genetics*, 12(11):jkac236.
- 691 Schield, D. R., Pasquesi, G. I. M., Perry, B. W., Adams, R. H., Nikolakis, Z. L., Westfall, A. K., Orton,
692 R. W., Meik, J. M., Mackessy, S. P., and Castoe, T. A. (2020). Snake Recombination Landscapes Are
693 Concentrated in Functional Regions despite PRDM9. *Molecular Biology and Evolution*, 37(5):1272–1294.
- 694 Singhal, S., Leffler, E. M., Sannareddy, K., Turner, I., Venn, O., Hooper, D. M., Strand, A. I., Li, Q., Raney,
695 B., Balakrishnan, C. N., Griffith, S. C., McVean, G., and Przeworski, M. (2015). Stable recombination
696 hotspots in birds. *Science*, page 6.
- 697 Smagulova, F., Brick, K., Pu, Y., Camerini-Otero, R. D., and Petukhova, G. V. (2016). The evolutionary
698 turnover of recombination hot spots contributes to speciation in mice. *Genes Dev.*, 30(3):266–280.
- 699 Smith, T. C. A., Arndt, P. F., and Eyre-Walker, A. (2018). Large scale variation in the rate of germ-line de
700 novo mutation, base composition, divergence and diversity in humans. *PLoS Genet*, 14(3):e1007254.
- 701 Spence, J. P. and Song, Y. S. (2019). Inference and analysis of population-specific fine-scale recombination
702 maps across 26 diverse human populations. *Science Advances*, 5(10):eaaw9206.
- 703 Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W., and Smadja, C. M. (2017). Variation in
704 recombination frequency and distribution across eukaryotes: patterns and processes. *Phil. Trans. R. Soc.*
705 *B*, 372(1736):20160455.
- 706 Tock, A. J. and Henderson, I. R. (2018). Hotspots for Initiation of Meiotic Recombination. *Front. Genet.*,
707 9:521.
- 708 Wallberg, A., Glémin, S., and Webster, M. T. (2015). Extreme Recombination Frequencies Shape Genome
709 Variation and Evolution in the Honeybee, *Apis mellifera*. *PLoS Genet*, 11(4):e1005189.
- 710 Wang, R. J., Thomas, G. W. C., Raveendran, M., Harris, R. A., Doddapaneni, H., Muzny, D. M., Capitano,
711 J. P., Radivojac, P., Rogers, J., and Hahn, M. W. (2020). Paternal age in rhesus macaques is positively
712 associated with germline mutation accumulation but not with measures of offspring sociability. *Genome*
713 *Res.*, 30(6):826–834.
- 714 Williams, A. L., Genovese, G., Dyer, T., Altemose, N., Truax, K., Jun, G., Patterson, N., Myers, S. R.,
715 Curran, J. E., Duggirala, R., Blangero, J., Reich, D., and Przeworski, M. (2015). Non-crossover gene
716 conversions show strong GC bias and unexpected clustering in humans. *eLife*, 4:e04637.

- 717 Wong, W. S. W., Solomon, B. D., Bodian, D. L., Kothiyal, P., Eley, G., Huddleston, K. C., Baker, R.,
718 Thach, D. C., Iyer, R. K., Vockley, J. G., and Niederhuber, J. E. (2016). New observations on maternal
719 age effect on germline de novo mutations. *Nat Commun*, 7(1):10486.
- 720 Zelkowski, M., Olson, M. A., Wang, M., and Pawlowski, W. (2019). Diversity and Determinants of Meiotic
721 Recombination Landscapes. *Trends in Genetics*, 35(5):359–370.
- 722 Úbeda, F. and Wilkins, J. F. (2011). The Red Queen theory of recombination hotspots. *Journal of Evolu-*
723 *tionary Biology*, 24(3):541–553.