

PRDM9 and the Evolution of Recombination Hotspots

Francisco Úbeda^{a,*}, Timothy W. Russell^a, Vincent A.A. Jansen^a

^a*School of Biological Sciences, Royal Holloway University of London, Egham, Surrey, TW20 0EX, UK*

Abstract

Recombination in mammals is not uniformly distributed along the chromosome but concentrated in small regions known as recombination hotspots. Recombination starts with the double-strand break of a chromosomal sequence and results in the transmission of the sequence that does not break (preventing recombination) more often than the sequence that breaks (allowing recombination). Thus recombination itself renders individual recombination hotspots inactive and over time should drive them to extinction in the genome. Empirical evidence shows that individual recombination hotspots die but, far from being driven to extinction, they are abundant in the genome: a contradiction referred to as the Recombination Hotspot Paradox. What saves recombination hotspots from extinction? The current answer relies in the formation of new recombination hotspots in new genomic sites driven by viability selection in favour of recombination. Here we formulate a population genetics model that incorporates the molecular mechanism initiating recombination in mammals (PRDM9-like genes), to provide an alternative solution to the paradox. We find that weak selection allows individual recombination hotspots to become inactive (die) while saving them from extinction in the genome by driving their re-activation (resurrection). Our model shows that when selection for recombination is weak, the introduction of rare variants causes recombination sites to oscillate between hot and cold phenotypes with a recombination hotspot dying only to come back. Counter-intuitively, we find that low viability selection leaves a hard selective sweep signature in the genome, with the selective sweep at the recombination hotspot being the hardest when fertility selection is the lowest. Our model can help to understand the rapid evolution of PRDM9, the co-existence of two types of hotspots, the life expectancy of hotspots, and the volatility of the recombinational landscape (with hotspots rarely being shared between closely related species).

Keywords:

Recombination Hotspot Paradox, PRDM9, Population genetics, Gene conversion, viability selection, Heteroclinic cycles

*Corresponding author

Email address: `f.ubeda@rhul.ac.uk` (Francisco Úbeda)

1. Introduction

1 The distribution of recombination in the genome - and thus crossover events - is key
2 to our understanding of the molecular mechanisms controlling recombination, the role
3 of recombination on evolution, and the implementation of tests linking genetic markers
4 with human disease (genome-wide association studies) (Boulton et al., 1997; Hey, 2004;
5 Rosenberg et al., 2010). In many mammals, recombination is not uniformly distributed in
6 the genome but concentrated in small chromosomal regions —known as recombination
7 hotspots— where recombination is ten to a thousand times more frequent than the
8 genome’s average (Lichten and Goldman, 1995; Petes, 2001; Myers et al., 2005; Paigen
9 and Petkov, 2010). While recombination hotspots are abundant in the mammalian
10 genome (for example, in the human genome there are more than twenty five thousand),
11 their mere existence is paradoxical and their life cycle is not fully understood (Boulton
12 et al., 1997; Pineda-Krch and Redfield, 2005; Myers et al., 2005).

13 Recombination is initiated by a double-strand break (DSB) and may result in the
14 conversion of the allelic sequence that breaks (active allele, enabling recombination)
15 into the allelic sequence that does not break (inactive allele, disabling recombination)
16 (Lichten and Goldman, 1995; Petes, 2001). The conversion of the allele that enables
17 recombination into the one that disables recombination should be faster in genomic re-
18 gions where recombination is higher (recombination hotspots). As a result individual
19 recombination hotspots should become inactive (this process is often referred as the
20 death of a hotspot; Coop and Myers (2007)) and, over evolutionary time, recombina-
21 tion hotspots should disappear from the genome (Boulton et al., 1997; Pineda-Krch and
22 Redfield, 2005). Empirical work shows that individual recombination hotspots die (Ptak
23 et al., 2004, 2005; Winckler et al., 2005; Coop et al., 2008; Myers et al., 2010; Stevison
24 et al., 2015) but, despite their self-destructive nature, recombination hotspots are abun-
25 dant in the mammalian genome (Myers et al., 2005; Baudat et al., 2013), thus posing
26 the Recombination Hotspot Paradox (Boulton et al., 1997; Pineda-Krch and Redfield,
27 2005): what saves recombination hotspots from extinction?

28 Due to its molecular, evolutionary and medical implications the Recombination
29 Hotspot Paradox has received much attention. Initial work aimed to test whether the
30 known beneficial effects of recombination —in particular how recombination may favor
31 proper chromosomal segregation during meiosis; thus avoiding the formation of aneu-
32 ploidy gametes (Hassold et al., 2000; Louis and Borts, 2003; Brick et al., 2012; Alves et al.,
33 2017)— can solve the paradox (Boulton et al., 1997; Pineda-Krch and Redfield, 2005;
34 Calabrese, 2007; Peters, 2008). These mathematical models found that the strength of
35 viability selection needed to maintain active alleles at recombination hotspots over evolu-
36 tionary time was too high to be realistic (Boulton et al., 1997; Pineda-Krch and Redfield,
37 2005; Calabrese, 2007; Peters, 2008). Furthermore, in these models when viability se-
38 lection prevents the extinction of hotspots in the genome, it does so by preventing the
39 death of individual hotspots, which is contrary to empirical observations (Ptak et al.,
40 2004, 2005; Winckler et al., 2005; Coop et al., 2008; Myers et al., 2010; Stevison et al.,
41 2015). Therefore, far from providing solutions to the Recombination Hotspot Paradox,
42 previous work demonstrates that the paradox is well grounded.

43 Recent advances in our understanding of the molecular mechanisms initiating re-
44 combination include the identification of gene PRDM9 in humans (and many mammals)
45 coding for protein PRDM9 that may bind a specific sequence at a target recombination
46 hotspot (Myers et al., 2010; Baudat et al., 2010). Binding specificity between PRDM9
47 and its target site is required for the initiation of recombination (Myers et al., 2010;
48 Baudat et al., 2010). This finding led to the verbal argument that when a target site has
49 its binding motif (active allele) replaced by the non-binding motif (inactive allele) due
50 to biased gene conversion, a mutant PRDM9 could create a new target site by coding
51 for a new binding motif (Myers et al., 2010; Baudat et al., 2010). Natural selection
52 would thus favor this rare mutant PRDM9 as long as recombination is advantageous for
53 the individual (Myers et al., 2010; Baudat et al., 2010). Lacking a mathematical model
54 to back this claim, it remained unclear whether selection would favor such mutant to
55 the extent of allowing the formation (henceforth birth) of new recombination hotspots
56 before an inactive allele arose. Furthermore, would the strength of selection required for
57 the birth of new hotspots be too high to be realistic?

58 Úbeda and Wilkins (2011) modeled a trans acting modifier locus with binding speci-
59 ficity —like PRDM9— showing that, for a strength of selection lower than in previous
60 models, new recombination hotspots can be born at new target sites, while existing re-
61 combination hotspots die (Úbeda and Wilkins, 2011). These findings were consistent
62 with empirical observations regarding the persistence of recombination hotspots in the
63 genome in spite of the death of individual recombination hotspots (Úbeda and Wilkins,
64 2011). The Red Queen hypothesis of recombination hotspots evolution refers to the
65 balance between death and birth of new hotspots driven by conversion and viability
66 selection (Úbeda and Wilkins, 2011), and is the prevailing explanation to the recombi-
67 nation hotspots paradox (Lesecque et al., 2014; Latrille et al., 2017).

68 In many respects, however, the Red Queen hypothesis needs further theoretical in-
69 vestigation (Latrille et al., 2017). One of these key theoretical aspects is the role of
70 viability selection in maintaining recombination hotspots, and the evolution of PRDM9
71 and target sequences (Ségurel et al., 2011; Latrille et al., 2017). Recent models include
72 variables that mask the effect of selection; for example drift, recurrent mutation, and
73 multiple locus targets (Úbeda and Wilkins, 2011; Latrille et al., 2017). While the intro-
74 duction of these variables is justified to make the models more realistic, they complicate
75 our understanding of the interplay between the key variables of these models, namely
76 conversion and selection.

77 Here we formulate a population genetics model aimed to explore the interplay be-
78 tween conversion and selection in the resolution of the Recombination Hotspot Paradox.
79 We start by considering an infinite population, without recurrent mutation and with a
80 single target locus, to eliminate the above mentioned confounding variables. We build on
81 the insight gained from this minimal model to interpret the results of an extended model
82 with a finite population and recurrent mutation. In doing so, we find an alternative solu-
83 tion to the Recombination Hotspot Paradox, one that does not require the formation of
84 new hotspots but relies on existing hotspots. Counter-intuitively, in our novel solution, it
85 is low viability selection regimes that allow the persistence of recombination hotspots in

86 spite of the death of individual ones (contrary to previous models) (Latrille et al., 2017).
87 Furthermore, sometimes, low viability selection accelerates the turnover of hotspots. We
88 also find that viability selection can maintain polymorphisms at the PRDM9 and target
89 loci. We apply these findings to explore the molecular signatures of selection in PRDM9
90 and target loci and consider their implications for genome-wide association studies.

91 **2. Methods**

92 *2.1. Two-locus n-alleles model*

93 We follow the classic Wright-Fisher population genetics framework Wright (1969);
94 Bürger (2000) to formulate a discrete time mathematical model of an infinite population
95 of diploid individuals carrying two loci with an arbitrary number of alleles in each locus.

96 This model represents the interaction between a gene (PRDM9-like) producing a
97 protein that binds a specific motif at a target recombination site (Figure 1), as it is
98 observed in humans and many mammals (Myers et al., 2010; Baudat et al., 2010, 2013).
99 The modifier locus A may carry alleles A_1, A_2, \dots, A_I each encoding a protein that at-
100 tempts to bind a motif at a target locus B . Locus B may carry alleles B_1, B_2, \dots, B_K
101 each corresponding to a base pair motif that the protein produced by locus A may at-
102 tempt to bind. In each generation, both modifier alleles in each diploid individual show
103 the same level of expression producing proteins that have equal probability of binding
104 the two target motifs (Figure 1). Therefore, in an individual with genotype $\frac{A_i B_k}{A_j B_l}$, the
105 probability that a protein produced by alleles A_i or A_j attempts to bind the motif of
106 alleles B_k or B_l is $\frac{1}{4}$ (Figure 1). The binding attempt of the protein A_i to the motif B_k
107 results in binding and a double-strand break of allele B_k with probability $b_{i,k}$. However,
108 the binding attempt may result in failure to bind and lack of any double-strand break
109 with probability $1 - b_{i,k}$ (where $0 < b_{i,k} < 1$) (Figure 1).

110 A double-strand break initiates recombination and the chromatid that breaks is often
111 repaired using its homologous chromatid as a template (Lichten and Goldman, 1995;
112 Petes, 2001) (Figure 1). During the repair process there might be a crossover event in or
113 near the target locus with probability r and none with probability $1 - r$ (where $0 < r < 1$)
114 (Lichten and Goldman, 1995; Petes, 2001) (Figure 1). In our model, we assume that
115 a crossover event between the modifier and target loci requires a double-strand break
116 at the target locus. However, if the modifier and target loci are far apart in the same
117 chromosome or in separate chromosomes, a crossover event between these loci may not
118 require a double-strand break. Whether a crossover event between the modifier and
119 target loci require a double-strand break at the target locus or not does not change any
120 of the qualitative results of our model (see the Supplemental Material for a formulation
121 of this model and Figure 2 for a summary of the results). During the repair process
122 there might also be conversion of the allelic motif that breaks into the allelic motif that
123 does not break with probability c and restoration to the allelic motif that breaks with
124 probability $1 - c$ (where $0 < c < 1$) (Szostak et al., 1983; Sun et al., 1991; Lichten and
125 Goldman, 1995; Petes, 2001) (Figure 1). Typically c takes the value $\frac{1}{2}$ (Szostak et al.,
126 1983; Sun et al., 1991; Lichten and Goldman, 1995; Petes, 2001). Notice that biased

127 gene conversion results in the over-transmission of the allele that is less likely to break
 128 (Boulton et al., 1997; Petes, 2001) (Figure 1).

129 Recombination ends up with Mendelian segregation of alleles into gametes. Following
 130 previous models (Boulton et al., 1997; Pineda-Krch and Redfield, 2005; Peters, 2008;
 131 Úbeda and Wilkins, 2011; Latrille et al., 2017), we assume that individuals undergoing
 132 recombination at the target locus have proper chromosomal segregation and do not
 133 suffer any fitness cost, while individuals that do not undergo recombination at the target
 134 locus have defective chromosomal segregation producing aneuploid (non-viable) gametes
 135 with probability f (where $0 < f < 1$) (Figure 1). Therefore, the fitness of individuals
 136 experiencing a recombination event is 1 but the fitness of individuals not experiencing
 137 a recombination event is $1 - f$ (Figure 1). Proper chromosomal segregation, however,
 138 often requires a crossover event rather than a recombination event (Baker et al., 1976;
 139 Koehler et al., 1996; Hassold and Hunt, 2001; Louis and Borts, 2003; Brick et al., 2012;
 140 Alves et al., 2017). Whether it is a crossover or a recombination event that determine
 141 the probability of proper chromosomal segregation does not change any of the qualitative
 142 results of our model (see the Supplemental Material for a formulation of this model and
 143 Figure 2 for a summary of the results).

144 Let $x_{i,k}$ be the frequency of haplotype A_iB_k in gametes. Notice that $0 \leq x_{i,k} \leq 1$ and
 145 $\sum_{i,k} x_{i,k} = 1$. Random union of gametes results in an embryo with genotype $\frac{A_iB_k}{A_jB_l}$ with
 146 frequency $x_{i,k}x_{j,l}$. The probability that this embryo reaches adulthood is independent of
 147 its genotype, but its genotype determines the outcome of meiosis in adults. In particular,
 148 the probability that during meiosis the protein produced by the modifier locus breaks
 149 targets B_k and B_l are $\bar{b}_{i,j,k} = \frac{1}{2}(b_{i,k} + b_{j,k})$ and $\bar{b}_{i,j,l} = \frac{1}{2}(b_{i,l} + b_{j,l})$ respectively, and the
 150 probability that it breaks one of the targets is $\bar{\bar{b}}_{i,j,kl} = \frac{1}{2}(\bar{b}_{i,j,k} + \bar{b}_{i,j,l})$. The probability that
 151 during meiosis a double-strand break is followed by a crossover event between alleles at
 152 locus A and B is r , and the probability that the motif that breaks is converted into the
 153 motif that does not break is c . Recombination at the target locus is followed by correct
 154 Mendelian segregation of haplotypes into gametes but in the absence of recombination
 155 segregation of haplotypes is incorrect with probability f . Haplotype segregation brings
 156 us back to the beginning of our census.

157 The frequency of haplotype A_iB_k in gametes in the next generation is:

$$\begin{aligned}
 x'_{i,k} = \frac{1}{\bar{w}} \sum_{j,l} \frac{1}{2} [& (\bar{\bar{b}}_{i,j,kl} + (1 - \bar{\bar{b}}_{i,j,kl})(1 - f)) x_{i,k} x_{j,l} \\
 & - \frac{1}{4} c (\bar{b}_{i,j,k} x_{i,k} x_{j,l} - \bar{b}_{i,j,l} x_{i,l} x_{j,k}) \\
 & - \frac{1}{2} (1 - c) r \bar{\bar{b}}_{i,j,kl} (x_{i,k} x_{j,l} - x_{i,l} x_{j,k})]
 \end{aligned} \tag{1}$$

158 where prime represents the next generation and:

$$\bar{w} = \sum_{i,k} \sum_{j,l} \frac{1}{2} [\bar{\bar{b}}_{i,j,kl} + (1 - \bar{\bar{b}}_{i,j,kl})(1 - f)] x_{i,k} x_{j,l} \tag{2}$$

159 is the population mean fitness. These changes in haplotype frequency underpin changes
 160 in the population mean crossover rate at the target locus:

$$\bar{r} = \frac{1}{2}r \sum_{i,k} \sum_{j,l} \bar{b}_{i,j,kl} x_{i,k} x_{j,l} \quad (3)$$

161 which is the phenotype whose evolution we are interested in.

162 Our model greatly differs from all other attempts to incorporate binding specificity
 163 (PRDM9-like genes) into the mechanism of recombination hotspots (Úbeda and Wilkins,
 164 2011; Latrille et al., 2017), as previous models relied on simulations while we present
 165 analytic results (although see Latrille et al. (2017) for a one locus model approximating
 166 the frequency of PRDM9-like alleles in an infinite population).

167 2.2. Two-locus two-allele model

168 We consider the above model in the particular case when there are two alleles (A_1, A_2)
 169 at the modifier locus and two alleles (B_1, B_2) at the only target locus, resulting in four
 170 different haplotypes ($A_1B_1, A_1B_2, A_2B_1, A_2B_2$). Henceforth, we assume that a match
 171 between the subscripts of the modifier allele producing the binding protein and the
 172 allelic sequence that is the target of this protein results in a double-strand break with
 173 probability b (where $0 < b < 1$) and a mismatch between the subscripts prevents a
 174 double-strand break. For our modelling purposes this translates into:

$$b_{i,k} = \begin{cases} b & \text{if } i = k \\ 0 & \text{if } i \neq k. \end{cases}$$

175 Notice that two of these haplotypes (A_1B_1, A_2B_2) correspond to haplotypes produc-
 176 ing a protein that matches its own recognition sequence (recombination enabling hap-
 177 lotypes) and the other two (A_1B_2, A_2B_1) correspond to haplotypes producing a protein
 178 that does not match its own recognition sequence (recombination disabling haplotypes).

179 The dynamic system describing the change in frequency over time of each of these
 180 haplotypes can be obtained from replacing generic subscripts i and k by specific sub-
 181 scripts 1 and 2 in equation (1). The frequency of haplotype A_iB_k in gametes in the next
 182 generation is:

$$\begin{aligned} \bar{w}x'_{1,1} &= (\frac{1}{4}b + \frac{1}{2}(1 - \frac{1}{2}b)(1 - f) + \frac{1}{4}bfx_{1,1} - \frac{1}{8}bcx_{1,2})x_{1,1} - \frac{1}{8}b(\frac{1}{2}c + (1 - c)r)D \\ \bar{w}x'_{1,2} &= (\frac{1}{4}b + \frac{1}{2}(1 - \frac{1}{2}b)(1 - f) - \frac{1}{4}bfx_{1,2} + \frac{1}{8}bcx_{1,1})x_{1,2} + \frac{1}{8}b(\frac{1}{2}c + (1 - c)r)D \\ \bar{w}x'_{2,1} &= (\frac{1}{4}b + \frac{1}{2}(1 - \frac{1}{2}b)(1 - f) - \frac{1}{4}bfx_{2,1} + \frac{1}{8}bcx_{2,2})x_{2,1} + \frac{1}{8}b(\frac{1}{2}c + (1 - c)r)D \\ \bar{w}x'_{2,2} &= (\frac{1}{4}b + \frac{1}{2}(1 - \frac{1}{2}b)(1 - f) + \frac{1}{4}bfx_{2,2} - \frac{1}{8}bcx_{2,1})x_{2,2} - \frac{1}{8}b(\frac{1}{2}c + (1 - c)r)D \end{aligned} \quad (4)$$

183 where

$$\bar{w} = \frac{1}{4}b + \frac{1}{2}(1 - \frac{1}{2}b)(1 - f) + \frac{1}{4}bf(x_{1,1}^2 + x_{2,2}^2 - x_{1,2}^2 - x_{2,1}^2) \quad (5)$$

184 is the population mean fitness and:

$$D = x_{1,1}x_{2,2} - x_{1,2}x_{2,1} \quad (6)$$

185 is the linkage disequilibrium.

186 To simplify the analysis, we define parameters $\alpha, \beta, \gamma,$ and δ as follows:

$$\bar{w}x'_{i,k} = \underbrace{\left(\frac{1}{4}b + \frac{1}{2}(1 - \frac{1}{2}b)(1 - f)\right)}_{\alpha} \pm \underbrace{\frac{1}{4}bf}_{\beta} x_{i,k} \pm \underbrace{\frac{1}{8}bc}_{\gamma} x_{i,l} x_{i,k} \pm \underbrace{\frac{1}{8}b(\frac{1}{2}c + (1 - c)r)}_{\delta} D \quad (7)$$

187 which allows us to re-write the system of equations (4) as follows:

$$\begin{aligned} \bar{w}x'_{1,1} &= (\alpha + \beta x_{1,1} - \gamma x_{1,2})x_{1,1} - \delta D \\ \bar{w}x'_{1,2} &= (\alpha - \beta x_{1,2} + \gamma x_{1,1})x_{1,2} + \delta D \\ \bar{w}x'_{2,1} &= (\alpha - \beta x_{2,1} + \gamma x_{2,2})x_{2,1} + \delta D \\ \bar{w}x'_{2,2} &= (\alpha + \beta x_{2,2} - \gamma x_{2,1})x_{2,2} - \delta D \end{aligned} \quad (8)$$

188 with population mean fitness:

$$\bar{w} = \alpha + \beta(x_{1,1}^2 - x_{1,2}^2 - x_{2,1}^2 + x_{2,2}^2). \quad (9)$$

189 Notice that $0 < \alpha, \beta, \gamma, \delta < 1$. This two-locus two-allele model shares some similarities
 190 with the well-known symmetric viability model of Karlin and Feldman (Karlin et al.,
 191 1970; Bürger, 2000), albeit our model is not symmetrical and therefore the results of the
 192 symmetric viability model do not carry over.

193 3. Results

194 3.1. Equilibria

195 We apply the equilibrium conditions ($x'_{i,k} = x_{i,k} = x_{i,k}^*$ for all i, k) to system (8) to
 196 find five equilibria with biological meaning; where all haplotype frequencies lie between
 197 (and including) 0 and 1. Let $\mathbf{x}^{*e} = (x_{1,1}^{*e}, x_{1,2}^{*e}, x_{2,1}^{*e}, x_{2,2}^{*e})$ denote the haplotype frequencies
 198 at equilibrium e where e is between one and five.

199 The first four equilibria correspond to the corners of the three dimensional simplex:

$$\begin{aligned}
\mathbf{x}^{*1} &= (1, 0, 0, 0) \\
\mathbf{x}^{*2} &= (0, 1, 0, 0) \\
\mathbf{x}^{*3} &= (0, 0, 1, 0) \\
\mathbf{x}^{*4} &= (0, 0, 0, 1).
\end{aligned} \tag{10}$$

200 Notice that equilibria 1 and 4, \mathbf{x}^{*1} and \mathbf{x}^{*4} , correspond to the fixation of one of the two
201 recombination enabling haplotypes, $x_{1,1}$ and $x_{2,2}$ respectively. Equilibria 2 and 3, \mathbf{x}^{*2}
202 and \mathbf{x}^{*3} , correspond to the fixation of one of the two recombination disabling haplotypes,
203 $x_{1,2}$ and $x_{2,1}$ respectively (Figure 2).

204 The last equilibrium can be obtained by noticing some symmetries of our model. In
205 particular, if at any point $x_{1,1} = x_{2,2}$ and $x_{1,2} = x_{2,1}$, this remains so in the future.
206 To see this, notice that if $x_{1,1} = x_{2,2}$ and $x_{1,2} = x_{2,1}$, the difference equations become
207 $x'_{1,1} = x'_{2,2}$ and $x'_{1,2} = x'_{2,1}$ and the changes in $x_{1,1}$ and $x_{1,2}$ are equal to the changes in
208 $x_{2,2}$ and $x_{2,1}$ respectively. Also note that if $x_{1,1} = x_{2,2}$ and $x_{1,2} = x_{2,1}$ and keeping in
209 mind that $x_{1,1} + x_{1,2} + x_{2,1} + x_{2,2} = 1$, we also have that $2x_{1,1} + 2x_{1,2} = 1$ and thus
210 $x_{1,2} = \frac{1}{2} - x_{1,1}$.

211 The existence of a one dimensional manifold which is invariant in the interior of
212 the state space implies that there is a symmetric equilibrium. The dynamics on this
213 manifold are described by a single difference equation:

$$\bar{w}x'_{1,1} = (\alpha + \beta x_{1,1} - \gamma(\frac{1}{2} - x_{1,1}))x_{1,1} - \delta(x_{1,1} - \frac{1}{4}) \tag{11}$$

214 with population mean fitness:

$$\bar{w} = \alpha + 2\beta(x_{1,1} - \frac{1}{4}). \tag{12}$$

215 Applying the equilibrium condition ($x'_{1,1} = x_{1,1} = x_{1,1}^*$) to the previous equation
216 yields the symmetric equilibrium:

$$\begin{aligned}
\mathbf{x}^{*5} &= (x_{1,1}^{*5}, \frac{1}{2} - x_{1,1}^{*5}, \frac{1}{2} - x_{1,1}^{*5}, x_{1,1}^{*5}) \\
x_{1,1}^{*5} &= \frac{1}{4} + \frac{1}{4} \frac{2\delta - \sqrt{(2\delta)^2 + (\gamma - \beta)^2}}{\gamma - \beta}.
\end{aligned} \tag{13}$$

217 At this equilibrium, the linkage disequilibrium is:

$$D^* = x_{1,1}^{*5} - \frac{1}{4} = \frac{1}{4} \frac{2\delta - \sqrt{(2\delta)^2 + (\gamma - \beta)^2}}{\gamma - \beta}, \tag{14}$$

218 and the population mean fitness is:

$$\bar{w}^* = \alpha + 2\beta(x_{1,1}^* - \frac{1}{4}) = \alpha + 2\beta D^*. \quad (15)$$

219 Notice that equilibrium \mathbf{x}^{*5} corresponds to a polymorphism where all haplotypes (re-
220 combination enablers and disablers) are preserved.

221 Finally, we can re-write the expression for equilibrium \mathbf{x}^{*5} in terms of the original
222 parameters of our model:

$$x_{1,1}^{*5} = \frac{1}{4} + \frac{1}{4} \frac{\frac{1}{2}c+(1-c)r - \sqrt{(\frac{1}{2}c+(1-c)r)^2 + (\frac{1}{2}c-f)^2}}{\frac{1}{2}c-f} \quad (16)$$

223 3.2. Stability

224 The stability of an equilibrium \mathbf{x}^{*e} of a map $\mathbf{x}' = \mathbf{g}(\mathbf{x})$ is determined by studying the
225 eigenvalues λ^e of the Jacobian matrix \mathbf{J} of the map evaluated at the equilibrium, that is
226 $\mathbf{J}|_{\mathbf{x}=\mathbf{x}^{*e}}$. For brevity, we will refer to the eigenvalues λ_i^e as the eigenvalues of equilibrium
227 \mathbf{x}^{*e} . If the modulus of all eigenvalues of equilibrium \mathbf{x}^{*e} are less than one ($|\lambda_i^e| < 1$ for
228 all $i = 1, \dots, n$), the equilibrium is linearly stable (where $|z|$ denotes the modulus of a
229 number z that may have real $\text{Re}(z)$ and imaginary $\text{Im}(z)$ components and is defined as
230 $|z| = \sqrt{\text{Re}(z)^2 + \text{Im}(z)^2}$). If the modulus of at least one eigenvalue of equilibrium \mathbf{x}^{*e} is
231 greater than one ($|\lambda_i^e| > 1$ for any $i = 1, \dots, n$), the equilibrium is linearly unstable.

The specifics of our model simplify the calculation of the Jacobian at equilibrium. In particular, our model describes changes in haplotype frequencies. To ensure that all frequencies add up to one at all times, the changes in frequency are normalized and the system is of the form:

$$\mathbf{x}' = \frac{\mathbf{g}(\mathbf{x})}{\bar{w}(\mathbf{x})} \quad (17)$$

where $\bar{w}(\mathbf{x}) = \mathbf{1}^T \mathbf{g}(\mathbf{x})$, $\mathbf{1}$ is a vector with all entries equal to one, and subscript T is the transpose operator. The Jacobian of this system is:

$$\mathbf{J} = D_x \frac{\mathbf{g}(\mathbf{x})}{\bar{w}(\mathbf{x})} = \frac{D_x \mathbf{g}(\mathbf{x})}{\bar{w}(\mathbf{x})} - \frac{\mathbf{g}(\mathbf{x}) \mathbf{1}^T D_x \mathbf{g}(\mathbf{x})}{\bar{w}(\mathbf{x})^2}, \quad (18)$$

where D_x is the total derivative with respect to x . Evaluated at equilibrium \mathbf{x}^* the Jacobian reduces to:

$$\mathbf{J}|_{\mathbf{x}=\mathbf{x}^*} = \frac{1}{\bar{w}(\mathbf{x}^*)} (\mathbf{I} - \mathbf{x}^* \mathbf{1}^T) D_x \mathbf{g}(\mathbf{x})|_{\mathbf{x}=\mathbf{x}^*}. \quad (19)$$

232 where \mathbf{I} is the identity matrix.

233 *3.2.1. Corner equilibria*

The eigenvalues of corner equilibria \mathbf{x}^{*1} and \mathbf{x}^{*4} are equal and given by:

$$\{\lambda_1^1, \lambda_2^1, \lambda_3^1, \lambda_4^1\} = \{\lambda_1^4, \lambda_2^4, \lambda_3^4, \lambda_4^4\} = \left(0, \frac{\alpha}{\alpha+\beta}, \frac{\alpha+\gamma}{\alpha+\beta}, \frac{\alpha-\delta}{\alpha+\beta}\right). \quad (20)$$

234 All eigenvalues of corner equilibrium \mathbf{x}^{*1} are real numbers, and \mathbf{x}^{*1} is stable if all
235 λ_{1-4}^1 lie between 1 and -1 .

- 236 1. Condition $-1 < \lambda_2^1 < 1$ is always satisfied.
- 237 2. Condition $-1 < \lambda_3^1 < 1$ implies the satisfaction of:
 - 238 i. $\lambda_3^1 < 1$ which requires that $\beta > \gamma$.
 - 239 ii. $\lambda_3^1 > -1$ which is always satisfied.
- 240 3. Condition $-1 < \lambda_4^1 < 1$ implies the satisfaction of:
 - 241 i. $\lambda_4^1 < 1$ which is always satisfied.
 - 242 ii. $\lambda_4^1 > -1$ which requires that $2\alpha + \beta - \delta > 0$ which is always satisfied for the
243 original parameters of our model.

244 To summarize, corner equilibria \mathbf{x}^{*1} and \mathbf{x}^{*4} are stable ($-1 < \lambda_{2-4}^1 < 1$) if $\beta > \gamma$
245 ($f > \frac{1}{2}c$ in terms of the original parameters) but unstable (saddles) ($-1 < \lambda_{2,4}^1 < 1$ but
246 $\lambda_3^1 > 1$) if $\beta < \gamma$ ($f < \frac{1}{2}c$) (see Table 1 and Figure 2).

The eigenvalues of corner equilibria \mathbf{x}^{*2} and \mathbf{x}^{*3} are equal and given by:

$$\{\lambda_1^2, \lambda_2^2, \lambda_3^2, \lambda_4^2\} = \{\lambda_1^3, \lambda_2^3, \lambda_3^3, \lambda_4^3\} = \left(0, \frac{\alpha}{\alpha-\beta}, \frac{\alpha-\gamma}{\alpha-\beta}, \frac{\alpha-\delta}{\alpha-\beta}\right). \quad (21)$$

247 All eigenvalues of corner equilibrium \mathbf{x}^{*2} are real numbers, and \mathbf{x}^{*2} is stable if all
248 λ_{1-4}^2 lie between 1 and -1 .

- 249 1. Condition $-1 < \lambda_2^2 < 1$ implies the satisfaction of:
 - 250 i. $\lambda_2^2 < 1$ which is never satisfied.
 - 251 ii. $\lambda_2^2 > -1$ which is always satisfied.
- 252 2. Condition $-1 < \lambda_3^2 < 1$ implies the satisfaction of:
 - 253 i. $\lambda_3^2 < 1$ which requires that $\beta < \gamma$.
 - 254 ii. $\lambda_3^2 > -1$ which is always satisfied for the original parameters of our model.

255 To summarise, corner equilibria \mathbf{x}^{*2} and \mathbf{x}^{*3} are unstable ($\lambda_2^2 > 1$). If $\beta < \gamma$ ($f < \frac{1}{2}c$)
256 these equilibria are saddles ($\lambda_2^2 > 1$ but $-1 < \lambda_3^2 < 1$) (see Table 1 and Figure 2).

257 *3.2.2. Heteroclinic orbit*

Here we show the existence of a heteroclinic orbit between the corner equilibria in our state space: $\dots\mathbf{x}^{*1} \rightarrow \mathbf{x}^{*2} \rightarrow \mathbf{x}^{*4} \rightarrow \mathbf{x}^{*3} \rightarrow \mathbf{x}^{*1}\dots$. To do so, we need to show that the subspaces in which the heteroclinic orbit travels are invariant. A set, $C \subseteq \mathbb{R}^n$, is an invariant set with respect to the map $x' = g(x)$ if, for every orbit ϕ it is true that $\phi_t(x) \in C \implies \phi_\tau(x) \in C$ for all $\tau > t$ where $t, \tau \in \mathbb{N}_+$. The subspaces in which our heteroclinic orbit travels are described by the lines joining each of the corners

of our simplex, namely: $(x_{1,1}, 1 - x_{1,1}, 0, 0)$, $(0, x_{1,2}, 0, 1 - x_{1,2})$, $(0, 0, 1 - x_{2,2}, x_{2,2})$, $(1 - x_{2,1}, 0, x_{2,1}, 0)$. Our system can be written in the form:

$$\bar{w}x'_{i,k} = (\alpha + \epsilon\beta x_{i,k} - \epsilon\gamma x_{i,l})x_{i,k} - \epsilon\delta(x_{i,k}x_{j,l} - x_{i,l}x_{j,k}), \quad (22)$$

258 where $\epsilon = 1$ for $(i, k) = (1, 1)$ and $(2, 2)$, and $\epsilon = -1$ for $(i, k) = (1, 2)$ and $(2, 1)$. From
 259 the system written in this form, it is easy to see that if $x_{i,k} = 0$ and $x_{i,l} = 0$ or $x_{j,k} = 0$
 260 for $(i, k, j, l) \in 1, 2$ then $x'_{i,k} = 0$. In particular for the heteroclinic orbit we consider,
 261 either when $x_{2,2} = 0$ then $x_{2,1} = 0$ and $x'_{2,2} = 0$, when $x_{2,1} = 0$ then $x_{1,1} = 0$ and
 262 $x'_{2,1} = 0$, when $x_{1,1} = 0$ then $x_{1,2} = 0$ and $x'_{1,1} = 0$, and when $x_{1,2} = 0$ then $x_{2,2} = 0$
 263 and $x'_{1,2} = 0$. This means that any subspace where $x_{i,k} = 0$ and $x_{i,l} = 0$ or $x_{j,k} = 0$ is
 264 invariant and thus all subspaces considered in our system are invariants.

265 When $\beta < \gamma$ ($f < \frac{1}{2}c$ in terms of the original parameters) all corner equilibria
 266 are saddles with one incoming and one outgoing eigenvector situated within the lines
 267 connecting the corner equilibria. Under the action of our system, the invariant subspaces
 268 have orbits which tend always away from one saddle equilibrium and towards another
 269 saddle equilibrium, thus implying the existence of a heteroclinic orbit. When $\beta < \gamma$,
 270 this heteroclinic orbit is stable (Russell et al., 2019).

271 3.2.3. Internal equilibrium

Calculating the eigenvalues of the internal equilibrium \mathbf{x}^{*5} using the original Jacobian matrix in (19) leads to intractable results. To attain eigenvalues that are tractable, we transform the vector \mathbf{x} into the vector \mathbf{y} using the linear transformation $\mathbf{y} = \mathbf{M}\mathbf{x}$ where:

$$\mathbf{M} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}. \quad (23)$$

The dynamics in the vicinity of the equilibrium for the transformed variables are:

$$\mathbf{y}' = \mathbf{M}\mathbf{x}' = \mathbf{M}\mathbf{J}\mathbf{x} = \mathbf{M}\mathbf{J}\mathbf{M}^{-1}\mathbf{y}, \quad (24)$$

where the matrix $\mathbf{M}\mathbf{J}\mathbf{M}^{-1}$ is given by:

$$\begin{aligned} \mathbf{M}\mathbf{J}\mathbf{M}^{-1} &= \mathbf{M}\frac{1}{\bar{w}^*}(\mathbf{I} - \mathbf{x}^*\mathbf{1}^T) D_x \mathbf{g}(\mathbf{x})|_{\mathbf{x}=\mathbf{x}^*} \mathbf{M}^{-1} \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ \frac{\beta - \gamma - 8D^*(\bar{w}^* + \delta)}{2\bar{w}^*} & -\frac{4D^*(\beta - \gamma) - 2(\bar{w}^* - \delta)}{2\bar{w}^*} & 0 & 0 \\ -\frac{\bar{w}^* + \frac{1}{2}\beta}{2\bar{w}^*} & 0 & 1 & \frac{\beta}{2\bar{w}^*} \\ -\frac{\frac{1}{2}(\beta - \gamma) + 2D^* + \bar{w}^*}{2\bar{w}^*} & 0 & \frac{\beta - \gamma}{2\bar{w}^*} & \frac{4D^*\gamma + 2\bar{w}^*}{2\bar{w}^*} \end{bmatrix}. \end{aligned} \quad (25)$$

The eigenvalues of the transformed matrix $\mathbf{M}\mathbf{J}\mathbf{M}^{-1}$ are equivalent to the eigenvalues of the original matrix \mathbf{J} but they are easier to find. In particular, the eigenvalues of matrix $\mathbf{M}\mathbf{J}\mathbf{M}^{-1}$ are:

$$\{\lambda_1^5, \lambda_2^5, \lambda_3^5, \lambda_4^5\} = \left(0, 1 + \frac{\gamma D^* + \sqrt{\Delta^*}}{\bar{w}^*}, 1 + \frac{\gamma D^* - \sqrt{\Delta^*}}{\bar{w}^*}, 1 - \frac{\delta + 2D^*(\beta - \gamma)}{\bar{w}^*}\right) \quad (26)$$

272 where $\Delta^* = (\gamma D^*)^2 + \frac{1}{4}\beta(\beta - \gamma)$.

273 The eigenvalues of internal equilibrium \mathbf{x}^{*5} can be either real or imaginary numbers.

274 1. Stability conditions derived from the second and third eigenvalues $\lambda_{2,3}^5$.

275 (a) Eigenvalues $\lambda_{2,3}^5$ are real numbers when $\Delta^* > 0$. If $\beta > \gamma$, the later condition
 276 is always satisfied, eigenvalues $\lambda_{2,3}^5$ are real numbers, and the stability of the
 277 internal equilibrium requires that $-1 < \lambda_{2,3}^5 < 1$. This requirement implies
 278 the satisfaction of four conditions:

- 279
- 280 i. Condition $\lambda_2^5 < 1$ requires that $\gamma D^* + \sqrt{\Delta^*} < 0$ which is never satisfied.
 - 281
 - 282 ii. Condition $\lambda_2^5 > -1$ requires that $\gamma D^* + \sqrt{\Delta^*} > -2\bar{w}^*$ which is always
 283 satisfied.
 - 284
 - 285 iii. Condition $\lambda_3^5 < 1$ requires that $\gamma D^* - \sqrt{\Delta^*} < 0$ which is always satisfied
 286 because $\gamma D^* < (\gamma D^*)^2 + \frac{1}{4}\beta(\beta - \gamma)$.
 - 287
 - 288 iv. Condition $\lambda_3^5 > -1$ requires that $\gamma D^* - \sqrt{\Delta^*} > -2\bar{w}^*$ which is always
 289 satisfied because $\alpha > \beta$ given the parametrisation of our model.

290

291 Notice that $\beta > \gamma$ implies that $D^* > 0$. In particular, from (14) we know that
 292 $D^* = \frac{1}{2(\beta - \gamma)} \left(\sqrt{\delta^2 + \frac{1}{4}(\beta - \gamma)^2} - \delta \right)$ and given that $\delta^2 + \frac{1}{4}(\beta - \gamma)^2 > \delta^2$ the
 293 sign of D^* is always equal to the sign of $\beta - \gamma$.

294

295 (b) Eigenvalues $\lambda_{2,3}^5$ are complex conjugate numbers when $\Delta^* < 0$ and thus condi-
 296 tion $\beta < \gamma$ is necessary for having complex eigenvalues. If $\beta > \gamma$ and the
 297 eigenvalues $\lambda_{2,3}^5$ are complex numbers, the stability of the internal equilibrium
 298 requires that $|\lambda_{2,3}^5| < 1$. This requirement implies the satisfaction of a single
 299 condition.

- 300
- 301 i. Condition $|\lambda_2^5| = |\lambda_3^5| < 1$ requires that $2\gamma\bar{w}^*D^* - \frac{1}{4}\beta(\beta - \gamma) < 0$. Replac-
 302 ing \bar{w}^* and D^* with their definitions from (15) and (14) respectively, yields
 303 the new condition $\alpha - \Omega < 4\delta < \alpha + \Omega$ where $\Omega = \frac{(2\gamma - \beta)\sqrt{\gamma(\beta^3 - \beta^2\gamma + \alpha^2\gamma)}}{\beta\gamma}$.
 304 The term Ω is equal to α if $\beta = \gamma$ but is greater than α if $\beta < \gamma$. This
 305 can be shown by calculating the derivative of Ω with respect to β , $\frac{\partial\Omega}{\partial\beta}$,
 306 which is negative when $\beta < \gamma$. This is true when $\alpha > \beta, 2\delta$ as is the
 307 case given the parametrisation of our model. Because Ω is greater than α
 308 when $\beta < \gamma$, the stability condition $\alpha - \Omega < 4\delta < \alpha + \Omega$ can be replaced
 309 by $0 < 4\delta < 2\alpha$ which is always satisfied given the parametrisation of our
 310 model. Therefore, when eigenvalues $\lambda_{2,3}^5$ are complex, their modulus is
 311 always less than one.

312

313 2. Stability conditions derived from the fourth eigenvalue λ_4^5 . Eigenvalue λ_4^5 is a real
 314 number and the stability of the internal equilibrium requires that $-1 < \lambda_4^5 < 1$.
 315 This requirement implies the satisfaction of two conditions:

316 i. Condition $\lambda_4^5 < 1$ requires that $-\delta - 2D^*(\beta - \gamma) < 0$ which is always satisfied
 317 because $\beta - \gamma$ and D^* have the same sign and thus their product is always
 318 positive.

319
 320 ii. Condition $\lambda_4^5 > -1$ requires that $\delta + 2D^*(\beta - \gamma) < 2\bar{w}^*$. Replacing D^* and
 321 \bar{w}^* with their definitions from (14) and (15) respectively, yields the new con-
 322 dition $2(\alpha + 2\beta D^*) > \sqrt{\frac{1}{4}(\beta - \gamma)^2 + \delta^2}$. Because $2(\alpha + 2\beta D^*) > 2\alpha - \beta$ and
 323 $\frac{1}{2}(\gamma - \beta) + \delta > \sqrt{\frac{1}{4}(\beta - \gamma)^2 + \delta^2}$ the later condition is true when $2\alpha - \beta >$
 324 $\frac{1}{2}(\gamma - \beta) + \delta$ which is always satisfied for the parametrisation of our model.
 325

326 To summarize, internal equilibrium \mathbf{x}^{*5} is unstable (saddle) ($\lambda_2^5 > 1$ but $-1 < \lambda_{3,4}^5 <$
 327 1) if $\beta > \gamma$ ($f > \frac{1}{2}c$) but stable ($|\lambda_{2,3}^5| < 1$ and $-1 < \lambda_4^5 < 1$) if $\beta > \gamma$ ($f < \frac{1}{2}c$) (see
 328 Table 1 and Figure 2).

329 3.3. Dynamics

330 3.3.1. Infinite population

331 When viability selection is strong ($f > \frac{1}{2}c$) the dynamics of our system tend towards
 332 the fixation of one of the recombination enabling haplotypes (\mathbf{x}^{*1} or \mathbf{x}^{*4}) (Figure 2
 333 and 3.a). In these two corner equilibria, an individual recombination hotspot remains
 334 inactive and the genomic recombinational landscape remains unchanged (Figure 3.a).
 335 Furthermore, the PRDM9-like gene does not evolve and remains monomorphic. An
 336 unchanging recombinational landscape and a non-evolving PRDM9 gene, are inconsistent
 337 with empirical observations on the life history of recombination hotspots controlled by
 338 PRDM9 (Ptak et al., 2004, 2005; Winckler et al., 2005; Coop et al., 2008; Myers et al.,
 339 2010; Stevison et al., 2015).

340 When viability selection is weak ($f < \frac{1}{2}c$) and initially all haplotypes are present in
 341 the population, the dynamics of our system oscillate towards a polymorphic equilibrium
 342 where all haplotypes (enabling and disabling) are present (\mathbf{x}^{*5}) (Figure 2 and 3.b). At
 343 this interior equilibrium, an individual recombination hotspot will see its activity reduced
 344 but not extinguished, and the genomic recombinational landscape remains unchanged
 345 (Figure 3.b). Furthermore, the PRDM9-like gene does not evolve but remains polymor-
 346 phic. An unchanging recombinational landscape and a non-evolving PRDM9 gene, are
 347 inconsistent with empirical observations (Ptak et al., 2004, 2005; Winckler et al., 2005;
 348 Coop et al., 2008; Myers et al., 2010; Stevison et al., 2015).

349 When viability selection is weak ($f < \frac{1}{2}c$) and initially one haplotype is present while
 350 the others are rare mutants, the dynamics of our system oscillate towards a heteroclinic
 351 cycle where fixation of one of the recombination enabling haplotypes alternates with
 352 fixation of one of the recombination disabling haplotypes ($\dots \mathbf{x}^{*1} \rightarrow \mathbf{x}^{*2} \rightarrow \mathbf{x}^{*4} \rightarrow \mathbf{x}^{*3} \rightarrow$

353 \mathbf{x}^{*1} ...) (Figure 2 and 3.c). Along this cycle, an individual recombination hotspot will
 354 alternate between becoming inactive (die) and becoming active (resurrect) (Figure 3.c).
 355 Therefore, the recombinational landscape becomes highly dynamic (Figure 3.c). Fur-
 356 thermore, the PRDM9-like gene is evolving fast with selective sweeps that are harder
 357 when viability selection is higher within the lower range ($f < \frac{1}{2}c$). A changing recombi-
 358 national landscape and a rapidly evolving PRDM9 gene, are consistent with empirical
 359 observations on the life history of recombination hotspots controlled by PRDM9 (Ptak
 360 et al., 2004, 2005; Winckler et al., 2005; Coop et al., 2008; Myers et al., 2010; Stevison
 361 et al., 2015).

362 3.3.2. Finite population

363 We finally modelled the cases of an infinite population without recurrent mutation, to
 364 better characterize the interaction between selection and conversion. In nature however,
 365 the population is finite and mutations are introduced recurrently. We carried out the
 366 numerical analysis of a model for a finite population with recurrent mutations at the
 367 modifier and target locus to gain insight on the effect of these two variables in our
 368 conclusions. The typical dynamics are summarized in Figure 4. In this figure it can be
 369 observed that the cycling remains with an alternation of hotspots and coldspots.

370 In particular, when viability selection is weak ($f < \frac{1}{2}c$) and initially one of the
 371 haplotypes is much more frequent than all the others, the haplotypes fluctuate around
 372 the boundary of the simplex, what is the heteroclinic cycle in the corresponding infinite
 373 population model ($\mathbf{x}^{*1} \rightarrow \mathbf{x}^{*2} \rightarrow \mathbf{x}^{*4} \rightarrow \mathbf{x}^{*3} \rightarrow \mathbf{x}^{*1}$) (Figure 4.b). Intuitively, selection
 374 and conversion favor the oscillation of haplotypes towards the boundary of the simplex
 375 where genetic drift pushes some of them to extinction (Figure 4.b). Extinction slows
 376 down the oscillatory dynamics but does not put an end to them, recurrent mutations
 377 re-introduce the missing variation and the system finds itself in the initial conditions
 378 that favor the heteroclinic cycle (Figure 4.b).

379 When viability selection is weak ($f < \frac{1}{2}c$) and initially all the haplotypes are frequent,
 380 the haplotypes fluctuate around the interior of the simplex, what is the polymorphic
 381 equilibrium in the infinite population model (x^{*5}) (Figure 4.a). Intuitively, selection and
 382 conversion favor the oscillation of haplotypes towards the interior of the simplex but
 383 genetic drift prevents them from settling (Figure 4.a). Because these oscillations remain
 384 far from the boundary the extinction of haplotypes is rarely observed (Figure 4.a). In
 385 the absence of extinction, hotspots and coldspots alternate rapidly. Genetic drift allows
 386 the transition from wide oscillations around the boundary to narrower oscillations within
 387 the interior and back.

388 4. Discussion

389 We find that strong selection (defined as selection bigger than conversion) fixes hap-
 390 lotypes which enable double-strand breaks (this translates into individual recombination
 391 hotspots that exhibit high activity and do not die over time (Figure 3.a)). This finding
 392 recovers the result of previous models (Boulton et al., 1997; Pineda-Krch and Redfield,
 393 2005; Calabrese, 2007; Peters, 2008). In our model however, weak selection (defined

394 as selection smaller than conversion) does not fix any particular haplotype; it either
 395 maintains all haplotypes in constant proportions (which translates into individual re-
 396 combination hotspots that exhibit moderate activity and do not die (Figure 3.b)), or
 397 the proportion of each haplotype cycles over time (which translates into individual re-
 398 combination hotspots that exhibit low and high activity, dying and resurrecting in a
 399 constant cycle (Figure 3.c)). These two types of recombination hotspots are novel. An
 400 equilibrium that maintains a polymorphism at a PRDM9-like locus and its target has
 401 not been described (Latrille et al., 2017). A cycle whereby the same set of alleles at a
 402 PRDM9-like locus and its target site rotate has not yet been described.

403 It is possible to gain an intuitive interpretation of our formal results if we consider
 404 a mutant gene playing a game against another gene from a gamete pool in a diploid
 405 individual. A mutant gene can play four strategies ($A_1B_1, A_1B_2, A_2B_1, A_2B_2$) and the
 406 gamete pool is formed by the same four strategies. The payoff of each gene interac-
 407 tion is summarized in the payoff matrix provided in Figure 5 and is determined by the
 408 individual fitness cost of not experiencing a DSB ($F_k = f > 0$), the allelic conversion
 409 benefit (or cost) of not experiencing (or experiencing) a DSB ($C_b = \frac{1}{2}c > 0$), and a
 410 recombination shuffling factor that determines which alleles benefit from conversion in
 411 double heterozygotes ($R_s = f(r) > 0$). Lets start by considering a population almost
 412 fixed for a recombination enabling haplotype A_1B_1 . If fitness cost is greater than con-
 413 version benefit ($F_k > C_b$), our resident population of A_1B_1 cannot be invaded by any
 414 alternative strategy ($1 > 1 - \frac{1}{2}F_k + \frac{1}{2}C_b$; Figure 5.a). Therefore strong selection favors
 415 highly active permanent recombination hotspots (Figure 3.a). If fitness cost is lower
 416 than conversion benefit ($F_k < C_b$; Figure 5.b), our resident population of A_1B_1 can be
 417 invaded by the rare mutant A_1B_2 ($1 - \frac{1}{2}F_k + \frac{1}{2}C_b > 1 - \frac{1}{2}F_k + R_s > 1$ when $C_b > 2R_s$;
 418 Figure 5.b) as it gains a transmission advantage that more than compensates for its fit-
 419 ness cost; once A_1B_2 becomes the resident haplotype, it can be invaded by rare mutant
 420 A_2B_2 ($1 - \frac{1}{2}F_k > 1 - F_k$; Figure 5.b) as it gains a fitness benefit and does not suffer a
 421 transmission disadvantage, once A_2B_2 becomes the resident it can be invaded by rare
 422 mutant A_2B_1 ($1 - \frac{1}{2}F_k + \frac{1}{2}C_b > 1$; Figure 5.b), and once A_2B_1 becomes the resident
 423 it can be invaded by rare mutant A_1B_1 ($1 - \frac{1}{2}F_k > 1 - F_k$; Figure 5.b) thus complet-
 424 ing a recurrent cycle. Therefore weak selection and abundance of only one haplotype,
 425 can favor recombination hotspots that alternate between low and high activity; dying
 426 and resurrecting in cyclic succession (Figure 3.c). When all haplotypes are frequent in
 427 the initial population, the abundance of double heterozygotes results in the shuffling
 428 of the transmission advantage between different haplotypes. Depending on intensity of
 429 the shuffling. either the previous cycle is maintained or the best strategy becomes to
 430 play a fixed proportion of each strategy. Therefore weak selection and abundance of all
 431 haplotypes, can favor recombination hotspots that exhibit moderate activity and do not
 432 die (Figure 3.b), providing an intuitive interpretation of our analysis.

433 These findings provide an alternative solution to the recombination hotspots paradox
 434 (Boulton et al., 1997). In the prevailing explanation (the Red Queen theory), individ-
 435 ual recombination hotspots die and are saved from extinction in the genome by the
 436 birth of new recombination hotspots at new target sites in the genome (Myers et al.,

2010; Baudat et al., 2010; Úbeda and Wilkins, 2011). Viability selection favors mutant PRDM9 alleles that bind new target sites (Úbeda and Wilkins, 2011). In our model, viability selection does not prevent the death of individual recombination hotspots but saves them from extinction in the genome by driving their resurrection in homozygous targets where the effect of conversion is negligible (Figure 3, 4). Selection favors mutant PRDM9 alleles that bind the alternative target allele within the same target site. Both theories succeed in explaining the life history of recombination hotspots characterized by: i. the death of individual recombination hotspots not leading to their extinction in the genome (notice however that in principle the Red Queen theory would require a never ending supply of targets to prevent the extinction); ii. rapid change of the recombinational landscape; iii. rapid evolution of PRDM9. In our model however, this life history is explained by the bottom range of viability selection parameters which seems more plausible from an empirical perspective). Furthermore, our model makes novel predictions that the Red Queen (at least in its present formulation) does not. In particular, our model predicts that: i. the molecular signature near recombination hotspots should be the one of multiple recurrent events of high crossover activity as opposed to a single even of high crossover activity; ii. viability selection can maintain polymorphisms in PRDM9 (Latrille et al. (2017)); iii. the same genetic architecture under the same selection regime can result in two different families of recombination hotspots, one family with alternation of high and low activity and another family with constant intermediate activity.

For the purpose of characterizing the interplay between selection and conversion on the evolution of recombination hotspots, our model makes a series of simplifying assumptions provided in the Methods section. Many of these assumptions are standard in population genetics models and relaxing all of them is beyond the scope of this manuscript. However, relaxing some of them will help us to better understand the empirical relevance of our model. In particular, we discuss the implications of considering multiple alleles and target loci and a finite population.

Our analysis assumes one modifier and one target locus with two alleles in each locus. In humans there are multiple alleles segregating at locus PRDM9 and multiple alleles at each of many target sites. We numerically explored how our conclusions change when either the number of alleles in each locus is increased or the number of target loci is increased. The dynamics in a model with three alleles remains very similar. If viability selection is weaker than gene conversion, the cycling is still observed although the fluctuations become irregular and unpredictable. This is consistent with intuition, as a modifier converts its specific target, it amplifies the frequency of any of the remaining targets. If one of the remaining targets attains a sufficiently high frequency, it will then allow selection on the accompanying modifier. The dynamics in a model with two targets also remains very similar. If viability selection is weaker than gene conversion, the cycling is still observed and the fluctuations between different targets can be either synchronized or not. Multiple targets allows selection on modifiers that match one of both targets. More realistic models would require considering larger number of alleles and target sites.

480 Our analysis assumes an infinite population without recurrent mutations, however
481 real populations are finite and mutations are recurrent. We numerically explored how
482 our conclusions change when we consider a finite population with recurrent mutations.
483 If viability selection is weaker than gene conversion, the cycling is still observed although
484 the cycles now drift in amplitude due to the stochastic effects. The dynamics may spend
485 more time in the vicinity of the interior of the simplex (the interior equilibrium in the
486 infinite population), where genetic drift rarely pushes any allele to extinction (Figure
487 4.a). The dynamics may spend more time in the vicinity of the boundary of the simplex
488 (the heteroclinic cycle in the infinite population), where genetic drift often pushes some
489 of alleles to extinction and that the dynamics become stuck. Once a suitable mutation
490 occurs the dynamics continue fluctuating (Figure 4.b). Stochasticity allows transitions
491 from oscillations mostly around the interior to mostly around the boundary.

492 Relaxing some of our assumptions in our model, suggests that in a finite population
493 with multiple alleles and target locus our main result holds; individual hotspots will die
494 but they will resurrect later in evolutionary time, thus precluding their extinction from
495 the genome in the long term. Population size, mutation rates, and number of target will
496 affect the turnover rate of recombination hotspots but not the qualitative behavior of
497 the selection conversion dynamics mediated by haplotype matching. This suggests that
498 our solution to the recombination hotspot paradox is robust although larger numbers
499 of target sites and their interplay with population size need to be modelled before any
500 conclusion can be reached.

501 **Acknowledgements**

502 [FU] thanks Amaya and Emma Úbeda for comments and support.

503 **Author Contributions**

504 [FU] conceived the research, [FU] formulated the model, [TR] and [VJ] analyzed the
505 model with feedback from [FU], [TR] carried out the numerical analysis, and [FU] wrote
506 the paper with feedback from [TR] and [VJ].

507 **Figure Captions**

508 **Figure 1.** *Model for recombination initiated by specificity of the double strand break.*
509 Summary of the sequence of events modelled. We start with the production of a PRDM9-
510 like protein with a recognition sequence that may match the target motif (same color
511 sequence in recognition and motif) or not (different color sequence in recognition and
512 motif). If protein and target bind, we follow the canonical DSB repair model for the
513 initiation of recombination (Szostak et al., 1983; Sun et al., 1991). Once recombination
514 (including crossover and conversion effects) has been completed, we model Mendelian
515 segregation of haplotypes with no fitness cost. If protein and target do not bind, there is
516 no recombination and we model Mendelian segregation of haplotypes with a fitness cost.

517 Notice that sister chromatids are represented at the beginning and end of the figure but
 518 are omitted from the middle part for clarity.

519 **Figure 2.** *Equilibria and heteroclinic cycle.* Summary of the equilibria with biological
 520 meaning, their stability and the basin of attraction of the heteroclinic cycle for three
 521 alternative models. Each panel corresponds to a different model given a specific value
 522 of the conversion rate c . Shades of green correspond to different values of the crossover
 523 rate $r = \{0, \frac{1}{2}, 1\}$ (with darker green corresponding to no-crossover $r = 0$). For each pair
 524 of values (c, r) , the equilibrium frequency of haplotype $x_{1,1}^*$ is plotted as a function of
 525 the fitness cost f . Red lines depict corner equilibria \mathbf{x}^{*1} and \mathbf{x}^{*4} corresponding to the
 526 fixation of recombination enabling haplotypes (notice that these are independent of the
 527 values of c and r). Blue lines depict corner equilibria \mathbf{x}^{*2} and \mathbf{x}^{*3} corresponding to the
 528 fixation of recombination disabling haplotypes (notice that these are independent of the
 529 values of c and r). Green lines depict twice corner equilibrium \mathbf{x}^{*5} corresponding to a
 530 polymorphism between recombination enabling and disabling haplotypes. Continuous
 531 lines depict stable equilibria while dashed lines depict unstable equilibria. The green
 532 colored area corresponds to the region in the space formed by the initial frequencies
 533 $(x_{1,1}^0, 0, 0, 1 - x_{1,1}^0)$ and the fitness cost f where the system tends to the heteroclinic
 534 cycle $(\dots \mathbf{x}^{*1} \rightarrow \mathbf{x}^{*2} \rightarrow \mathbf{x}^{*4} \rightarrow \mathbf{x}^{*3} \rightarrow \mathbf{x}^{*1} \dots)$ as opposed to the internal equilibrium \mathbf{x}^{*5} .
 535 Shades of green correspond to different values of the crossover rate $r = \{0, \frac{1}{2}, 1\}$ (with
 536 darker green corresponding to no-crossover $r = 0$). (i) The first panel corresponds to the
 537 case presented in the main text where selection is determined by double-strand breaks,
 538 and crossover events between the PRDM-9 and its target loci require a double-strand
 539 break at the target locus. (ii) The second panel corresponds to the case where selection
 540 is determined by crossover events, and crossover events between the PRDM-9 and its
 541 target loci require a double-strand break at the target locus. (iii) The third panel corre-
 542 sponds to the case where selection is determined by double-strand breaks, and crossover
 543 events between the PRDM-9 and its target loci do not require a double-strand break at
 544 the target locus.

545 **Figure 3.** *Dynamics of the system.* Examples of the three types of dynamics we find
 546 in our system. Each panel corresponds to a different combination of parameter values
 547 (f, b) and initial conditions $(x_{1,1}^0, x_{1,2}^0, x_{2,1}^0, x_{2,2}^0)$, while parameter values r, c remain fixed
 548 across panels, in particular $(r, c) = (1, \frac{1}{2})$. Sub-panel (i) depicts the frequency of all
 549 haplotypes $(x_{1,1}, x_{1,2}, x_{2,1}, x_{2,2})$ at time t as a point in the three dimensional simplex.
 550 Arrows indicate in which direction the dynamics progress as time goes by. The color of
 551 the line depicts the population mean recombination activity of the target (see legend).
 552 Sub-panel (ii) stacks three plots, namely: each of the haplotype frequencies against gener-
 553 erational time, the population mean recombination activity as a line against time, and
 554 the population mean recombination activity as heat map against time. Panel (a) corre-
 555 sponds to parameter values $(f, b) = (0.44, 0.50)$ and initial conditions $(0, x_{1,2}^0, 1 - x_{1,2}^0, 0)$
 556 where $x_{1,2}^0 = 0.33$ or $x_{1,2}^0 = 0.66$. (a.i) shows that when the initial condition is $x_{1,2}^0 = 0.33$
 557 the system tends to corner equilibrium \mathbf{x}^{*1} . When the initial condition is $x_{1,2}^0 = 0.66$ the
 558 system tends to the other stable corner equilibrium \mathbf{x}^{*4} . In both cases the target site
 559 at equilibrium is a recombination hotspot (target colored). (a.ii) shows that when the

560 initial condition is $x_{1,2}^0 = 0.33$ the recombination enabling haplotype $x_{1,1}$ becomes fixed.
 561 There are no changes at the modifier locus coding for PRDM9-like proteins. The popu-
 562 lation mean recombination activity reaches and remains over time at its highest (1). The
 563 target site becomes and remains a recombination hotspot over time. Panel (b) corre-
 564 sponds to parameter values $(f, b) = (0.22, 0.25)$ and initial conditions $(x_{1,1}^0, 0, 0, 1 - x_{1,1})$
 565 where $x_{1,1} = 0.80$. (b.i) shows that the system tends to internal equilibrium \mathbf{x}^{*5} where
 566 the target site is what we called a recombination warmspot. (b.ii) shows that the fre-
 567 quency of all haplotypes oscillate in their approach to equilibrium where all haplotypes
 568 (recombination enabling and disabling) are present. There are oscillations at the locus
 569 coding for PRDM9-like proteins in the approach to equilibrium but these changes cease
 570 when equilibrium is reached. The population mean recombination activity oscillates be-
 571 tween high and low as it approaches an intermediate value (0.5) at equilibrium. The
 572 target site oscillates between hot and cold phenotypes as it approaches a warm pheno-
 573 type at equilibrium. Panel (c) corresponds to parameter values $(f, b) = (0.22, 0.75)$ and
 574 initial conditions $(x_{1,1}^0, 0, 0, 1 - x_{1,1}^0)$ where $x_{1,1}^0 = 0.90$. (c.i) shows that the system tends
 575 to the heteroclinic cycle $(\dots \mathbf{x}^{*1} \rightarrow \mathbf{x}^{*2} \rightarrow \mathbf{x}^{*4} \rightarrow \mathbf{x}^{*3} \rightarrow \mathbf{x}^{*1} \dots)$. (c.ii) shows that the
 576 frequency of all haplotypes oscillate in their approach to the heteroclinic cycle where
 577 there is an alternation between near fixation of one of the recombination enabling hap-
 578 lotypes and near fixation of one of the recombination disabling haplotypes. There are
 579 oscillations at the locus coding for PRDM9-like proteins, oscillations that become in-
 580 creasingly pronounced as the system approaches the heteroclinic cycle. The population
 581 mean recombination activity oscillates between high and low, oscillations that become
 582 increasingly pronounced as the system approaches the heteroclinic cycle. The target
 583 site oscillates between hot and cold phenotypes with it hot and cold character becoming
 584 more marked as the system approaches the heteroclinic cycle.

585 **Figure 4.** *Comparison with finite populations.* Examples of the correspondence
 586 between dynamics in the infinite and finite population models. Each panel corresponds
 587 to a different combination of parameter values (f, b) and (μ, N) where μ is the mutation
 588 rate and N is the population size. Parameter values $(r, c) = (1, \frac{1}{2})$ and initial conditions
 589 $(x_{1,1}^0, x_{1,2}^0, x_{2,1}^0, x_{2,2}^0) = (0.99, \frac{1}{3}0.01, \frac{1}{3}0.01, \frac{1}{3}0.01)$ remain fixed across panels. Sub panel
 590 (i) stacks three plots, namely: each of the haplotype frequencies against generational
 591 time, the population mean recombination activity against time, and the population mean
 592 recombination activity as a heat map against time. Sub panel (ii) depicts the frequency of
 593 all haplotypes $(x_{1,1}, x_{1,2}, x_{2,1}, x_{2,2})$ at time t as a point in the three dimensional simplex.
 594 Arrows indicate in which direction the dynamics progress as time goes by. The color
 595 of the line depicts the population mean recombination activity of the target site (see
 596 legend). Panel (a) corresponds to parameter values $(f, b) = (0.22, 1.00)$ and $(\mu, N) =$
 597 $(10^{-5}, 10^4)$. The target site oscillates between hot and cold phenotypes rapidly and no
 598 haplotype becomes fixed. Panel (b) corresponds to parameter values $(f, b) = (0.22, 1.00)$
 599 and $(\mu, N) = (10^{-6}, 10^4)$. The target site oscillates between hot and cold phenotypes
 600 slowly and haplotypes often become fixed.

601 **Figure 5.** *Evolutionary game.* Payoff matrix of a game played by each haplotype
 602 against a haplotype pool. The payoff is determined by the possibility of a diploid geno-

603 type containing that haplotype experiencing a fitness cost (F_k) due to the absence of a
604 double-strand break, a conversion benefit (C_b)—or conversion cost ($-C_b$)— due to the
605 conversion of the opponent’s haplotype into the player’s haplotype—or the conversion
606 of the player’s haplotype into the opponent’s haplotype, and a reshuffling benefit or cost
607 due to the formation of the player’s or the opponent’s haplotype due to the formation of
608 new combinations of alleles. In the first matrix we assume that the fitness cost is greater
609 than the conversion benefit ($F_k > C_b$). Starting with a population fixed for haplotype
610 A_1B_1 , A_1B_1 is the mutant strategy that gives the highest payoff (in grey in the matrix).
611 No mutant haplotype can invade and A_1B_1 is the only evolutionary stable strategy. In
612 the second matrix we assume that the fitness cost is smaller than the conversion benefit
613 ($F_k < C_b$). Starting with a population fixed for haplotype A_1B_1 , A_1B_2 is the mutant
614 strategy that gives the highest payoff (in grey in the matrix) and should take over the
615 population. When A_1B_2 has become the resident strategy, A_2B_2 is the mutant strategy
616 that gives the highest payoff (in grey in the matrix) and should take over the popula-
617 tion. Using the same logic becomes obvious that in this second game there is no pure
618 evolutionary stable strategy but a continuous cycling of strategies.

619 **Table 1.** *Stability.* The *eigenvalue* column contains the eigenvalues corresponding to
620 each equilibrium with biological meaning (\mathbf{x}^{*1-5}). The *stability* column summarizes the
621 analysis of the stability of each equilibrium using their eigenvalues. This analysis shows
622 that the stability of all equilibria is determined by a single condition, namely whether
623 $\beta > \gamma$ or not .

624 Alves, I., Houle, A. A., Hussin, J. G., Awadalla, P., 2017. The impact of recombination
625 on human mutation load and disease. *Phil. Trans. R. Soc. B* 372 (1736), 20160465.

626 Baker, B. S., Carpenter, A. T., Esposito, M. S., Esposito, R. E., Sandler, L., 1976. The
627 genetic control of meiosis. *Annual review of genetics* 10 (1), 53–134.

628 Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G.,
629 De Massy, B., 2010. Prdm9 is a major determinant of meiotic recombination hotspots
630 in humans and mice. *Science* 327 (5967), 836–840.

631 Baudat, F., Imai, Y., De Massy, B., 2013. Meiotic recombination in mammals: localiza-
632 tion and regulation. *Nature Reviews Genetics* 14 (11), 794.

633 Boulton, A., Myers, R. S., Redfield, R. J., 1997. The hotspot conversion paradox and the
634 evolution of meiotic recombination. *Proceedings of the National Academy of Sciences*
635 94 (15), 8058–8063.

636 Brick, K., Smagulova, F., Khil, P., Camerini-Otero, R. D., Petukhova, G. V., 2012.
637 Genetic recombination is directed away from functional genomic elements in mice.
638 *Nature* 485 (7400), 642.

639 Bürger, R., 2000. *The mathematical theory of selection, recombination, and mutation.*
640 John Wiley & Sons.

- 641 Calabrese, P., 2007. A population genetics model with recombination hotspots that are
642 heterogeneous across the population. *Proceedings of the National Academy of Sciences*
643 104 (11), 4748–4752.
- 644 Coop, G., Myers, S. R., 2007. Live hot, die young: transmission distortion in recombina-
645 tion hotspots. *PLoS genetics* 3 (3), e35.
- 646 Coop, G., Wen, X., Ober, C., Pritchard, J. K., Przeworski, M., 2008. High-resolution
647 mapping of crossovers reveals extensive variation in fine-scale recombination patterns
648 among humans. *science* 319 (5868), 1395–1398.
- 649 Hassold, T., Hunt, P., 2001. To err (meiotically) is human: the genesis of human aneu-
650 ploidy. *Nature Reviews Genetics* 2 (4), 280.
- 651 Hassold, T., Sherman, S., Hunt, P., 2000. Counting cross-overs: characterizing meiotic
652 recombination in mammals. *Human Molecular Genetics* 9 (16), 2409–2419.
- 653 Hey, J., 2004. What’s so hot about recombination hotspots? *PLoS biology* 2 (6), e190.
- 654 Karlin, S., Feldman, M. W., et al., 1970. Linkage and selection: two locus symmetric
655 viability model. *Theoretical population biology* 1 (1), 39–71.
- 656 Koehler, K. E., Hawley, R. S., Sherman, S., Hassold, T., 1996. Recombination and
657 nondisjunction in humans and flies. *Human molecular genetics* 5 (Supplement_1),
658 1495–1504.
- 659 Latrille, T., Duret, L., Lartillot, N., 2017. The red queen model of recombination hot-spot
660 evolution: a theoretical investigation. *Phil. Trans. R. Soc. B* 372 (1736), 20160463.
- 661 Lesecque, Y., Glémin, S., Lartillot, N., Mouchiroud, D., Duret, L., 2014. The red queen
662 model of recombination hotspots evolution in the light of archaic and modern human
663 genomes. *PLoS genetics* 10 (11), e1004790.
- 664 Lichten, M., Goldman, A. S., 1995. Meiotic recombination hotspots. *Annual review of*
665 *genetics* 29 (1), 423–444.
- 666 Louis, E. J., Borts, R. H., 2003. Meiotic recombination: too much of a good thing?
667 *Current Biology* 13 (24), R953–R955.
- 668 Myers, S., Bottolo, L., Freeman, C., McVean, G., Donnelly, P., 2005. A fine-scale map
669 of recombination rates and hotspots across the human genome. *Science* 310 (5746),
670 321–324.
- 671 Myers, S., Bowden, R., Tumian, A., Bontrop, R. E., Freeman, C., MacFie, T. S.,
672 McVean, G., Donnelly, P., 2010. Drive against hotspot motifs in primates implicates
673 the *prdm9* gene in meiotic recombination. *Science* 327 (5967), 876–879.
- 674 Paigen, K., Petkov, P., 2010. Mammalian recombination hot spots: properties, control
675 and evolution. *Nature Reviews Genetics* 11 (3), 221.

- 676 Peters, A., 2008. A combination of cis and trans control can solve the hotspot conversion
677 paradox. *Genetics*.
- 678 Petes, T. D., 2001. Meiotic recombination hot spots and cold spots. *Nature Reviews*
679 *Genetics* 2 (5), 360.
- 680 Pineda-Krch, M., Redfield, R., 2005. Persistence and loss of meiotic recombination
681 hotspots. *Genetics*.
- 682 Ptak, S. E., Hinds, D. A., Koehler, K., Nickel, B., Patil, N., Ballinger, D. G., Przeworski,
683 M., Frazer, K. A., Pääbo, S., 2005. Fine-scale recombination patterns differ between
684 chimpanzees and humans. *Nature genetics* 37 (4), 429.
- 685 Ptak, S. E., Roeder, A. D., Stephens, M., Gilad, Y., Pääbo, S., Przeworski, M., 2004.
686 Absence of the tap2 human recombination hotspot in chimpanzees. *PLoS biology* 2 (6),
687 e155.
- 688 Rosenberg, N. A., Huang, L., Jewett, E. M., Szpiech, Z. A., Jankovic, I., Boehnke,
689 M., 2010. Genome-wide association studies in diverse populations. *Nature Reviews*
690 *Genetics* 11 (5), 356.
- 691 Russell, T. W., , Russell, M. J. Úbeda, F., Jansen, V. A. A., 2019. Stable cycling in
692 quasi-linkage equilibrium: fluctuating dynamics under gene conversion and selection.
693 (Submitted).
- 694 Ségurel, L., Leffler, E. M., Przeworski, M., 2011. The case of the fickle fingers: how the
695 prdm9 zinc finger protein specifies meiotic recombination hotspots in humans. *PLoS*
696 *biology* 9 (12), e1001211.
- 697 Stevison, L. S., Woerner, A. E., Kidd, J. M., Kelley, J. L., Veeramah, K. R., McManus,
698 K. F., Project, G. A. G., Bustamante, C. D., Hammer, M. F., Wall, J. D., 2015.
699 The time scale of recombination rate evolution in great apes. *Molecular biology and*
700 *evolution* 33 (4), 928–945.
- 701 Sun, H., Treco, D., Szostak, J. W., 1991. Extensive 3'-overhanging, single-stranded dna
702 associated with the meiosis-specific double-strand breaks at the arg4 recombination
703 initiation site. *Cell* 64 (6), 1155–1161.
- 704 Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J., Stahl, F. W., 1983. The double-
705 strand-break repair model for recombination. *Cell* 33 (1), 25–35.
- 706 Úbeda, F., Wilkins, J., 2011. The red queen theory of recombination hotspots. *Journal*
707 *of evolutionary biology* 24 (3), 541–553.
- 708 Winckler, W., Myers, S. R., Richter, D. J., Onofrio, R. C., McDonald, G. J., Bontrop,
709 R. E., McVean, G. A., Gabriel, S. B., Reich, D., Donnelly, P., et al., 2005. Comparison
710 of fine-scale recombination rates in humans and chimpanzees. *Science* 308 (5718), 107–
711 111.

712 Wright, S., 1969. Evolution and the genetics of populations: Vol. 2. The theory of gene
713 frequencies.