

Research



**Cite this article:** Jeffries DL, Gerchen JF, Scharmann M, Pannell JR. 2021 A neutral model for the loss of recombination on sex chromosomes. *Phil. Trans. R. Soc. B* **376**: 20200096.  
<https://doi.org/10.1098/rstb.2020.0096>

Accepted: 13 March 2021

One contribution of 12 to a theme issue 'Challenging the paradigm in sex chromosome evolution: empirical and theoretical insights with a focus on vertebrates (Part I)'.

**Subject Areas:**

evolution, computational biology, theoretical biology

**Keywords:**

sex chromosomes, recombination, neutral model, sexually antagonistic selection

**Author for correspondence:**

Daniel L. Jeffries  
e-mail: [dljeffries86@gmail.com](mailto:dljeffries86@gmail.com)

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5448716>.

# A neutral model for the loss of recombination on sex chromosomes

Daniel L. Jeffries, Jörn F. Gerchen, Mathias Scharmann and John R. Pannell

Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

DLJ, 0000-0003-1701-3978; MS, 0000-0001-8523-6888; JRP, 0000-0002-0098-7074

The loss of recombination between sex chromosomes has occurred repeatedly throughout nature, with important implications for their subsequent evolution. Explanations for this remarkable convergence have generally invoked only adaptive processes (e.g. sexually antagonistic selection); however, there is still little evidence for these hypotheses. Here we propose a model in which recombination on sex chromosomes is lost due to the neutral accumulation of sequence divergence adjacent to (and thus, in linkage disequilibrium with) the sex determiner. Importantly, we include in our model the fact that sequence divergence, in any form, reduces the probability of recombination between any two sequences. Using simulations, we show that, under certain conditions, a region of suppressed recombination arises and expands outwards from the sex-determining locus, under purely neutral processes. Further, we show that the rate and pattern of recombination loss are sensitive to the pre-existing recombination landscape of the genome and to sex differences in recombination rates, with patterns consistent with evolutionary strata emerging under some conditions. We discuss the applicability of these results to natural systems.

This article is part of the theme issue 'Challenging the paradigm in sex chromosome evolution: empirical and theoretical insights with a focus on vertebrates (Part I)'.

## 1. Introduction

A strikingly convergent pattern among species with genetic sex determination is the loss of recombination between their sex chromosomes [1]. This results in a lower efficacy of purifying or positive selection and opens the door to the accumulation of deleterious mutations, loss of function and altered patterns of gene expression. Given these important implications, the frequent loss of recombination on sex chromosomes across such a diverse range of taxa [2] demands an explanation. What is it about the regions adjacent to sex-determining (SD) genes that predisposes them to the loss of recombination? Despite the considerable attention this question has received, we still lack a fully satisfying answer.

So far, several hypotheses have been put forward, as recently reviewed in Ponnikas *et al.* [3]. Charlesworth & Wall [4] proposed that selection could favour expansion of the non-recombining region (NRR) if it encompasses deleterious recessive mutations, as they would then be sheltered by the fixed heterozygosity of the sex-linked region. More recently, Úbeda *et al.* [5] proposed that recombination suppression could be selected for in cases where the SD locus arises near to a meiotic driver or a drive suppressor, in order to ensure balanced sex ratios. Finally, the most widely cited explanation for suppressed recombination on sex chromosomes is sexually antagonistic (SA) selection. Under this hypothesis, selection should favour the suppression of recombination between the SD locus and a physically linked SA locus (i.e. a locus with opposite fitness effects in the two sexes). This is because increasing linkage disequilibrium (LD) between these loci will maintain each allele at the SA locus in its appropriate genetic background, thus resolving intragenomic conflict. First presented by

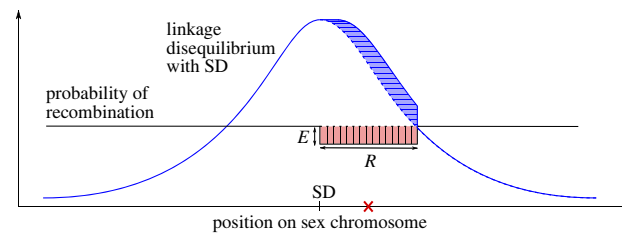
Fisher [6], this theory was later expanded theoretically by Charlesworth & Charlesworth [7], Bull [8] and Rice [9]. The SA selection hypothesis is appealing because it makes immediate intuitive sense for species in which males and females differ in morphology, behaviour, life history or physiology.

Despite the plausibility of the SA hypothesis, we still lack strong evidence for a causal relationship between selection of any kind and the loss of recombination on sex chromosomes. In several taxa, it seems that regions of low recombination likely pre-dated (and perhaps facilitated) the linkage of SA variants, meiotic drivers or deleterious recessive alleles to a SD locus [10–15]. Furthermore, recombination loss is seen in systems where sex chromosome turnovers are extremely common, for example in several families of fish [16–18]. Selection for linkage between a SD locus and, for instance, a SA locus that confers an advantage to the heterogametic sex, is expected to have a stabilizing effect on that SD system, as any new SD system that arises would have to overcome that advantage [19,20]. Thus, while it seems likely that selection has played a role in the evolution of suppressed recombination around SD loci, it remains unclear how important these processes have been.

Each of the hypotheses summarized above supposes an adaptive reason for the loss of recombination on sex chromosomes and they all invoke selection for genetic linkage with the sex determiner as a mechanism. By contrast, the possibility that recombination has been lost on sex chromosomes for non-adaptive reasons has received little attention. While it has been discussed briefly by Charlesworth [1,21] and Kent *et al.* [22], non-adaptive explanations have only been addressed at any length in two studies. Bengtsson & Goodfellow [23] developed a theoretical model (later coined the ‘attrition’ model [24]) to assess how neutral diversity can be expected to behave in the pseudoautosomal region (PAR) of mammalian sex chromosomes. They found that sequence divergence between the sex chromosomes should accumulate in the absence of selection if recombination rate was of the same order of magnitude as the mutation rate or lower.

More recently, Ironside [25] proposed a verbal model concerning the role of neutral inversions. On an autosome, a neutral inversion will momentarily prevent recombination in heterozygous individuals but will ultimately be fixed or lost via genetic drift. In either case, recombination will be free to resume. By contrast, if such an inversion arises and fixes on a Y or W chromosome and either encompasses the SD locus or overlaps with an existing NRR, it will be forever maintained in a heterozygous state by the extreme balancing selection on the SD locus. As such, recombination will be permanently halted in this region. While this model is intuitive, it does not explain the loss of recombination on sex chromosomes where no inversion is observed. But the idea that heterozygosity should be maintained around the SD locus should apply to any mutation that is in strong LD with the SD locus, not just to inversions. Thus, the key concept of Ironside’s model could apply much more generally.

Here, we propose a model for the non-adaptive loss of recombination on sex chromosomes that encompasses aspects of the models from both Bengtsson & Goodfellow [23] and Ironside [25]. In our model, frequency-dependent selection and the genetics of sex determination maintain polymorphism indefinitely at the SD locus itself. If neutral or nearly neutral mutations (henceforth referred to simply as neutral) occur at a rate that is sufficiently high relative to the local

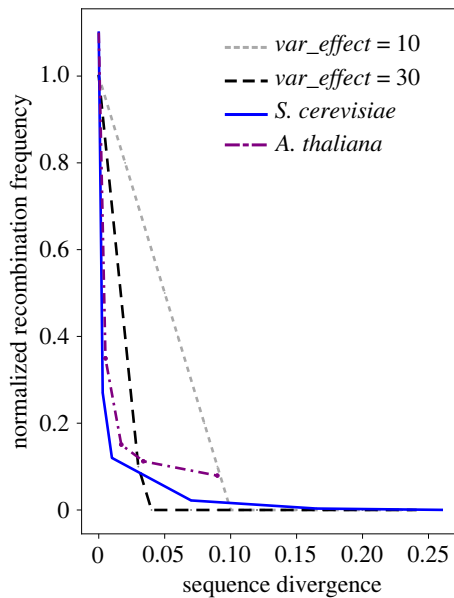


**Figure 1.** Schematic representation of our model for loss of recombination on sex chromosomes. For a given heterozygous position (red cross), initial recombination probability ( $r_{0i}$ , dashed black line) is reduced by a factor of  $E$  in a window around it of width  $W$  (red vertical line-shaded area). This in turn increases LD with the SD for a portion of the chromosome (blue curved line and horizontal line-shaded area) and so facilitates the further accumulation of differences between the sex chromosomes. (Online version in colour.)

recombination rate, then short-range LD should allow the two haplotypes close to the SD locus to begin to diverge [23], as would be the case for any two populations of partially isolated sequences [26,27]. The novelty of our model comes from the notion that heterozygosity at any site reduces the likelihood of recombination at and close to that site. As such, neutral divergence that builds up in LD with the SD locus will, itself, reduce the recombination rate in that region, creating a positive feedback loop and facilitating the build-up of more divergence (figure 1). We suggest that this recursive process might eventually lead to a point where recombination no longer occurs at an appreciable level in the regions adjacent to the SD locus.

The negative relationship between recombination rate and sequence divergence has a strong empirical basis and is a widespread characteristic of genome evolution. Recombination between homologous sequences is fundamental to the reproductive process across much of the tree of life, but it opens the way to ectopic recombination (between non-allelic sites), which can result in massive, often lethal structural rearrangements and aneuploidy. It is thus not surprising that the recombination machinery has evolved to avoid recombination between dissimilar sequences, specifically via strict homology searching [28] and the mismatch repair pathway, which is conserved across prokaryotes and eukaryotes [29]. While the exact biomolecular processes involved are complex and differ between organisms, one apparent consistency is the significant inhibitory effect that even a few mismatched bases can have on recombination rates between homologous sequences [30–33] (figure 2). Given the maintenance of constant heterozygosity at SD loci, this could play a significant role in the evolution of sex chromosomes, yet to our knowledge, is still to be explored in this context.

Here, we used computer simulations to identify conditions under which a NRR on a Y-chromosome could arise and expand, and assess the applicability of this model to the natural world. Intuitively, divergence between sex chromosomes should accumulate adjacent to the SD locus as long as the mutation rate is sufficiently high relative to the recombination rate, and the rate at which neutral divergence builds up should be slowed with increasing population sizes. We therefore focused especially on these three core parameters. We also performed simulations to test the effects of both heterochiasmy (a difference in the rate and distribution of recombination events between sexes) and male-biased mutation rates, as both are common in eukaryotes, and are expected to influence sex chromosome evolution [10,12,34,35]. Lastly, as



**Figure 2.** The relationship between sequence divergence and normalized recombination rate found in *Saccaromyces cerevisiae* (blue line; [31]) and *Arabidopsis thaliana* [33]. The light and dark grey lines represent the relationship between sequence divergence and recombination probability as specified in our simulations using the  $E$  parameter. (Online version in colour.)

recombination also varies substantially over the genome, we examine the effect of a non-uniform recombination landscape, including local hotspots and coldspots (shared by both sexes), on the rate and pattern of recombination loss.

## 2. Simulation methods

To test our model, we simulated a 100 kb long stretch of DNA in a dioecious population with a fixed number of randomly mating diploid individuals and non-overlapping generations. We assumed an XY system of sex determination, with sex determined by a single base mutation located midway along with the simulated DNA sequence. In male meioses, half of the gametes received a Y male-determining allele, while the other half received its X homologue. Simulations began with no polymorphism in the population other than the SD nucleotide. In each generation, neutral point mutations were allocated at random positions along the sequence at rate  $\mu$ . If a mutation fell at an already polymorphic site, that mutation was reallocated. We also performed a test set of simulations in which we simply discarded mutations that landed at already polymorphic sites and found close agreement between the two approaches (see electronic supplementary material, figure S1).

Recombination events occurred at a sequence-wide rate of  $r$ , with their locations along the sequence,  $i$ , being drawn from a vector of recombination probabilities at each position  $r_i$ . Each simulation started with the vector  $r_i$  set with the initial (baseline) recombination probabilities at generation zero,  $r_{0i}$ . Based on this initial vector, recombination probabilities were calculated for each meiosis as a function of divergence between haplotypes at each position. Specifically, for a given target site  $i$ :

$$r_i = r_{0i} \times \left(1 - \frac{Eh}{W}\right),$$

where  $E$  is the factor by which each heterozygous position reduces recombination probability,  $h$  is the number of heterozygous positions in the given window and  $W$  is the width of the window centred on position  $i$ .

We assumed a linear decline in the recombination rate with sequence divergence (figure 2). After exploring the sensitivity of our model to parameter values, we set  $W=1000$  and  $E=30$ , with a single mutation thus reducing the baseline recombination rate by 3% ( $E=30$ ) for a 1000 bp window around it. To explore the impact of male mutation bias (MMB), we ran simulations in which male versus female rates differed by a constant factor for a given run, while maintaining a constant sex-averaged mutation rate at  $\mu=7.5 \times 10^{-6}$ . We explored the effect of the initial recombination landscape ( $r_i$ ) by comparing three scenarios that differed in the spatial pattern of the initial recombination probabilities: uniform, heterochiasmatic and hotspot/coldspot. In the heterochiasmatic model, the male recombination rate was higher towards the ends of the simulated sequence and lower at the centre, while in the hotspot/coldspot model, initial recombination probabilities fluctuated in a sinusoidal manner along the sequence in both sexes. In all cases, global recombination rates (averaged across the whole sequence and across both sexes) remained constant.

Because of constraints on computer time, we were limited to simulating small populations only ( $N=50$ – $200$ ); in the Results and discussion (§3), we consider the limitations this poses for our interpretations. Given the low population sizes investigated, we varied the mutation rate  $\mu$  and initial sequence-wide recombination rate  $r$  around  $N$ -scaled rates ( $\theta=4N\mu$  and  $\rho=4Nr$ ) that correspond to values typical for eukaryotes [36,37]. In simulations exploring different  $N$ , we kept  $\theta$  and  $\rho$  constant by scaling  $\mu$  and  $r$  accordingly. Simulations were run for 10 000, 20 000 or 30 000 generations depending on the parameters or scenarios being tested. All parameter combinations and simulation details are set out in table 1.

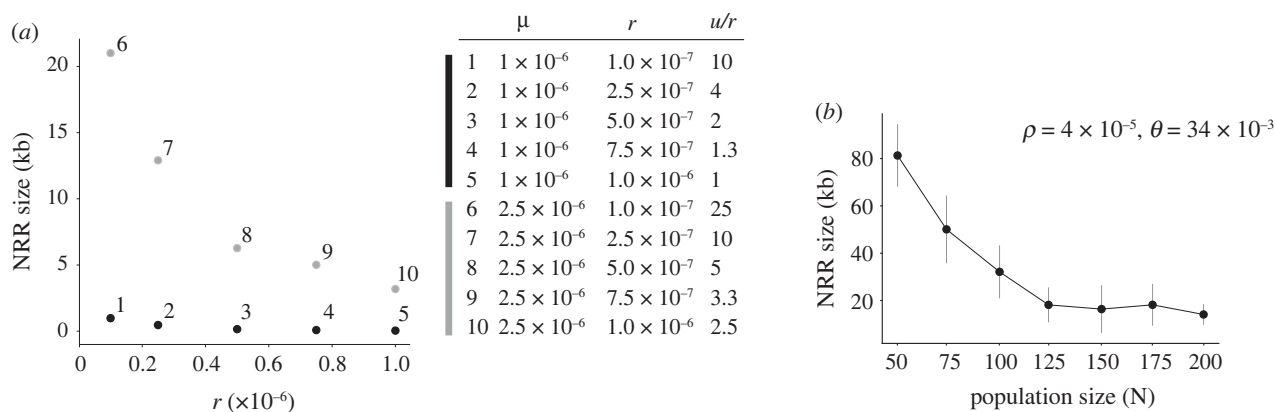
For each set of simulations, we recorded the size of the NRR every 100 generations. This was defined as a contiguous region containing the SD in which recombination probabilities in the vector  $r_i$  had reached zero in the entire population. We simultaneously recorded average sequence divergence in female meioses (i.e. between X–X pairs; dXX), the average sequence divergence in male meioses (i.e. between X–Y pairs; dXY) and the male–female  $F_{ST}$  at each position on the simulated stretch of DNA. All reported statistics (dXX, dXY,  $F_{ST}$ , recombination probabilities and NRR size) were averaged over 25 replicate runs.

Simulations were implemented in C++11. Full details and instructions for the installation and running of the model can be found on GitHub (<https://github.com/mscharmann/sexecevo>).

## 3. Results and discussion

### (a) Key results of the simulations

Our simulations identified conditions that promoted the evolution of reduced recombination around the SD locus, and those in which reduced recombination did not evolve. As predicted by the analytical work of Bengtsson & Goodfellow [23], recombination was lost more rapidly when  $\mu$  was high relative to  $r$  (e.g.  $N=100$ ,  $\mu \geq 1 \times 10^{-6}$ ,  $r \leq 1 \times 10^{-6}$ , figure 3a). Importantly, however, the rate at which



**Figure 3.** (a) Parameter space exploration for  $\mu$  and  $r$  when  $N = 100$ ,  $E = 30$  and  $W = 1000$ . Each point represents the averaged NRR size across all 25 iterations for each simulation after 20 000 generations. (b) Effect of population size on the size of the NRR after 20 000 generations. Plotted points and vertical lines represent the mean and standard deviation of the size of the NRR across all 25 iterations of each simulation.

recombination decreased depended not just on the relative values of  $\mu$  and  $r$ , but on their absolute values. This is best seen by comparing simulations 1 and 7 in figure 3a. In both,  $\mu/r = 10$ , but mutation and recombination rates are 2.5-fold higher in simulation 7. These conditions resulted in a faster initiation of recombination loss and a faster subsequent rate of loss in simulation 7 than in simulation 1 (electronic supplementary material, figure S2a). Given the above, it is somewhat surprising that simulations assuming a greater mutation rate in males than females revealed no significant increase in the rate at which the NRR expanded, with variation between the 25 simulations for each value of MMB being far greater than between sets of simulations with differing MMB values (electronic supplementary material, figure S2e). However, we note that, while the Y chromosome might experience an elevated mutation rate relative to the X chromosome, the X spends only a third of its time in males, so should experience the lowest mutation rate of any genome compartment. As recombination loss depends on polymorphism arising on both the Y and the X, it is possible that this reduction in the X mutation rate decreases the overall effect of a male-biased mutation in our model.

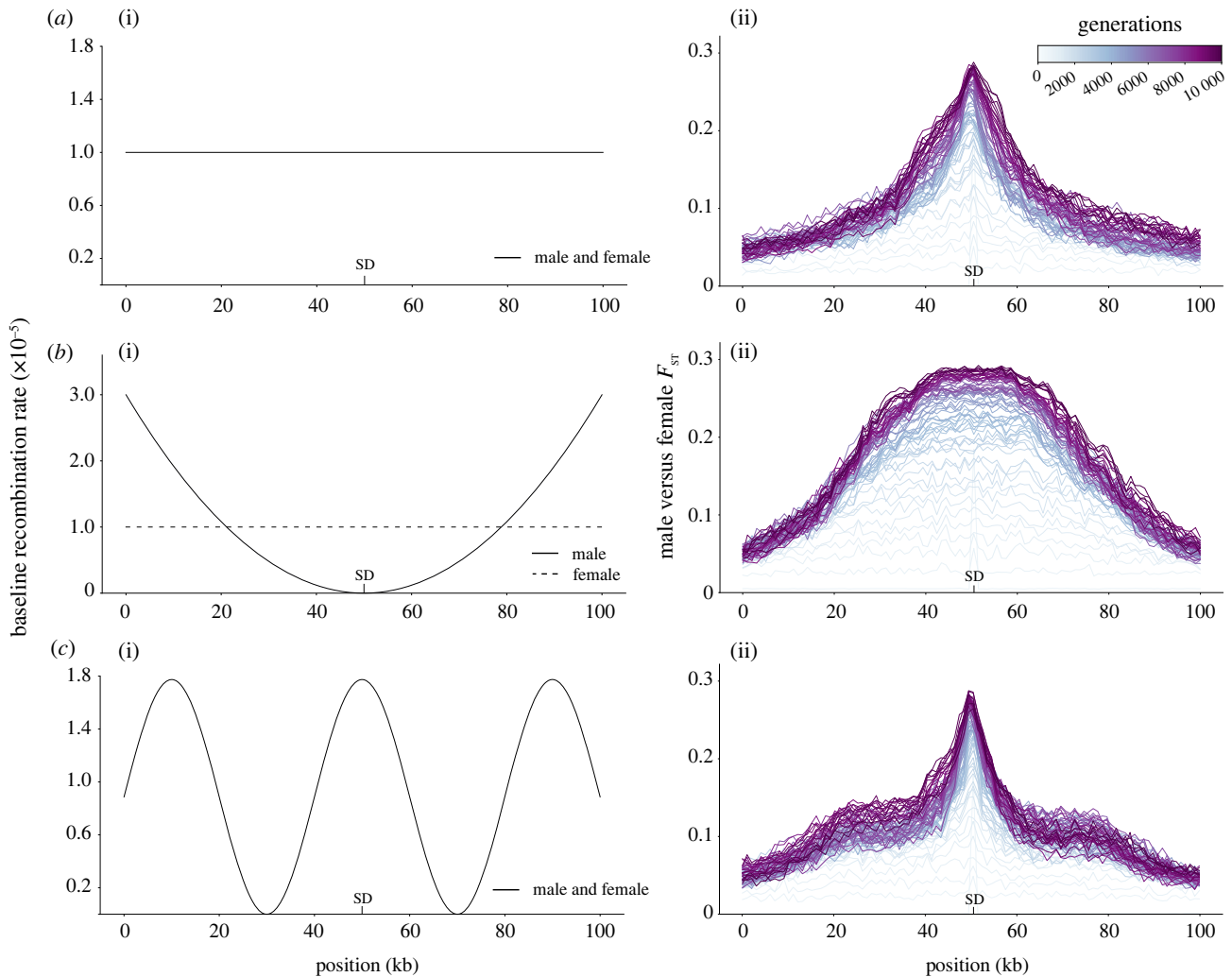
Recombination loss was also heavily dependent on population size, taking longer to begin, and progressing more slowly in larger populations (figure 3b; electronic supplementary material, figure S1b), presumably because of the slower coalescence for large  $N$  and thus slower divergence in sequence between homologous segments along with the simulated sequence. Interestingly, the size reached by the NRR over the course of our simulations dropped rapidly with increasing population size but levelled off with population sizes above 125 (figure 3b). This likely reflects the relationship between population size and the probability of fixation of neutral mutations [38], although more detailed simulations in which individual allele frequencies are tracked would be required to formally test this hypothesis.

The dynamics of recombination decline under our model are evidently complex and depend on the interaction between  $\mu$ ,  $r_i$  and  $N$ . To appreciate this complexity, consider the fate of a new mutation on one of the gametologs. Once a mutation arises near to the SD locus, it might be lost from the gametolog on which it arose, or it may be fixed there. Alternatively, if recombination between the new mutation and the SD locus

transfers it to the alternative gametolog, it may now be fixed or lost from either or both of them. The proportion of new mutations that succumb to each fate should depend on both the local recombination rate and the local effective population size. When both are low, mutations have a good chance of drifting to fixation on one gametolog before recombination communicates the mutation to the other. However, the number of mutations that eventually fix between gametologs is ultimately limited to a proportion of all mutations that arise, thus, the rate at which recombination is lost will always be inherently linked to the mutation rate. In addition, though many mutations will fail to fix between the gametologs, any polymorphism that exists in the intermediary stages of loss or fixation will contribute to heterozygosity and thus potentially further contribute to the reduction of recombination.

### (b) Limitations of the model and simulations

At the heart of our model is the relationship between recombination rate and sequence divergence. We assumed that recombination probability in a given region reduced linearly by a factor of  $E$  as sequence divergence increased, eventually reaching zero when sequence divergence =  $1/E$ . As expected, increasing the effect of each mutation on the reduction of recombination probability ( $E$ ) resulted in recombination loss at a lower value of  $\mu$  for a given  $r$  (electronic supplementary material, figure S2c). Similarly, increasing the width of the window affected by each mutation ( $W$ ) also resulted in an increase in the rate of recombination loss, though this effect was moderate compared to that of  $E$  (electronic supplementary material, figure S2d). Outside of these tests, simulations were performed with  $E = 30$ , which is similar to the initial drop in recombination for low levels of sequence divergence observed in *Saccharomyces cerevisiae* [31], *Arabidopsis thaliana* [33] (figure 2) and mammals [39]. However, in reality, this relationship appears to be more L-shaped, with very rare recombination events occasionally seen at high-sequence divergences (greater than 10%). This L-shaped relationship is likely governed by the molecular mechanisms of heteroduplex formation and the reliance on stretches of 100% sequence identity over a minimal length of DNA (termed 'minimal efficient processing segment', MEPS) [40,41]. However, MEPS length is known to differ substantially among taxa, from tens of bases in *S. cerevisiae* [30,32,42,43], to hundreds or thousands of bases in mammals [39,44]. In turn,



**Figure 4.** The effect of three different starting recombination landscapes (i) in the rate and pattern of neutral sex chromosome evolution ( $F_{ST}$ ; ii). (a) Uniform, (b) heterochiasmatic and (c) hotspot/coldspot starting recombination landscapes were tested. All simulations were run for 10 000 generations, and male versus female  $F_{ST}$  was averaged over 25 replicates. (Online version in colour.)

the relationship between recombination rate and sequence divergence is also likely to be highly variable throughout nature. Our simplified model likely gives us a first-order sense of what might happen given this diverse range of contexts, but it is important to note that rare recombination events at high-sequence divergences could homogenize sex chromosomes [45], reducing the expansion of non-recombination under our model. Therefore, a more nuanced approach to this aspect of our model, for instance using the random walk approach of Fujitani and Kobayashi [46] would undoubtedly be worthwhile in future work.

Our mutational model was also highly simplified compared to natural situations. First, we considered only the effect of heterozygous point mutations on recombination rates and not analogous effects that might arise from a structural variation such as inversions, insertions and deletions, which are likely to suppress recombination more strongly. Second, our simulations did not account for the effect of mutations that, by their nature, inhibit recombination (i.e. regardless of whether or not they are heterozygous). In mammals, for example, recombination initiation and the presence of hotspots are heavily reliant on specific sequence motifs that bind the chromatin-modifying protein PRDM9 [47]. Such motifs number in the tens of thousands in mammalian genomes [48], and so are likely to fall close to a sex determiner

wherever one arises. A single mutation at a PRDM9-binding site should substantially reduce the rate of recombination, regardless of whether it is heterozygous or homozygous. It has also been proposed that the accumulation of transposable elements in sex-linked regions could lead to silencing via methylation and chromatin alterations, which would also suppress recombination, regardless of heterozygosity [22]. These additional forms of mutation would surely increase the potential for neutral loss of recombination on sex chromosomes on top of that which we have modelled. As such, our model likely provides a conservative estimate of the parameter space in which we might expect this process to occur at the population sizes simulated.

### (c) Relevance of our model: how often might it apply in nature?

Due to computational limitations, we were only able to consider small populations from 50 to 200 individuals. As shown by our results, and as expected on the basis of the longer coalescent times, neutral loss of recombination occurred more slowly for populations with larger effective sizes. Thus, it is difficult to extrapolate how low the recombination rate and how high the mutation rate would have to be to allow for the neutral expansion of an NRR in more realistic scenarios. We

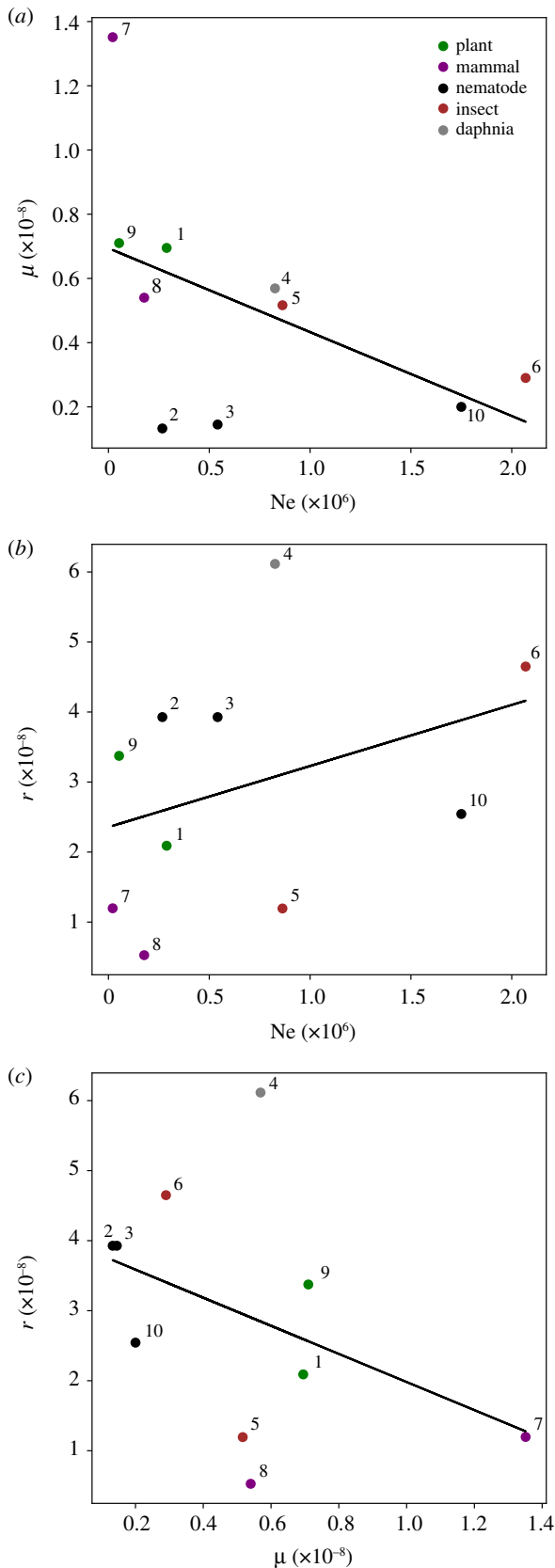
**Table 1.** Details of parameter values used in all simulations performed. Each row summarizes all parameter ranges used for each set of simulations. The focal test parameter of each set is highlighted in italics. Theoretical linkage map lengths for each simulation are given below the recombination rate.

simulation set	<i>N</i>	generations ( <i>x1000</i> )	$\mu$	<i>r</i>	$\theta$	$\rho$	<i>W</i>	<i>E</i>	MMB	ancestral recombination landscape
1	100	20–30	$1 \times 10^{-6}$ – $1 \times 10^{-5}$	$1 \times 10^{-7}$ – $1 \times 10^{-6}$ (1–10 cM)	$4 \times 10^{-4}$ – $4 \times 10^{-3}$	$4 \times 10^{-5}$ – $4 \times 10^{-4}$	1000	10 & 30	1	uniform
2	50–200	20	$3.75 \times 10^{-6}$ – $1.5 \times 10^{-5}$	$5 \times 10^{-8}$ – $2 \times 10^{-7}$ (0.5–2 cM)	$3 \times 10^{-3}$	$4 \times 10^{-5}$	1000	10	1	uniform
3	100	20	$7.5 \times 10^{-6}$	$1 \times 10^{-7}$ (1 cM)	$3 \times 10^{-3}$	$4 \times 10^{-5}$	1000	6–14	1	uniform
4	100	20	$7.5 \times 10^{-6}$	$1 \times 10^{-7}$ (1 cM)	$3 \times 10^{-3}$	$4 \times 10^{-5}$	100–1000	10	1	uniform
5	100	10	$7.5 \times 10^{-6}$	$1 \times 10^{-7}$ (1 cM)	$3 \times 10^{-3}$	$4 \times 10^{-5}$	1000	10	1	heterochiasmatic
6	100	10	$7.5 \times 10^{-6}$	$1 \times 10^{-7}$ (1 cM)	$3 \times 10^{-3}$	$4 \times 10^{-5}$	1000	10	1	hotspot / coldspot
7	100	20	$7.5 \times 10^{-6}$	$1 \times 10^{-7}$ (1 cM)	$3 \times 10^{-3}$	$4 \times 10^{-5}$	1000	10	1–10	uniform

note, however, that while genome-wide recombination rates, mutation rates and effective population sizes vary strongly across eukaryotes, a general trend (from the limited estimates currently available) seems to be that taxa with low population sizes tend to also possess higher mutation rates (figure 5*a*) and lower recombination rates (figure 5*b*).

More important than genome-wide averages in mutation and recombination rates are local conditions close to the SD. For example, our simulations of a fluctuating recombination landscape showed that recombination loss proceeded quickly through pre-existing coldspots but slowly through hotspots (figure 4*c*). Interestingly, this generated patterns of divergence consistent with evolutionary strata (i.e. large regions across which recombination was lost at roughly the same time [49]). Furthermore, in situations where recombination rates differ between the sexes (either genome-wide or locally), regions in high LD with an SD locus or an existing NRR will experience similarly sex-biased conditions. Again, this was reflected in the results of our simulations of heterochiasmatic recombination rates (figure 4*b*), in which NRR expansion occurred more rapidly compared to simulations with a uniform recombination landscape in both sexes. Lastly, regions that experience LD with the SD will, to a proportional extent, have lower effective population sizes by virtue of their census numbers in the population, with fully sex-linked regions of the X and Y having  $\frac{3}{4}$  and  $\frac{1}{4}$  the effective size of autosomes, respectively. These effective size ratios may be further decreased by factors unrelated to census size: for example, sexual selection [50,51] on the heterogametic sex or mating strategies that increase variance in reproductive success (e.g. polygyny [52]) should tend to reduce the effective size of regions in LD with the SD locus.

One lineage with characteristics that might conform to the requirements of our model is humans. As in other mammals, genome-wide mutation rates in humans are high, genome-wide recombination rates are low and population sizes are low (figure 5). Furthermore, male recombination rates in humans are almost half those of females [56], male mutation rates are up to 20-fold higher [35] and mating systems in humans are thought to have been historically polygynous [52], contributing to a further potential reduction in the effective size of the sex-linked region. If these characteristics were shared by our mammalian ancestors, the neutral processes that we have modelled could perhaps have played a role in the loss of recombination on human or other mammalian sex chromosomes early on in their evolution. Of course, mammalian sex chromosomes are old and their various Y chromosome strata ceased to recombine hundreds of millions of years ago, so current patterns are not relevant to a process that might have been important in the initial expansion of a NRR. Bengtsson & Goodfellow [23] noted that recombination rates in the human PAR are particularly high and suggested that the maintenance of neutral sex-linked variation adjacent to the PAR boundary thus seemed unlikely. Given that at least one crossover is typically required per chromosome arm in order to ensure proper segregation during meiosis, chromosomes for which recombination has been lost over most of their length will, by necessity, possess high recombination rates in the PAR, as is the case for most mammalian sex chromosomes [57]. Thus, if the expansion of a NRR began under the influence of processes modelled here, it seems likely that the process would eventually become arrested by the increasing recombination rates in the shrinking PAR.



**Figure 5.** Relationships between mutation rates, recombination rates and effective population size in species where high-confidence estimates are available. Black lines represent simple linear regressions. Species include: 1. *Arabidopsis thaliana*, 2. *Caenorhabditis briggsae*, 3. *Caenorhabditis elegans*, 4. *Daphnia pulex*, 5. *Drosophila melanogaster*, 6. *Heliconius melpomene*, 7. *Homo sapiens*, 8. *Mus musculus*, 9. *Oryza sativa*, 10. *Pristionchus pacificus*. Effective population sizes and mutation rates were taken from Lynch *et al.* [53]; human and mouse recombination rates were calculated from the genetic maps in Bhéner *et al.* [54] and Cox *et al.* [55], respectively. All other recombination rates were taken from Stapley *et al.* [37]. (Online version in colour.)

While it is not clear from our simulations whether the neutral processes modelled here could ever be solely responsible for the observed recombination loss on sex chromosomes, it seems plausible that the process could act in concert with other causes of recombination suppression. We found that regions evolving the complete loss of recombination were always flanked by regions of high  $F_{ST}$  between males and females and a low (but non-zero) recombination probability (e.g. figure 4). Consequently, such regions would experience LD with the SD locus. It is well known that the ability of selection to maintain a SA polymorphism on a sex chromosome depends on strong LD between the SA and SD loci [58]. Thus, the reduction of the recombination rate through neutral sequence divergence in the region around an already sex-linked region could facilitate selection for further reduction of recombination if a SA locus, a meiotic driver or a deleterious recessive mutation arose in this region [10,12]. Conversely, both positive and negative selective interference caused by selection for such genes near to the boundary of a NRR would reduce effective population size in this region [59], thus increasing the potential for the processes modelled here to play a role. Under such circumstances, neutral and non-neutral processes might mutually facilitate one another, increasing the likelihood or rate of recombination suppression. Of course, even in such cases, there are also several processes that can counteract the progressive loss of recombination on sex chromosomes, whatever its mechanism. For example, sex chromosome turnovers would erase any divergence between gametologs [20,60], as can sex reversals [61].

In conclusion, our simulations indicate that, in principle, recombination can be lost on sex chromosomes as a result of a neutral process of homologous sequence divergence in LD with a SD locus. Theoretically, this process should also apply to any region of a genome that is strictly heterozygous, e.g. around mating-type loci [62], supergenes such as those on the social chromosomes of certain ant species [63,64] or loci governing dimorphisms such as distyly in plants [65]. However, further work will be necessary to determine how often this process operates in nature. Our simulations predict that the rate and pattern of recombination loss, as modelled here, should be strongly affected by population size. It would therefore be interesting to compare patterns of recombination loss on sex chromosomes to past demographic histories, perhaps comparing lineages with orthologous sex determiners but differing extents of recombination loss. Ultimately, directly testing for the mechanism of recombination loss on sex chromosomes will be challenging. However, we argue that if we are to consider adaptive explanations, they must be set against a neutral null model. A case in point here is the commonly observed pattern of evolutionary strata on sex chromosomes. While it is often proposed that strata result from selection for linkage between SD and SA loci (e.g. [66]), it is clear that neutral processes can, in principle, result in similar patterns of sequence variation between gametologs when there are non-uniform ancestral recombination landscapes.

**Data accessibility.** The simulation program written and used in this study is implemented in C++11. Full details and instructions for installation and running of this model can be found on GitHub (<https://github.com/mscharmann/sexrecevo>), along with python scripts used to plot the outputs of the model.

**Authors' contributions.** D.L.J. conceived the model, helped write the simulation program and wrote the manuscript. J.F.G. and M.S. wrote the

simulation program and commented on the manuscript. J.R.P. wrote the manuscript and provided guidance throughout the study.

**Competing interests.** We declare we have no competing interests.

**Funding.** This work was funded by the Swiss National Science Foundation (grants 310030\_185196 and 31003A\_163384) and the University of Lausanne.

## References

- Charlesworth D, Charlesworth B, Marais G. 2005 Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **95**, 118–128. (doi:10.1038/sj.hdy.6800697)
- Bachtrog D *et al.* 2014 Sex determination: why so many ways of doing it? *PLoS Biol.* **12**, e1001899. (doi:10.1016/j.tig.2018.04.001)
- Ponnikas S, Sigeman H, Abbott JK, Hansson B. 2018 Why do sex chromosomes stop recombining? *Trends Genet.* **34**, 492–503. (doi:10.1016/j.tig.2018.04.001)
- Charlesworth B, Wall JD. 1999 Inbreeding, heterozygote advantage and the evolution of neo-X and neo-Y sex chromosomes. *Proc. R. Soc. B* **266**, 51–56. (doi:10.1098/rspb.1999.0603)
- Úbeda F, Patten MM, Wild G. 2015 On the origin of sex chromosomes from meiotic drive. *Proc. R. Soc. B* **282**, e20141932. (doi:10.1098/rspb.2014.1932)
- Fisher RA. 1931 The evolution of dominance. *Biol. Rev.* **6**, 345–368. (doi: 10.1111/j.1469-185X.1931.tb01030.x)
- Charlesworth D, Charlesworth B. 1980 Sex differences in fitness and selection for centric fusions between sex-chromosomes and autosomes. *Genet. Res.* **35**, 205–214. (doi:10.1017/s0016672300014051)
- Bull JJ. 1983 *Evolution of sex determining mechanisms*. Menlo Park, CA: The Benjamin/Cummings Publishing Company, Inc.
- Rice WR. 1987 The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution* **41**, 911–914. (doi:10.1111/j.1558-5646.1987.tb05864.x)
- Bergero R, Gardner J, Bader B, Yong L, Charlesworth D. 2019 Exaggerated heterochiasmy in a fish with sex-linked male coloration polymorphisms. *Proc. Natl Acad. Sci. USA* **116**, 6924–6931. (doi:10.1073/pnas.1818486116)
- Bergero R, Charlesworth D. 2019 Reply to Wright *et al.*: how to explain the absence of extensive Y-specific regions in the guppy sex chromosomes. *Proc. Natl Acad. Sci. USA* **116**, 12 609–12 610. (doi:10.1073/pnas.1906633116)
- Sardell JM, Kirkpatrick M. 2020 Sex differences in the recombination landscape. *Am. Nat.* **195**, 361–379. (doi:10.1086/704943)
- Koerich LB, Wang X, Clark AG, Carvalho AB. 2008 Low conservation of gene content in the *Drosophila* Y chromosome. *Nature* **456**, 949–951. (doi:10.1038/nature07463)
- Bachtrog D, Mahajan S, Bracewell R. 2008 Massive gene amplification on a recently formed *Drosophila* Y chromosome. *Nat. Ecol. Evol.* **3**, 1587–1597. (doi:10.1038/s41559-019-1009-9)
- Rifkin JL, Beaudry FEG, Humphries Z, Choudhury BI, Barrett SCH, Wright SI. 2020 Widespread recombination suppression facilitates plant sex chromosome evolution. *Mol. Biol. Evol.* **38**, 1018–1030. (doi:10.1093/molbev/msaa271)
- Myosho T, Takehana Y, Hamaguchi S, Sakaizumi M. 2015 Turnover of sex chromosomes in *Celebensis* group Medaka fishes. *G3* **5**, 2685–2691. (doi:10.1534/g3.115.021543)
- Sardell JM, Josephson MP, Dalziel AC, Peichel CL, Kirkpatrick M. 2020 Contrasting tempos of sex chromosome degeneration in sticklebacks. *BioRxiv*. 2020.09.17.300236. (doi:10.1101/2020.09.17.300236)
- Lubieniecki KP, Lin S, Cabana EI, Li J, Lai YYY, Davidson WS. 2015 Genomic instability of the sex-determining locus in Atlantic salmon (*Salmo salar*). *G3* **22**, 2513–2522. (doi:10.1534/g3.115.020115)
- Pokorná M, Kratochvíl L. 2009 Phylogeny of sex-determining mechanisms in squamate reptiles: are sex chromosomes an evolutionary trap? *Zool. J. Linn. Soc.* **150**, 168–183. (doi:10.1111/j.1096-3642.2008.00481.x)
- Blaser O, Neuenschwander S, Perrin N. 2014 Sex-chromosome turnovers: the hot-potato model. *Am. Nat.* **183**, 140–146. (doi:10.1086/674026)
- Charlesworth D. 2017 Evolution of recombination rates between sex chromosomes. *Phil. Trans. R. Soc.* **372**, e20160456. (doi:10.1098/rstb.2016.0456)
- Kent TV, Uzunović J, Wright SI. 2017 Coevolution between transposable elements and recombination. *Phil. Trans. R. Soc.* **372**, e20160458. (doi:10.1098/rstb.2016.0458)
- Bengtsson BO, Goodfellow PN. 1987 The effect of recombination between the X and Y chromosomes of mammals. *Ann. Hum. Genet.* **51**, 57–64. (doi:10.1111/j.1469-1809.1987.tb00865.x)
- Ellis N, Yen P, Neiswanger K, Shapiro LJ, Goodfellow PN. 1990 Evolution of the pseudoautosomal boundary in Old World monkeys and great apes. *Cell* **63**, 977–986. (doi:10.1016/0092-8674(90)90501-5)
- Ironside JE. 2010 No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome evolution. *Bioessays* **32**, 718–726. (doi:10.1002/bies.200900124)
- Nei M, Feldman MW. 1972 Identity of genes by descent within and between populations under mutation and migration pressures. *Theor. Popul. Biol.* **3**, 460–465. (doi:10.1016/0040-5809(72)90017-2)
- Nordborg M, Inman H. 2002 Molecular population genetics. *Curr. Opin. Plant Biol.* **5**, 69–73. (doi:10.1016/s1369-5266(01)00230-8)
- Renkawitz J, Lademann CA, Jentsch S. 2014 Mechanisms and principles of homology search during recombination. *Nat. Rev. Mol. Cell Biol.* **15**, 369–383. (doi:10.1038/nrm3805)
- Spies M, Fishel R. 2015 Mismatch repair during homologous and homeologous recombination. *Cold Spring Harb. Perspect. Biol.* **7**, a022657. (doi:10.1101/cshperspect.a022657)
- Datta A, Adjiri A, New L, Crouse GF, Jinks Robertson S. 1996 Mitotic crossovers between diverged sequences are regulated by mismatch repair proteins in *Saccaromyces cerevisiae*. *Mol. Cell Biol.* **16**, 1085–1093. (doi:10.1128/mcb.16.3.1085)
- Datta A, Hendrix M, Lipsitch M, Jinks-Robertson S. 1997 Dual roles for DNA sequence identity and the mismatch repair system in the regulation of mitotic crossing-over in yeast. *Proc. Natl Acad. Sci. USA* **94**, 9757–9762. (doi:10.1073/pnas.94.18.9757)
- Ira G, Haber JE. 2002 Characterization of *RAD51*-independent break-induced replication that acts preferentially with short homologous sequences. *Mol. Cell Biol.* **22**, 6384–6392. (doi:10.1128/mcb.22.18.6384-6392.2002)
- Li L, Jean M, Belzile F. 2006 The impact of sequence divergence and DNA mismatch repair on homeologous recombination in *Arabidopsis*. *Plant J.* **45**, 908–916. (doi:10.1111/j.1365-3113X.2006.02657.x)
- Acosta A, Martínez-Pacheco ML, Díaz-Barba K, Porras N, Gutiérrez-Mariscal M, Cortez D. 2019 Deciphering ancestral sex chromosome turnovers based on analysis of male mutation bias. *Genome Biol. Evol.* **11**, 3054–3067. (doi:10.1093/gbe/evz221)
- Wilson Sayres MA, Makova KD. 2011 Genome analyses substantiate male mutation bias in many species. *Bioessays* **33**, 938–945. (doi:10.1002/bies.201100091)
- Krasovec M, Eyre-Walker A, Sanchez-Ferandin S, Piganeau G. 2017 Spontaneous mutation rate in the smallest photosynthetic eukaryotes. *Mol. Biol. Evol.* **34**, 1770–1779. (doi:10.1093/molbev/msx119)
- Stapley J, Feulner PGD, Johnston SE, Santure AW, Smadja CM. 2017 Variation in recombination frequency and distribution across eukaryotes: patterns and processes. *Phil. Trans. R. Soc. B* **372**, e20160455. (doi:10.1098/rstb.2016.0455)



38. Kimura M. 1962 On the probability of fixation of mutant genes in a population. *Genetics* **47**, 713–719. (doi:10.1093/genetics/47.6.713)
39. Hasty P, Rivera-Pérez J, Bradley A. 1991 The length of homology required for gene targeting in embryonic stem cells. *Mol. Cell Biol.* **11**, 5586–5591. (doi:10.1128/mcb.11.11.5586)
40. Fujitani Y, Yamamoto K, Kobayashi I. 1995 Dependence of frequency of homologous recombination on the homology length. *Genetics* **140**, 797–809. (doi:10.1093/genetics/140.2.797)
41. Fujitani Y, Kobayashi I. 1999 Effect of DNA sequence divergence on homologous recombination as analyzed by a random-walk model. *Genetics* **153**, 1973–1988. (doi:10.1093/genetics/153.4.1973)
42. Chen W, Jinks-Robertson S. 1999 The role of the mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast. *Genetics* **151**, 1299–1313. (doi:10.1093/genetics/151.4.1299)
43. Shen P, Huang HV. 1986 Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. *Genetics* **112**, 441–457. (doi:10.1093/genetics/112.3.441)
44. Mundia MM, Desai V, Magwood AC, Baker MD. 2014 Nascent DNA synthesis during homologous recombination is synergistically promoted by the rad51 recombinase and DNA homology. *Genetics* **197**, 107–119. (doi:10.1534/genetics.114.161455)
45. Grossen C, Neuenschwander S, Perrin N. 2012 The evolution of XY recombination: sexually antagonistic selection versus deleterious mutation load. *Evolution* **66**, 3155–3166. (doi:10.1111/j.1558-5646.2012.01661.x)
46. Fujitani Y, Kobayashi I. 2003 Asymmetric random walk in a reaction intermediate of homologous recombination. *J. Theor. Biol.* **220**, 359–370. (doi:10.1006/jtbi.2003.3167)
47. Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, De Massy B. 2010 PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* **327**, 836–840. (doi:10.1126/science.1183439)
48. Baker CL, Walker M, Kajita S, Petkov PM, Paigen K. 2014 PRDM9 binding organizes hotspot nucleosomes and limits Holliday junction migration. *Genome Res.* **24**, 724–732. (doi:10.1101/gr.170167.113)
49. Lahn BT, Page DC. 1999 Four evolutionary strata on the human X chromosome. *Science* **286**, 964–967. (doi:10.1126/science.286.5441.964)
50. Sayres MA W, Lohmueller KE, Nielsen R. 2014 Natural selection reduced diversity on human Y chromosomes. *PLoS Genet.* **10**, e1004064. (doi:10.1371/journal.pgen.1004064)
51. Wilder JA, Mobasher Z, Hammer MF. 2004 Genetic evidence for unequal effective population sizes of human females and males. *Mol. Biol. Evol.* **21**, 2047–2057. (doi:10.1093/molbev/msh214)
52. Hammer MF, Mendez FL, Cox MP, Woerner AE, Wall JD. 2008 Sex-biased evolutionary forces shape genomic patterns of human diversity. *PLoS Genet.* **4**, e1000202. (doi:10.1371/journal.pgen.1000202)
53. Lynch M, Ackerman MS, Gout J-F, Long H, Sung W, Thomas WK, Foster PL. 2016 Genetic drift, selection and the evolution of the mutation rate. *Nat. Rev. Genet.* **17**, 704–714. (doi:10.1038/nrg.2016.104)
54. Bhérec C, Campbell CL, Auton A. 2017 Refined genetic maps reveal sexual dimorphism in human meiotic recombination at multiple scales. *Nat. Commun.* **8**, e14994. (doi:10.1038/ncomms14994)
55. Cox A *et al.* 2009 A new standard genetic map for the laboratory mouse. *Genetics* **182**, 1335–1344. (doi:10.1534/genetics.109.105486)
56. Mank JE. 2009 The evolution of heterochiasmy: the role of sexual selection and sperm competition in determining sex-specific recombination rates in eutherian mammals. *Genet. Res.* **91**, 355–363. (doi:10.1017/S0016672309990255)
57. Smagulova F, Gregoret IV, Brick K, Khil P, Camerini-Otero RD, Petukhova GV. 2011 Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. *Nature* **472**, 375–378. (doi:10.1038/nature09869)
58. Jordan CY, Charlesworth D. 2012 The potential for sexually antagonistic polymorphism in different genome regions. *Evolution* **66**, 505–516. (doi:10.1111/j.1558-5646.2011.01448.x)
59. Hill WG, Robertson A. 1966 The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294. (doi:10.1017/S001667230800949X)
60. Jeffries DL *et al.* 2018 A rapid rate of sex-chromosome turnover and non-random transitions in true frogs. *Nat. Commun.* **9**, e4088. (doi:10.1038/s41467-018-06517-2)
61. Rodrigues N, Studer T, Dufresnes C, Perrin N. 2018 Sex-chromosome recombination in common frogs brings water to the fountain-of-youth. *Mol. Biol. Evol.* **35**, 942–948. (doi:10.1093/molbev/msy008)
62. Branco S *et al.* 2017 Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism. *Proc. Natl Acad. Sci. USA* **114**, 7067–7072. (doi:10.1073/pnas.1701658114)
63. Martinez-Ruiz C, Pracana R, Stolle E, Paris CI, Nichols RA, Wurm Y. 2020 Genomic architecture and evolutionary antagonism drive allelic expression bias in the social supergene of red fire ants. *eLife* **9**, e55862. (doi:10.7554/eLife.55862)
64. Schwander T, Libbrecht R, Keller L. 2014 Supergenes and complex phenotypes. *Curr. Biol.* **24**, 288–294. (doi:10.1016/j.cub.2014.01.056)
65. Brennan AC. 2017 Distyly supergenes as a model to understand the evolution of genetic architecture. *Am. J. Bot.* **104**, 5–7. (doi:10.3732/ajb.1600363)
66. Wright AE *et al.* 2017 Convergent recombination suppression suggests role of sexual selection in guppy sex chromosome formation. *Nat. Commun.* **8**, e14251. (doi:10.1038/ncomms14251)