Modeling the stepwise extension of recombination suppression on sex chromosomes and other supergenes through deleterious mutation sheltering

Authors: Paul Jay[1*], Emilie Tezenas[1,2,3], Amandine Véber[3], Tatiana Giraud[1]

Affiliations: [1]Ecologie Systématique Evolution, Bâtiment 360, CNRS, AgroParisTech, Université Paris-Saclay, 91400 Orsay, France
[2]Univ. Lille, CNRS, UMR 8198 – Evo-Eco-Paleo, F-59000 Lille, France
[3] MAP5, CNRS, Université de Paris, 75006 Paris, France

*corresponding author: paul.yann.jay@gmail.com

Abstract: Many organisms have sex chromosomes with large non-recombining regions that have expanded stepwise, generating "evolutionary strata" of differentiation. The reasons for this remain poorly understood, but the principal hypotheses proposed to date are based on antagonistic selection due to differences between sexes. However, it has proved difficult to obtain empirical evidence of a role for sexually antagonistic selection in extending recombination suppression, and antagonistic selection has been shown to be unlikely to account for the evolutionary strata observed on fungal mating-type chromosomes. There may, therefore, be other mechanisms involved in the extension of non-recombining regions. We show here, by mathematical modeling and stochastic simulation, that recombination suppression on sex chromosomes and around supergenes can expand in a stepwise manner under a wide range of parameter values simply because it shelters recessive deleterious mutations, which are ubiquitous in genomes. Permanently heterozygous alleles, such as the male-determining allele in XY systems, protect linked chromosomal inversions against the expression of their recessive mutation load, leading to the successive accumulation of inversions around these alleles without the need for antagonistic selection. Similar results were obtained with models assuming recombination-suppressing mechanisms other than chromosomal inversions, and for supergenes other than sex chromosomes, including those without XY-like asymmetry, such as fungal mating-type chromosomes. However, inversions capturing a permanently heterozygous allele were found to be less likely to spread when the mutation load was lower (e.g. under conditions of large effective population size, low mutation rates and high dominance coefficients). This may explain why sex chromosomes remain homomorphic in some organisms but are highly divergent in others. Here, we explicitly state and model a simple and testable hypothesis explaining the existence of stepwise extensions of recombination suppression on sex chromosomes, which can also be applied to mating-type chromosomes and supergenes in general.

Key words: associative overdominance, evolutionary strata, deleterious mutations, sex chromosomes, supergenes, inversions

Introduction

Sex chromosomes, mating-type chromosomes and supergenes in general are widespread in nature and control striking polymorphisms, such as sexual dimorphism or color polymorphism, in many organisms, including humans (*1–5*). Supergenes display puzzling genomic architectures, defined by extensive regions lacking recombination, encompassing multiple genes (*3*, *4*). Sex chromosomes, in particular, often appear to have evolved by stepwise extension of non-recombining regions, generating "evolutionary strata" of differentiation between haplotypes (*4*, *6*, *7*). It is generally thought that the reason for this gradual expansion of recombination suppression on sex chromosomes is that selection favors the linkage of sex-determining genes to sexually antagonistic loci [i.e. to alleles advantageous in one sex but disadvantageous in the other (*7–9*)]. However, theoretical concerns have been raised about this idea (*10*), and it has proved difficult to obtain empirical support for a role of antagonistic selection in the extension of recombination suppression (*6*, *11*, *12*). Furthermore, a gradual expansion of recombination suppression has been observed around many fungal mating-type loci (*5*) and at other supergenes, such as those controlling wing color in butterflies or social structure in ants (*13*, *14*). Some type of antagonistic selection may exist between morphs determined by color or social supergenes, but antagonistic selection between fungal mating types is highly unlikely, given their life cycles and the results of genomic investigations (*15*, *16*). For example, there has been a repeated stepwise extension of recombination suppression in anther-smut fungi despite the lack of a haploid phase consisting of cells of different mating types potentially expressing contrasted life history traits, leaving little room for antagonistic selection (*15*, *16*). Alternative explanations have been put forward for the expansion of non-recombining regions on sex chromosomes (*11*), such as meiotic drive (*17*), genetic drift (*18*), deleterious-mutation sheltering (*19*, *20*) and the neutral accumulation of sequence divergence (*21*), but the conditions in which such mechanisms could apply also appear to be restricted.

We develop here a general model for testing the idea that recombination suppression may extend stepwise around sex-determining or mating-type loci because it provides the fitness advantage of sheltering deleterious mutations segregating in nearby regions. Related hypotheses have been explored before, but in the restricted specific context of inbreeding (*19*, *20*). We model here a more general hypothesis, based on the idea that genomes carry numerous deleterious recessive variants, as suggested by studies in a wide range of biological systems and by the pervasiveness of inbreeding depression in nature (*22–27*). We use mathematical modeling and stochastic simulations to test the hypothesis that permanently heterozygous alleles, such as male-determining alleles in XY systems, protect linked chromosomal inversions against the expression of their recessive mutation load, potentially leading to an accumulation of inversions around permanently heterozygous alleles, generating evolutionary strata.

The rationale behind this model is illustrated in Figure 1. Consider a diploid population carrying partially recessive, deleterious mutations. The combined effects of recombination, mutation, selection and drift result in individuals carrying different numbers of deleterious variants genome-wide and within particular genomic regions (Figure 1A). Chromosomal inversions may occur at any position and suppress recombination when heterozygous, thereby capturing a specific combination of deleterious variants (i.e. a haplotype). Inversions capturing fewer deleterious variants than the population average for the region concerned have a fitness advantage and should, therefore, increase in frequency. Such inversions are advantageous due to associative overdominance, i.e. the inversion itself is neutral but it

captures a combination of alleles that is advantageous when heterozygous (*28*, *29*). However, as the frequency of an inversion increases, homozygotes for this inversion become more common. Homozygotes are at a strong disadvantage due to the recessive deleterious variants carried by these inversions, and selection against homozygotes would therefore be expected to prevent such inversions from reaching high frequencies (Figure 1B).

Now, consider an inversion that, by chance, captures a permanently heterozygous allele, such as the male-determining allele in an XY system. If this Y-linked inversion captures fewer deleterious variants than the population average, it should increase in frequency without ever suffering the deleterious consequences of having its load expressed. The recessive deleterious mutations captured by the sex-linked inversion are, indeed, fully associated with the permanently heterozygous, male-determining allele, and will, therefore, never occur as homozygotes. Unlike autosomal inversions, Y-linked inversions retain their fitness advantage with increasing frequency (Figure 1C). Hence, less heavily loaded Y-linked inversions would be expected to spread, becoming fixed in the population of Y chromosomes, resulting in a suppression of recombination between the X and Y chromosomes in the region covered by the inversion.

The successive fixation of additional inversions linked to this Y-fixed inversion should cause the non-recombining region to expand further, by the same process, thereby leading to the formation of a chromosome with a large non-recombining region around the sex-determining locus —i.e. a sex chromosome with evolutionary strata of differentiation (Figure 1D). We use the term “chromosomal inversions” here for simplicity, but the proposed mechanism would hold for any mechanism suppressing recombination, and we model several types of recombination suppressor. The proposed mechanism would be expected to apply to any genomic region with a permanently or near-permanently heterozygous locus, such as mating-type loci in fungi or supergenes in many organisms. The accumulation of deleterious mutations following recombination suppression has been extensively studied (*30*, *31*, *31–33*), but we investigate here its converse: that deleterious mutations could be a cause, and not only a consequence, of recombination suppression.

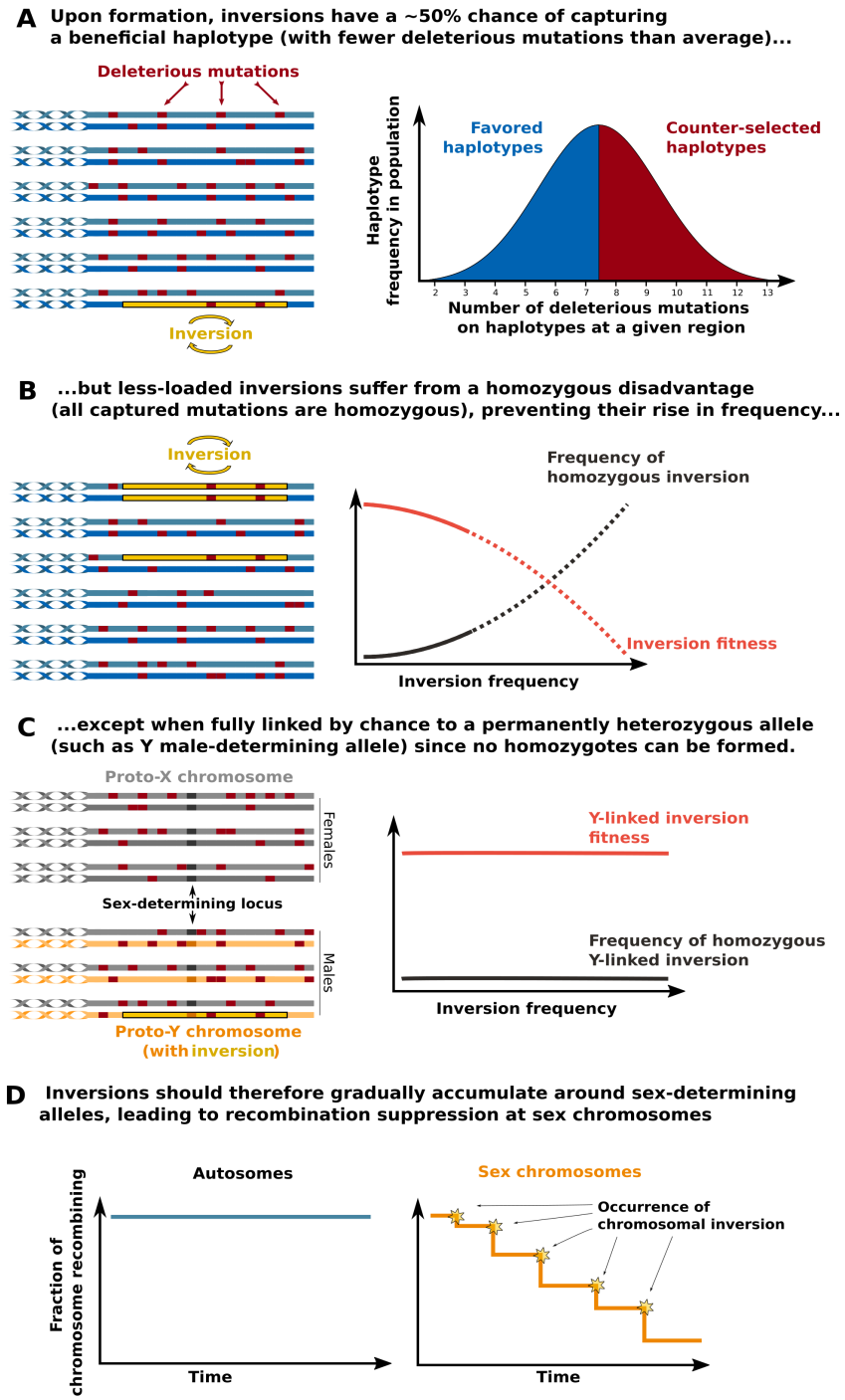

Figure 1 | Schematic diagram of the model

A, Within any population and for any genomic region, diploid individuals (here represented by two homologous chromosomes) and haplotypes (i.e. combinations of mutations, here one chromosome) carry variable numbers of partially recessive, deleterious mutations. A substantial fraction (about 50%) of haplotypes have fewer deleterious mutations than the average and should be favored by selection. Chromosomal inversions therefore have a significant chance of capturing beneficial haplotypes. **B**, The increase in frequency of a beneficial inversion leads to an increase in the frequency of homozygotes for the inversion, which have a fitness disadvantage because they are homozygous for all the deleterious recessive mutations carried by the inversion. The fitness of the inversion therefore decreases with increasing

inversion frequency, keeping the frequency of the inversion low. **C**, Permanent heterozygosity at a Y-chromosome sex-determining allele protects linked inversions from this disadvantage of homozygosity. Hence, beneficial inversions (carrying fewer deleterious mutations than average) should spread and become fixed in the population of Y chromosomes. **D**, Unlike those on autosomes, inversions capturing the Y sex-determining allele never suffer from homozygous disadvantage and should therefore accumulate on proto-Y chromosomes, leading to a suppression of recombination between the X and Y chromosomes. A similar mechanism should operate for any recombination suppressor acting in *cis* (not just chromosomal inversions) and any locus with permanently heterozygous alleles.

Results

We explored this general hypothesis of stepwise recombination suppression around permanently heterozygous alleles with infinite population deterministic models and individual-based simulations. Very little is known about the dynamics of deleterious mutations in genomic regions under particular recombination regimes, such as those generated by polymorphic inversions. We therefore had to use simulations to explore realistic scenarios involving various levels of drift. In both the infinite population model and individual-based simulations, we modeled diploid populations with only partially recessive deleterious mutations, occurring at a rate u, with heterozygous and homozygous individuals suffering from $1-hs$ and $1-s$ reductions in fitness, respectively, and with multiplicative effects of mutations on fitness. We first considered that all mutations occurring in genomes had the same dominance (h) and selection (s) coefficients, and then relaxed this hypothesis. Individuals were considered to have two pairs of chromosomes, one of which harbored a locus with at least one allele permanently or almost permanently heterozygous (see Methods for details). Several situations were considered, mimicking those encountered in XY sex-determination systems, in fungal mating-type systems and in overdominant supergenes. We analyzed the evolution of recombination modifiers suppressing recombination across the fragment in which they reside (i.e. *cis*-modifiers), either exclusively in heterozygotes (mimicking for example chromosomal inversions), or in both heterozygotes and homozygotes (e.g. histone modifications). Each of these recombination modifiers, which was assumed to be neutral in itself, appeared in a single haplotype, and was, thus, in linkage disequilibrium with a specific set of mutations, such that its fitness was exclusively dependent on the number of deleterious alleles within the segment captured. We first compared the dynamics of inversion-mimicking mutations in an autosome with those capturing a male-determining allele in an XY system (males are XY and females are XX, the male-determining allele being permanently heterozygous), and we then considered other types of recombination modifiers and heterozygosity rules.

As noted by several authors (*28*, *34*, *35*), inversions capturing fewer deleterious mutations than the population average should increase in frequency. In infinite populations, the number of mutations harbored by individuals within a genomic segment of size n follows a binomial distribution of parameters n and q, with q the mean frequency of mutations at mutation-selection equilibrium (Figure 2a and Figure S1). On average, individuals harbor nq mutations on each of their chromosomal segments of size n. Under realistic parameter values, the vast majority of large chromosomal regions therefore carry several deleterious mutations. For instance, considering s=0.001, h=0.1, u=1e-09 and n=2Mb, more than 99.999% of chromosomal fragments carry a least one mutation, the mean number of mutations being nq = 20 (Figure S1). If m is the number of recessive deleterious mutations captured by

a given inversion, the mean fitness of individuals homozygous for the inversion (W_{II}), heterozygous for the inversion (W_{NI}) or lacking the inversion (W_{NN}) in an infinite population can be easily expressed as a function of *n*, *q*, *h*, *s*, and *m* [see Methods; (*28*, *34*)]. Once formed, inversions should increase in frequency if heterozygotes for the inversion are fitter than homozygotes without the inversion ($W_{NI}>W_{NN}$), which is the case if the inversion carries fewer mutations than the population average [$m < nq$; (*28*)].

Repeated sampling from binomial distributions with a wide range of parameters showed that more than half the inversions occurring in genomes captured fewer deleterious mutations than the population average (figure 2b; hereafter referred to as "less-loaded inversions"). Indeed, the distribution of mutation number across individuals is almost symmetric when $n*q$ is high (as the binomial distribution converges to a normal distribution; e.g. Figure S1), but, when $n*q$ is low, the distribution is zero-inflated, increasing the probability of inversions being less-loaded. Therefore, a substantial fraction of inversions occurring in genomes are beneficial when they form (i.e. when rare enough not to occur as homozygotes). For instance, with *h* values ranging from 0 to 0.5 and *s* values ranging from 0.001 to 0.25, between 36% and 98% (mean=70%) of the 2 Mb inversions occurring in the genome are beneficial, carrying fewer recessive deleterious mutations than average. Simulations in finite populations of different sizes confirmed that most inversions (mean=66% in the range of parameter values studied) had a fitness advantage upon formation. The simulations also showed that inversions could be favored if they captured mutations that were rarer than average (Figures S2-3).

The deterministic increase in frequency of an inversion (I) on an autosome or capturing an XY-like sex-determining locus can easily be determined with a two-locus two-allele model. For instance, the change in frequency of an inversion on the proto-Y chromosome can be expressed as:

$$\Delta F_{YI} = \frac{F_{YI}(F_{XIf}W_{II} + (1 - F_{XIf})W_{NI} - \overline{W_m}) \pm W_{NI}D}{\overline{W_m}},$$

where F_{XIf} is the frequency of the inversion on the proto-X chromosome in females, *r* is the rate of recombination between the inversion and the sex-determining locus, *D* is their linkage disequilibrium and $\overline{W_m}$ is mean male fitness (Figure 2c ; see Methods and Appendix for details). Based on this model, and initially assuming that inverted and non-inverted segments no longer accumulate deleterious mutations after their formation (i.e. W_{II}, W_{NI} and W_{NN} are fixed parameters), we simulated the evolutionary trajectory of inversions on autosomes and of inversions capturing the male-determining allele on the Y chromosome under a wide range of parameter values. We found that the frequency of less-loaded inversions tended to remain low in autosomes, whereas these inversions became fixed in the population of Y chromosomes (Figure 2c), as expected according to our hypothesis.

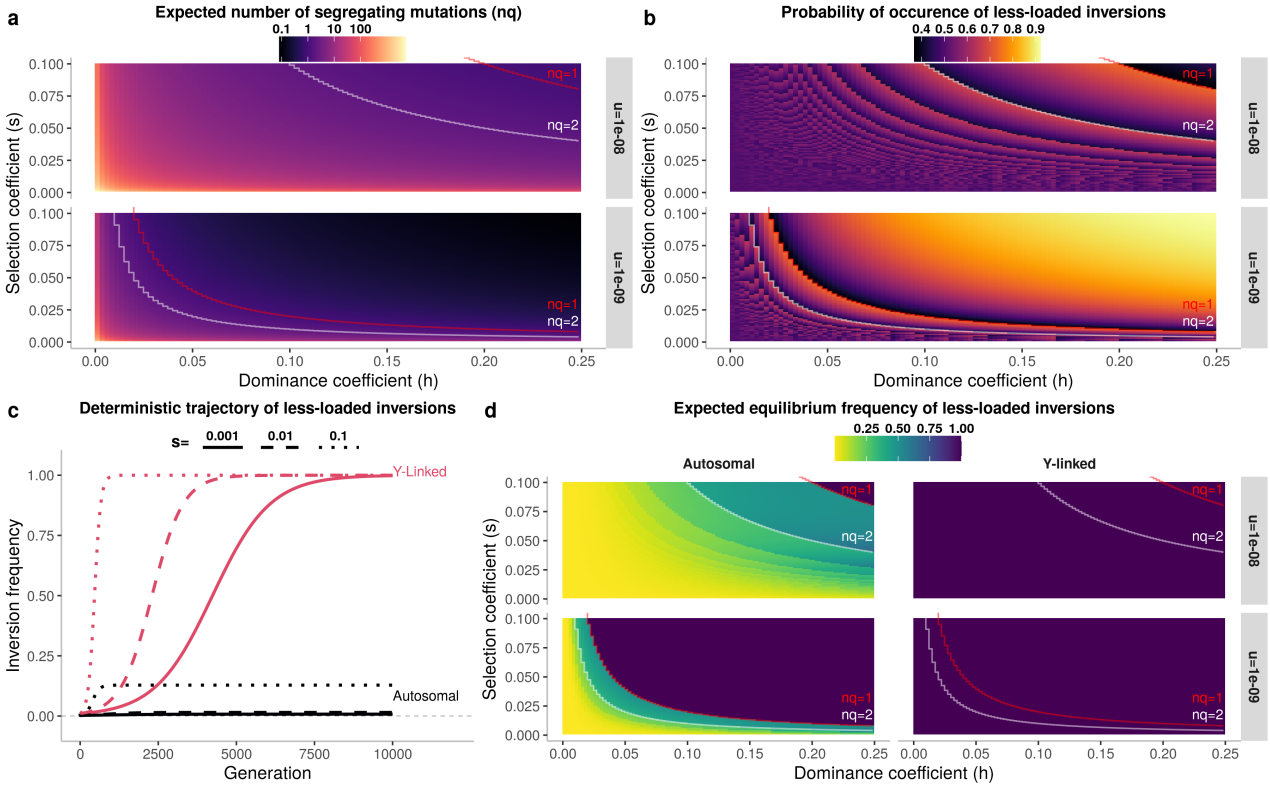

Figure 2 | In infinite populations, all less-loaded chromosomal inversions become fixed on the Y sex chromosome, whereas most remain at a low frequency on the autosome

a, Expected number of segregating recessive deleterious mutations within a segment of *n*=2 Mb as a function of the dominance coefficient (*h*), selection coefficient (*s*) and mutation rate (*u*). Mutations were considered to be at mutation-selection equilibrium frequency (see methods). Parameters resulting in one and two expected mutations are highlighted by red and white lines, respectively, and are also shown on panels b and d. **b**, Probability of occurrence of a less-loaded inversion covering a segment of *n*=2 Mb as a function of the dominance coefficient (*h*), selection coefficient (*s*) and mutation rate (*u*). Less-loaded inversions are defined as those for which $m<nq$, where *m* is the number of mutations in the inversion (i.e., fewer than the average for the population). The number of mutations in inversions can take only integer values (*m*=0, 1, 2, etc.), and the transition between these values affects the probability of occurrence of less-loaded inversions, yielding the non-continuous graphs in panels b and d. **c**, Deterministic change in inversion frequency, for the case of 2 Mb inversions capturing the male-determining allele on the Y-chromosome or inversions on an autosome. For clarity, the frequency displayed for Y-linked inversions is the frequency of inversions in the population of Y chromosomes. The figure illustrates the case of inversions carrying a number of mutations 5% lower than the population average (m=[0.95*nq]), mutations being at their mutation-selection equilibrium frequency with $u=10^{-8}$ and *h*=0.1. Figures S4-8 illustrate the fate of inversions with intermediate degrees of linkage to the male-determining allele, with various numbers of mutations relative to the population average, or inversions linked to a locus with variable heterozygosity rules or in the X chromosome population. **d**, Expected equilibrium frequency of less-loaded 2 Mb inversions (i.e. with *m*=0,1 ,2, …, nq) capturing the male-determining allele, or on an autosome, weighted by their probability of occurrence (Methods, equation 5, 6, and 7). This panel extends the result from panel c for a range of dominance coefficients (*h*), selection coefficients (*s*) and mutation rates (*u*). As above, the frequency displayed for Y-linked inversions is the frequency of inversions on the population of Y chromosomes. Above the line nq=1, the mean number of segregating mutations in a given inversion (of

length n) is less than 1, indicating that all less-loaded inversions are mutation-free. Such inversions can also become fixed on autosomes, leading to an expected equilibrium frequency of 1 for both autosomes and the Y chromosome. Figure S9 shows the overall frequencies of all inversions, whether less or more loaded than average, and over a larger range of parameters.

We therefore expressed the equilibrium frequency of inversions ($F_{equ,m}$, Figure 2d) as a function of m, the number of deleterious mutations carried by the inversion, to calculate the threshold values at which autosomal and Y-linked inversions could be maintained or fixed (see appendix). Assuming that q<<1 and s<<1, we found that inversions on autosomes become fixed when $m<qhn/(1-h)$, whereas they should stabilize at a frequency of $\frac{W_{NI}-W_{NN}}{(2W_{NI}-W_{II}-W_{NN}}$, the equilibrium frequency of an overdominant locus, when $qhn/(1-h)<m<nq$ (see Methods). Inversions on autosomes capturing more than $qhn/(1-h)$ recessive deleterious mutations do, indeed, suffer from a homozygote disadvantage (WNI > WII, figure S1), preventing them from reaching high frequencies (Figure 2c). This is the case for most autosomal inversions under realistic parameter values (e.g. $m>qhn/(1-h)$ for more than 99.99 % of inversions when n=2 Mb, u=1e-09, h=0.1 and s=0.001, Figure S1). By contrast, permanently heterozygous alleles protect inversions from this homozygote disadvantage, allowing these inversions to become fixed. All inversions capturing the male-determining allele become, therefore, fixed if they carry fewer than average mutations (i.e. $m<nq$). Thus, contrary to the argument proposed in a previous study that only mutation-free inversions can become fixed (*35*), we found that inversions can carry deleterious mutations and nevertheless become fixed in the population, provided that they carry fewer than $qhn/(1-h)$ mutations if they are located on autosomes and fewer than nq mutations if they capture a permanently heterozygous allele (on the Y chromosome, for example).

The expected equilibrium frequency of all inversions occurring in a given genomic region can be expressed as :

$$[F_{equ}] = \sum_{m=0}^{n} P_m \times F_{equ,m}$$

with P_m being the probability of occurrence of inversions with m mutations and $F_{equ,m}$ the equilibrium frequency of inversions with m mutations (Figure 2d; see Methods). For permanently heterozygous inversions, we, thus, have:

$$[F_{equ}] = \sum_{m=0}^{\lfloor nq \rfloor} \binom{m}{n} q^m (1-q)^{n-m},$$

and for autosomal inversions:,

$$[F_{equ}] = \sum_{m=0}^{\lfloor nq \frac{h}{1-h} \rfloor} \binom{m}{n} q^m (1-q)^{n-m} + \sum_{m=\lfloor nq \frac{h}{1-h} \rfloor + 1}^{\lfloor nq \rfloor} \binom{m}{n} q^m (1-q)^{n-m} \frac{W_{NI}-W_{NN}}{2W_{NI}-W_{NN}-W_{II}}.$$

With a realistic range of parameters, inversions are much more likely to become fixed if they capture the male-determining allele on the Y chromosome than if they are unlinked to this allele (e.g. on an autosome; Figure 2c-d and Figures S4-9). For instance, with u=1e-09, h=0.1, s=0.001, and n=2 Mb, 47% of inversions occurring on the Y chromosome would be expected to become fixed, versus only 0.000045% of inversions on autosomes.

If deleterious mutations continue to arise after the formation of the inversion, the dynamics of the inversion become harder to predict deterministically. In infinite populations, the accumulation of

mutations after the formation of the inversion can be approximated by $p_t = \frac{u}{sh} + (p_0 - \frac{u}{sh})e^{-sht}$ where P_t is the number of mutations in the inversion at time *t* (*28*, *34*). This assumes that inverted segments evolve under mutation-selection dynamics with recombination. However, such assumptions do not hold in many situations. Indeed, low-frequency inversions in finite populations should almost never occur as homozygotes and should therefore evolve with almost no recombination. This is also true for inversions capturing a permanently heterozygous allele, which never undergoes recombination. We confirmed, by simulation, that the above approximations for an infinite population depart strongly from the situation observed in finite populations (e.g. Figure S13). In finite populations, low-frequency or permanently heterozygous inversions tend to evolve under a Muller's ratchet-like dynamic, with the mean fitness of inversions decreasing in a stepwise manner due to the sequential loss, by drift, of the inverted haplotypes with the lowest mutational load. Following their formation, autosomal and Y-linked inversions tend to accumulate more mutations than the average for the population, contrasting with predictions for infinite populations (Figure S13). Only inverted segments reaching relatively high frequencies in autosomes eventually recombine when homozygous, and their dynamics of mutation accumulation therefore involve a mixture of a Muller's ratchet-like regime (when rare, at the start of their spread) and a mutation-selection-drift regime with recombination (when they reach intermediate frequencies). Little is known about the transition between these regimes (*36*, *37*). We therefore used individual-based simulations to study the fate of inversions accumulating deleterious mutations. This also made it possible to relax the previous assumption (*35*) that the time to inversion fixation is much shorter than the time taken for the inversions to accumulate deleterious mutations. Individual-based simulations also made it possible to relax the assumption that all genomic positions are independent.

Simulations of the spread of inversions in finite populations with recurrent mutation confirmed the tendency identified in the deterministic model without mutation accumulation: over most of the parameter space, inversions are much more likely to spread if they capture the sex-determining allele on the Y chromosome than if they are located on autosomes (Figure 3 and Figures S10-16). Many autosomal inversions carrying a mutation load segregated for hundreds of generations. For example, with N=1000, s=0.01, h=0.1 and u=1e^{-08}, 73 of 10,000 inversions of 2 Mb in length continued to segregate after 500 generations. However, all these autosomal inversions were lost at the end of the simulations (i.e. after 10,000 generations, Figure 3a). These inversions initially spread because, by chance, they had a lower-than-average mutation load, but their homozygote disadvantage prevented them from reaching fixation. They were eventually lost because they accumulated further mutations relative to non-inverted segments (Figure 3b and S13).

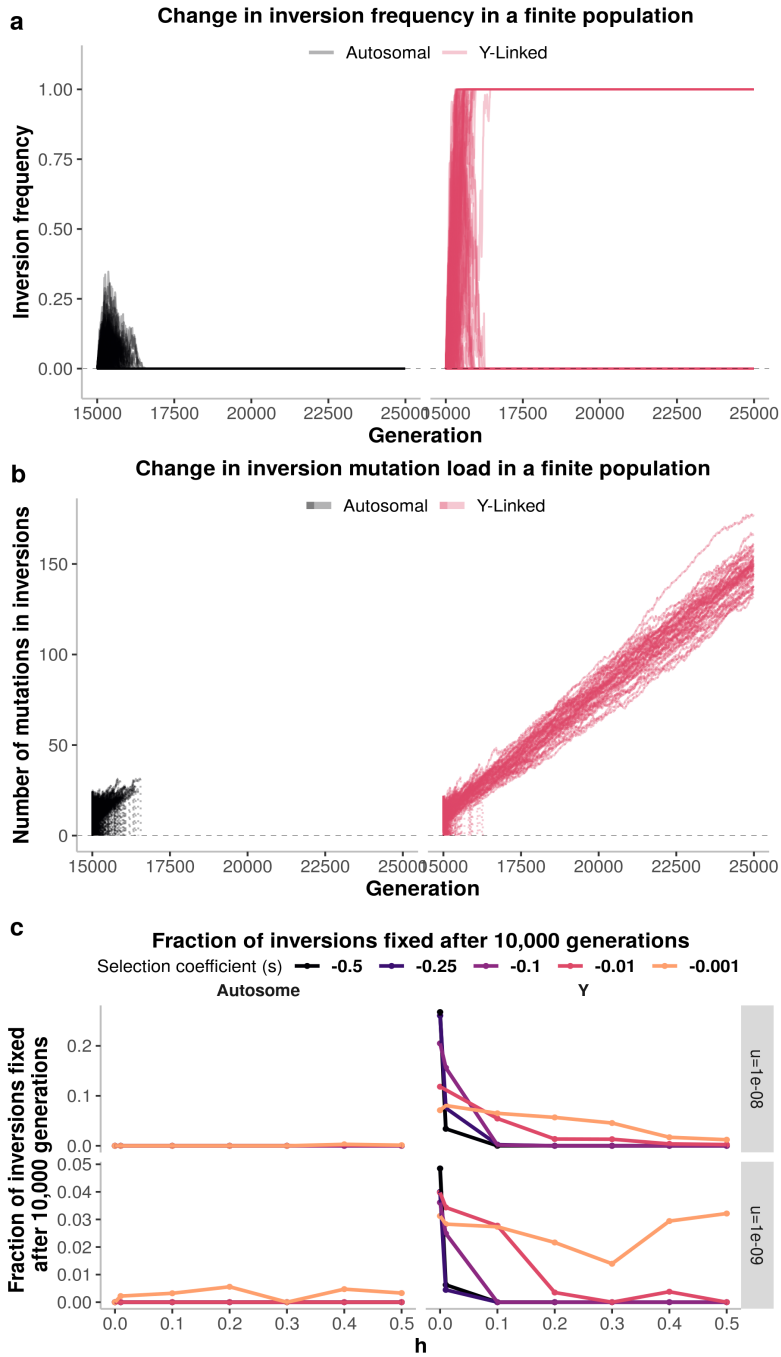

Figure 3 | In finite populations, a substantial fraction of less-loaded chromosomal inversions on the Y sex chromosome become fixed, whereas all less-loaded inversion on autosomes are lost

a, Change in inversion frequency in stochastic simulations of 1000 individuals experiencing recessive deleterious mutations at a rate u=1e-08, all mutations having the same dominance and selection coefficients (here, *h*=0.1 and *s*=0.01). The figure displays the frequency of 10,000 independent inversions on each of an autosome and a proto-Y chromosome, with each line representing a specific inversion (i.e. a simulation). The evolutionary trajectories of inversions of different sizes, in larger populations, on the X-chromosome or with other parameter combinations are displayed in figures S10-14. **b**, Change in the mean number of mutations carried by the inverted segments in each of the 10,000 simulations. Each line represents one simulation. A zoom-view of the autosomal inversion dynamics is shown in Figure S13. **c**, Proportions, in stochastic simulations, of 2 Mb inversions that were fixed after 10,000 generations, for different parameter value combinations. This extends the result from panel **a** to other values of dominance

coefficient (*h*), selection coefficient (*s*) and mutation rate (*u*). For each parameter combination, we simulated 10,000 inversions of 2 Mb capturing a random genomic fragment. Only inversions not lost after 20 generations are considered here. Results for inversions of different sizes are shown in Figure S15 and results for different population sizes are shown in Figure S16.

By contrast, substantial fractions of less-loaded inversions capturing the permanently heterozygous sex-determining allele on the Y chromosome spread until they became fixed in the Y chromosome population; this was the case even for inversions that were not mutation-free (Figure 3 and Figure S10-16). For example, for s=0.01, h=0.1, u=1e^{-08} and N=1000, 49 of the 10,000 Y-linked inversions (2 Mb) became fixed in the Y chromosome population, whereas all inversions on the autosome were lost. New mutations occurred on Y-linked inversions, but they did not accumulate rapidly enough to prevent these 49 inversions from spreading and reaching fixation (Figure 3a-b and Figure S13). Permanent heterozygosity effectively results in directional selection for the less-loaded inversions, and not only for those free from mutations (Figure 3), leading to rapid fixation before the accumulation of too many new deleterious mutations.

Similar results were obtained when: i) two or more permanently heterozygous alleles segregated at a mating incompatibility locus, modeling plant self-incompatibility or fungal mating-type systems [Figure S5 and S14; (*5*, *38*)]; ii) the alleles were not permanently heterozygous, but were strongly overdominant, thus occurring mostly in the heterozygous state, as for several supergenes controlling color polymorphism [Figure S8 ; (*13*, *39*, *40*)]; and iii) the inversion was in strong but incomplete linkage with the permanently heterozygous allele (e.g. 0.1 cM away from the allele, Figure S4-8). Our model can thus account for the existence of inversions very close, but not fully linked to a mating-type locus, as reported in the chestnut blast fungus (*41*, *42*). Similar results were also obtained when we considered recombination modifiers suppressing recombination even when homozygous, as for histone modifications or methylation (*43*), rather than solely when heterozygous, as for inversions (Figure S14). Similar results were also observed if we assumed that the fitness effects of mutations occurring across the genome were drawn from a gamma distribution (i.e. the mutations segregating in populations had different fitness effects; Figure S17).

We, thus, found that neutral recombination modifiers spread in a very large range of conditions, if they captured permanently heterozygous alleles (Figure 3C and Figure S15-16). However, as expected, for the inversions to benefit from a heterozygote advantage by associative overdominance, deleterious mutations segregating in the genome had to be partially or completely recessive ($h<0.5$; Figure 3c and Figures S15-16). Furthermore, as for any variant, even beneficial inversions benefiting from a selective advantage could be lost by drift, the probability of such loss depending on population size and the relative fitness of the inversion, in terms of the number and type of mutations initially captured relative to the average for the population (Figures 3 and S15-16). Inversions occurring in regions in which large numbers of mutations segregate are more likely than those in mutation-poor regions to capture many fewer mutations than average, and therefore to have a stronger relative advantage. In other words, inversions in mutation-dense regions have a wider fitness distribution. The probability of Y-linked inversion fixation, thus, increases with increasing inversion size, mutation recessiveness and mutation rate (Figure 3c and Figures S15-16). As expected (*44*), we found that, with increasing population size, beneficial inversions took more generations to spread to fixation (Figure S11-12). This longer time favored the further accumulation of deleterious mutations in inversions,

lowering their fitness and, in some cases, preventing their fixation. The probability of inversion fixation in the Y chromosome population was, therefore, inversely correlated with population size, but was always much higher than that of fixation on autosomes (Figure S15-16).

Inversions would, therefore, be expected to accumulate successively around permanently heterozygous alleles across evolutionary time, under a large range of conditions, leading, for example, to the formation of non-recombining sex chromosomes with a typical pattern of evolutionary strata. We studied the formation of such strata by the sequential occurrence of multiple inversions, by simulating the evolution of large chromosomes experiencing the occurrence of multiple chromosomal inversions that can overlap with each other, under parameter values typical of those observed in mammals (Figures 4, 5 and S19). We simulated, over 100,000 generations, populations of N=10,000 or N=1,000 individuals, carrying two 100 Mb chromosomes, one of which harbored a mammalian-type sex-determining locus (XY males and XX females). Individuals experienced only deleterious or nearly-neutral mutations, with mutation rates and fitness effects similar to those observed in humans (fitness effect being drawn from a gamma distribution). The dominance coefficient of each mutation was chosen uniformly at random from a wide set of values (see Methods for details). At the start of each simulation, we randomly sampled *k* genomic positions that could be used as inversion breakpoints, with *k* being 10, 100, 1000 or 10,000. These genomic positions represent inversion hotspots, such as those that can be generated by repeats in genomes (*45*). In each generation, we introduced *j* inversions, *j* being sampled from a Poisson distribution. Simulation times were limited by using inversion rates such that one inversion, on average, occurred in each generation, throughout the whole population. The two breakpoints of each inversion were chosen, at random, from the k positions. It was, therefore, possible for two independent inversions to appear at the same position, allowing, in particular, the reversion of an inversion to its ancestral orientation, thereby restoring recombination. Reversion may be favored, for example, by an inversion accumulating enough deleterious mutations after its fixation on the Y chromosome to carry a larger number of mutations than the population average (*46*). We assumed that inversions partially overlapped by another inversion or that captured a smaller inversion could not be reversed, i.e., recombination could not be restored in such situations even if subsequent inversions re-used the same breakpoints. Indeed, reversions of partially overlapped inversions do not restore ancestral arrangements but instead result in complex reshufflings of gene order and orientation and are therefore unlikely to restore recombination. We ran 10 simulations for each set of parameters.

In all simulations assuming relatively large numbers of potential inversion breakpoints (k=100, 1000, 10,000), the Y chromosome progressively stopped recombining with the X chromosome as it accumulated inversions fully linked to the male sex-determining allele (Figures 4, 5 and S19). After the occurrence of an initial inversion capturing the male sex-determining allele, multiple inversions partially overlapping this inversion or other Y-fixed inversions were selected, thereby generating a growing chaos of overlapping chromosomal rearrangements. The non-recombining region, thus, extended around the sex-determining locus in a stepwise manner, perfectly reflecting the evolution of sex chromosomes and other supergenes with evolutionary strata (Figure 4). Some events in the gradual extension of recombination suppression were reversed, due to the occasional occurrence of beneficial reversions (Figures 4 and 5). The accumulation of overlapping inversions was, however, more rapid than the occurrence of beneficial reversions, leading to a progressive extension of the non-recombining region (Figure 4 and 5). By contrast, when we assumed a smaller number of potential breakpoints (*k* =

10), recombination suppression between sex chromosomes evolved in only two of the 10 simulations, and across only a small genomic region (Figure 5), owing to the more frequent occurrence of beneficial reversions.

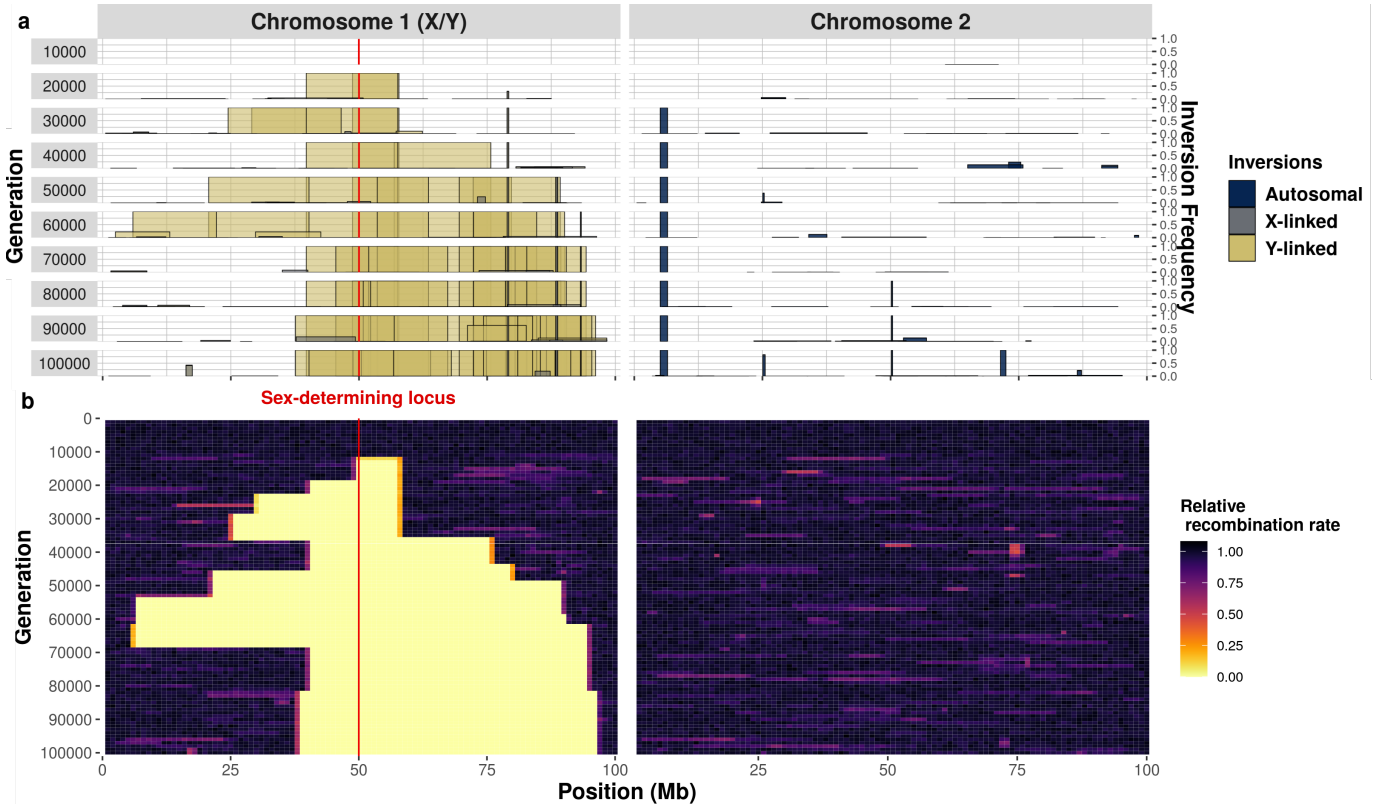

Figure 4 | Successive accumulation of inversions around a male-determining allele in an XY system, leading to the formation of non-recombining sex chromosomes

A simulation of *N*=1000 individuals, each with two pairs of 100 Mb chromosomes, over 100,000 generations. Chromosome 1 harbors an X/Y sex-determining locus at 50 Mb (individuals are XX or XY). In each generation, one inversion appears, on average, in the whole population, in an individual sampled uniformly at random, with the two recombination breakpoints sampled uniformly at random from *k*=100 potential breakpoints. **a,** Overview of chromosomal inversion frequency and position for 10 different generations. The width of the box represents the position of the inversion and the height of the box indicates inversion frequency. Inversions appearing on the Y chromosome are depicted in yellow, those appearing on the X chromosomes are depicted in gray. The colors are not entirely opaque, so that regions with overlapping inversions appear darker. Previously fixed inversions may be lost due to the occurrence of beneficial reversions and selection. **b,** Changes in the relative rate of recombination over the entire course of the simulation. The numbers of recombination events occurring at each position (binned in 1 Mb windows) are recorded at the formation of each offspring, across all homologous chromosomes in the population. Only recombination events between the X and the Y chromosomes are shown for chromosome 1 (i.e., recombination events between the two X chromosomes in females are not shown). Unlike chromosome 1, chromosome 2 harbors no permanently heterozygous alleles. All inversions on this chromosome suffer from homozygote disadvantage and very few inversions therefore become fixed on chromosome 2. See Figure S19 for a simulation with *N*=10,000 individuals.

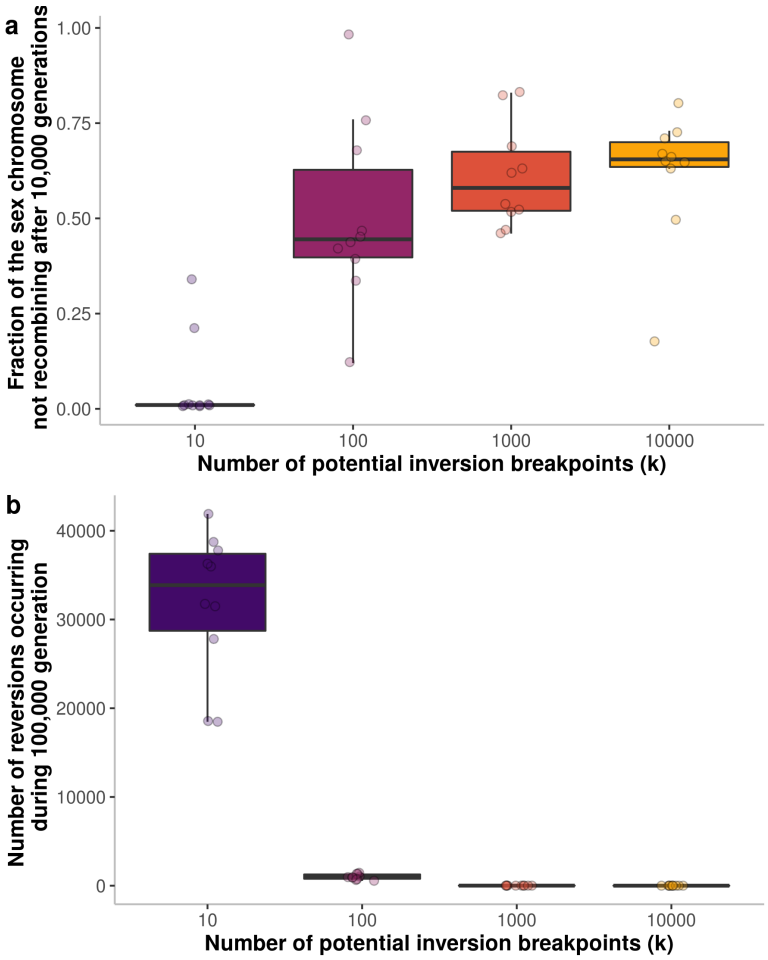

Figure 5 | Effect of the number of potential inversion breakpoints on the evolution of recombination suppression.

For each number of breakpoints, 10 simulations were conducted. Each dot represents the result of a simulation with *N*=1000. **a,** Fraction of the length of the Y sex chromosome not recombining after 100,000 generations**. b,** Number of reversions occurring over the course of the 100,000 generations. Boxplot elements: central line: median, box limits: 25th and 75th percentiles, whiskers: 1.5x interquartile range.

Discussion

Our results show that recombination suppression on sex chromosomes and other supergenes can evolve simply because genomes harbor many partially recessive, deleterious variants. Our model for the evolution of sex chromosomes, and supergenes in general, is based on simple and widespread phenomena: i) inversions (or any recombination suppressor) can be favored solely because they contain fewer deleterious mutations than the population average, a situation applying to a substantial fraction of the inversions formed; ii) such inversions tend to display overdominance: they are beneficial in the heterozygous state but suffer from a homozygote disadvantage, which prevents them from reaching high frequencies and becoming fixed on autosomes; iii) when, by chance, inversions capture a permanently heterozygous allele, they do not suffer from this homozygote disadvantage and are therefore able to spread until they are fully associated with the permanently heterozygous allele (e.g. they become fixed in the Y chromosome population). These three phenomena have been reported independently in several studies, but, to our knowledge, never in interaction [see references (*28*, *35*) for i, (*29*, *47*) for ii, and (*5*, *19*) for iii]. The combined influence of the mechanisms related to ii and iii has been shown to promote sex chromosome-autosome fusion in highly inbred populations (*20*). We show here that such mechanisms can readily lead to the stepwise extension of the non-recombining region on sex chromosomes themselves, without the need for inbreeding. Moreover, unlike previous studies [e.g. (*18*, *35*, *48*)], we show that the higher probability of inversion fixation on Y chromosomes is not restricted to mutation-free inversions, but applies to any inversion capturing fewer deleterious mutations than the average.

We did not include possible decreases in fertility due to segregation issues arising during meiosis as a result of inversions in our model (*49*); such issues would, clearly, decrease the probability of inversion fixation as a function of the fitness cost, as for any other model explaining the evolution of the stepwise extension of recombination suppression involving inversions. However, our model holds for any mechanism of recombination suppression (it is not limited to inversions) and the other possible mechanisms [such as methylation (*43*)] would not be expected to decrease fertility. Given the lack of knowledge about inversion rates in natural conditions and the computation challenge represented by the simulation of complex patterns of recombination, we used reasonably high inversion rates in our sex-chromosome evolution simulations, making it possible to observe the stepwise extension of a non-recombining region within 100,000 generations. A recent study suggested that inversions may not be too rare and that inversion breakpoints may be widely distributed throughout the genome (*45*). The use of different inversion rates might result in much shorter or much longer times for the stepwise extension of the non-recombining region, but should not change the final outcome. In nature, the stepwise extension of the non-recombining region between sex or mating-type chromosomes often occurs over time scales of the order of tens of millions of generations [e.g. about 250 M years in human (*50*)], suggesting that the fixation of inversions may be a relatively rare event.

On autosomes, inversions maintained at low frequencies because of their homozygote disadvantage tend to be lost rapidly because they accumulate further deleterious mutations. On the Y chromosome, contrary to previous suggestions (*28*), we show that the accumulation of further deleterious mutations following the formation of an inversion is generally too slow to prevent the fixation of less-loaded inversions. Deleterious mutations nevertheless accumulate following fixation of the inversion, leading to Y-linked inversion degeneration, as observed on many non-recombining sex

chromosomes (*30*). This may lead to selection for reversion of the inversion, thereby restoring recombination (*46*). The question as to whether this can actually occur is common to all mechanisms explaining evolutionary strata, including sexual antagonism.

We found inversion reversion to be rare, unless there were no more than 10 inversion breakpoints over the two 100 Mb chromosomes. Indeed, for reversions to invade the population and restore recombination, the following conditions must be met: i) sufficient deleterious mutations must accumulate in the initial inversion to render reversion beneficial and ii) a partially overlapping inversion should not have invaded the population before the occurrence of a beneficial reversion. Indeed, partially overlapping inversions result in a complex reshuffling of gene order and orientation, preventing the restoration of recombination even in situations in which reversion could be selected. When the number of potential breakpoints is relatively high, the probability of partially overlapping inversions occurring is higher than the probability of a reversion occurring, regardless of the relative rates of inversions and deleterious mutations.

The reversion of inversions has been reported only in very rare cases in which inversions occur at specific breakpoints rich in repeated elements (*51*, *52*), as in our simulations in which only a few possible breakpoints were present in the genome. The number of genomic positions at which inversions can occur in natural conditions is unknown, but this number is likely to be high given the chaos of rearrangements observed in some sex and mating-type chromosomes with hundreds of different breakpoints (*16*, *45*, *53*, *54*). A recent study reported the recurrent appearance of inversions at the same positions, but also the existence of numerous potential breakpoints of inversions in human genomes (*45*). Moreover, chromosomal rearrangements rapidly accumulate in recently established regions of non-recombination (*53*) and overlapping inversions are observed on many sex chromosomes, which should prevent the restoration of recombination by reversions (*16*, *55–58*). Therefore, over a wide range of realistic parameter values, the reversion of inversions should not prevent the stepwise formation of non-recombining sex chromosomes. When there are few potential inversion breakpoints in the genome, the mechanism stabilizing inversions through dosage compensation considered in a recent model (*46*) may act in addition to the mechanism of sex-chromosome evolution studied here, as these two mechanisms are not mutually exclusive. However, by contrast to the framework of our model, dosage compensation mechanisms require XY-like asymmetry and cannot, therefore, explain evolutionary strata on fungal mating-type chromosomes or bryophyte UV sex chromosomes (*5*, *54*).

The theory proposed here applies to any locus with at least one permanently or nearly permanently heterozygous allele. It requires only that, within a population, individuals carry different numbers of partially recessive deleterious mutations in a genomic region that can be subjected to recombination suppression. This situation is probably frequent in diploid, dikaryotic or heterokaryotic organisms. For instance, the mean distance between two heterozygous positions is ~1000 bp in primates (*59*), and it has been estimated that most of mutations are probably partially recessive and deleterious (*22*, *60*). Chromosomal rearrangements spanning hundreds of kilobases are frequently observed (*45*, *61*, *62*), and would, therefore, be expected to capture several deleterious variants.

We show that inversions are more likely to spread in regions harboring many segregating deleterious mutations, in which they have a greater chance of capturing highly advantageous haplotypes. Our model therefore predicts that species harboring a large number of deleterious recessive variants, due to their mating system or high mutation rate, for example, will be more prone to the

evolution of large non-recombining regions with evolutionary strata on sex chromosomes than species with a low mutation load. Furthermore, in species with large population sizes, the time required for an inversion to become fixed may exceed the time for which the number of deleterious mutations in the inversion remains below average. This would prevent some inversions from becoming fixed, potentially lowering the rate of expansion of non-recombining regions on sex chromosomes in species with large population sizes. Variations in population size, mutation rates and mating systems across lineages may, therefore, account for the large number of different sex-chromosome structures in nature, with some organisms maintaining homomorphic sex chromosomes and others evolving highly differentiated sex chromosomes with multiple strata (*1*). In plants, for example, selection at the haploid stage seems to result in the efficient purging of recessive deleterious mutations (*63*), potentially accounting for the smaller non-recombining region observed on the sex chromosomes of plants and algae than in animals (*64*, *65*). In fungi, evolutionary strata around mating-type genes have been reported only in species with an extended dikaryotic stage (*5*). This is consistent with our model, in which recombination suppression evolves through the sheltering of deleterious mutations, which cannot occur in haploid organisms.

In conclusion, given its simplicity and wide scope of application, our model of sex chromosome evolution is a powerful alternative to other explanations (*6*), although the various theories are not mutually exclusive. The strength of our model lies in the absence of strong assumptions, such as sexually antagonistic selection, XY asymmetry or small population sizes. Our model, based on the often overlooked observation that recessive deleterious mutations are widespread in genomes within natural populations, can explain the evolution of stepwise recombination suppression over a wide range of realistic parameter values. Furthermore, it can also explain why some supergenes, such as those in butterflies and ants, display evolutionary strata (*13*, *14*), and why many fungal mating-type loci display a stepwise extension of non-recombining regions despite the absence of sexually antagonistic selection or XY-like asymmetry in these organisms (*5*, *16*). Our model therefore provides a general and simple framework for understanding the evolution of non-recombining regions around loci carrying permanently heterozygous alleles.

Materials and Methods

Infinite population deterministic model

We consider the discrete-time evolution of an infinite size, randomly mating population experiencing only deleterious recessive mutations, with heterozygotes and homozygotes suffering from a *1-hs* and *1-s* reduction in fitness, respectively. At all *n* sites, mutations are at the same mutation-selection equilibrium frequency, denoted *q*. We used non-approximated values of *q* as derived by (*66*):

$$q = \frac{h(1+u)}{2(2h-1)}\left[1-\sqrt{1-\frac{4(2h-1)u}{sh^2(1+u)^2}}\right]$$

We assume that the sites are independent. The number of mutations carried by a chromosomal segment of length n then follows a binomial distribution of parameters (*n*, *q*). We follow the frequency of an inversion I of size *n*, considering that it captures *m* mutations and that it appears in a population carrying only non-inverted segments. The mean fitness of a non-inverted homozygote can be computed as follows (see Appendix, section 2):

$$\overline{W_{NN}} = (1-2q(1-q)hs-q^2s)^n.$$

Note that in the parameter regimes where we can make the approximation $q \approx u/hs$ with $q << 1$, we have $\overline{W_{NN}} \approx (1-2u)^n$, in accordance with mutation load theory (*34*, *67*). Similarly, an individual who is heterozygous for an inversion with *m* mutations has a mean fitness of:

$$\overline{W_{NI}} = (q(1-s)+(1-q)(1-hs))^m(q(1-hs)+1-q)^{n-m}$$

Assuming that $q << 1$, Nei et al., 1967 (*28*), considered that individuals heterozygous for an inversion had no mutations homozygous, such as $\overline{W_{NI}} \approx (1-hs)^m(1-hs)^{nq} \approx (1-hs)^{nq+m}$. Note that we use non-approximated values in computations. An individual who is homozygous for a segment I with *m* mutations is homozygous for all these mutations. Its fitness can therefore be expressed as:

$$W_{II} = (1-s)^m.$$

The inversion frequency trajectory can be determined with a simple two-locus two-allele model. We considered four different situations, depending on the possible heterozygosity at the locus with permanently heterozygous alleles. The Appendix, sections 4-8, presents the evolution in time of the frequency of the inversion in detail in these four situations. Here, we briefly describe the results for inversions more or less linked to a permanently heterozygous allele in a XY system. The change in frequency of inversions on the Y chromosome, on the X chromosome or on autosomes between generations t and $t+1$ (Figure 2c) is described by

$$
\begin{aligned}
F_{XIm}^{t+1} &= \frac{F_{XIf}^{t}\left(W_{II}F_{YI}^{t} + W_{NI}F_{YN}^{t}\right) + rW_{NI}D^{t}}{\overline{W_m^t}},\\
F_{XNm}^{t+1} &= \frac{F_{XNf}^{t}\left(W_{NN}F_{YN}^{t} + W_{NI}F_{YI}^{t}\right) - rW_{NI}D^{t}}{\overline{W_m^t}},\\
F_{XIf}^{t+1} &= \frac{F_{XIm}^{t}F_{XIf}^{t}W_{II} + \frac{1}{2}W_{NI}[F_{XIm}^{t}F_{XNf}^{t} + F_{XIf}^{t}F_{XNm}^{t}]}{\overline{W_f^t}},\\
F_{XNf}^{t+1} &= \frac{F_{XNm}^{t}F_{XNf}^{t}W_{NN} + \frac{1}{2}W_{NI}[F_{XNm}^{t}F_{XIf}^{t} + F_{XNf}^{t}F_{XIm}^{t}]}{\overline{W_f^t}},\\
F_{YI}^{t+1} &= \frac{F_{YI}^{t}(F_{XIf}^{t}W_{II} + F_{XNf}^{t}W_{NI}) - rW_{NI}D^{t}}{\overline{W_m^t}},\\
F_{YN}^{t+1} &= \frac{F_{YN}^{t}(F_{XNf}^{t}W_{NN} + F_{XIf}^{t}W_{NI}) + rW_{NI}D^{t}}{\overline{W_m^t}},\\
F_{XI} &= \frac{2}{3}F_{XIf} + \frac{1}{3}F_{XIm},\\
F_{I} &= \frac{1}{4}F_{YI} + \frac{3}{4}F_{XI}
\end{aligned}
\qquad \text{(Equation 1)}
$$

where F_{YI}^{t} is the frequency of inversions in the population of Y chromosomes at time *t*, F_{XIf}^{t} is the frequency of the inversion on the X chromosome in females (respectively F_{XIm}^{t} in males), *r* is the rate of recombination between the inversion and the sex-determining locus, *D* is their linkage disequilibrium and $\overline{W_m^t}$ is the mean male fitness (Figure 2c ; see Appendix, section 8 for details). When *r*=0.5, this system of equations describes the evolution of inversions on autosomes. Unless stated otherwise, the deterministic simulations presented here (Figures 2c and S4-8) were performed with an initial *D*=-0.01 or *D*=0.01, depending on the allele which the inversion appeared linked to. Results for various initial linkage disequilibrium values are presented in Figure S6. When the inversion captures the male-determining allele, *r*=0, $F_{XIm} = F_{XIf} = 0$, and $F_{XNf} = F_{XNm} = 1$ at any time *t*. The equations for F_{YI}^{t}, the frequency of inversions capturing the male-determining allele then reduces to

$$\mathrm{F}_{YI}^{t+1} = F_{YI}^{t}\frac{W_{NI}}{\overline{W_m^t}},$$

$$\mathrm{F}_{YN}^{t+1} = F_{YN}^{t}\frac{W_{NN}}{\overline{W_m^t}}.$$

with $\overline{W_m^t} = F_{YN}^{t}W_{NN} + F_{XI}^{t}W_{NI}$. Substituting F_{YN} by $1 - F_{YI}$ and withdrawing F_{YI}^{t}, we have

$$\Delta F_{YI} = \frac{F_{YI}^{t}(1 - F_{YI}^{t})\left(W_{NI} - W_{NN}\right)}{\overline{W_m^t}}$$

The search for F_{YI} such that $\Delta F_{YI} = 0$ readily gives two equilibria : 0 and 1. Since $\overline{W_m^t} > 0$ and $0 \leq F_{YI} \leq 1$, we can conclude that:

$$\boxed{\begin{array}{ll} \text{If } W_{NI} > W_{NN}, & F_{YI} \longrightarrow 1,\\ \text{If } W_{NI} < W_{NN}, & F_{YI} \longrightarrow 0 \end{array}}$$

In the case of an inversion appearing on an autosome (*r*=0.5), $F_{XIm} = F_{XIf} = F_{YI}$ and $\overline{W_m} = \overline{W_f}$, so that the change in inversion frequency in the population is:

$$F_I^{t+1} = \frac{(F_I^t)^2 W_{II} + F_I^t F_N^t W_{NI}}{\overline{W^t}},$$

and

$$\Delta F_I = \frac{F_I}{\overline{W}}\Big(F_I^2\left(2W_{NI} - W_{II} - W_{NN}\right) + F_I\left(W_{II} + 2W_{NN} - 3W_{NI}\right) + W_{NI} - W_{NN}\Big) = \frac{F_I}{\overline{W}} \times P(F_I),$$

with $P(X) = X^2\left(2W_{NI} - W_{II} - W_{NN}\right) + X\left(W_{II} + 2W_{NN} - 3W_{NI}\right) + W_{NI} - W_{NN}$.
The equilibria are 0 and the roots of the polynomial P, which are 1 and $\frac{(W_{NI}-W_{NN})}{(2W_{NI}-W_{II}-W_{NN})}$. We thus have:

If $W_{NI} > W_{NN}$ and $W_{II} > W_{NI}$,	$F_I \longrightarrow 1$,
If $W_{NI} > W_{NN}$ and $W_{II} < W_{NI}$,	$F_I \longrightarrow \dfrac{W_{NN} - W_{NI}}{W_{II} + W_{NN} - 2W_{NI}}$
If $W_{NI} < W_{NN}$,	$F_I \longrightarrow 0$.

Inversion equilibrium frequencies therefore depend, as expected, on the relative fitness of homozygotes and heterozygotes for the inversion and of the non-inverted homozygotes. We thus derive the conditions for the inversion to be favoured or disfavoured as a function of the number of mutations captured by the inversion (*m*). A straightforward computation gives (see the Appendix, section 9):

$$W_{II} > W_{NI} \quad \text{if and only if} \quad m < \beta_1(q,h,s) \times n,$$
$$W_{NI} > W_{NN} \quad \text{if and only if} \quad m < \beta_2(q,h,s) \times n$$

with

$$\beta_1(q,h,s) = \frac{\ln(q(1-hs)+1-q)}{\ln\left(\frac{(1-s)(q(1-hs)+1-q)}{q(1-s)+(1-q)(1-hs)}\right)} = \frac{\ln(1-qhs)}{\ln\left(\frac{(1-s)(1-qhs)}{1-hs-qs(1-h)}\right)}, \qquad \text{(Equation 2)}$$

$$\beta_2(q,h,s) = \frac{\ln\left(\frac{1-2q(1-q)hs-q^2s}{q(1-hs)+1-q)}\right)}{\ln\left(\frac{q(1-s)+(1-q)(1-hs)}{q(1-hs)+1-q}\right)}. \qquad \text{(Equation 3)}$$

Assuming *q*<<1 and *s*<<1, these quantities can be approximated by $\beta_1 \approx q\dfrac{h}{1-h}$ and $\beta_2 \approx q$ [the latter in accordance with a previous study (*28*)]. When $qhn/(1-h) < m < nq$, inversions should thus go to fixation on Y chromosomes and stabilize at intermediate frequency on autosomes. When $m < qhn/(1-h)$, inversions should go to fixation on autosomes and on the Y chromosome. Observe that the closer *h* is to ½ (i.e., the scenario without dominance), the smaller the difference between the thresholds β_1 and β_2 is. When h=0.5, $\beta_1 = \beta_2$, inversions should therefore go to fixation on autosome when they have fewer mutations than average, as they have then no homozygous disadvantage preventing their fixation. In contrast, when *h* is small, β_1 is significantly smaller than β_2, showing that the condition for heterozygotes to be favoured over non-inverted homozygotes ($W_{NI} > W_{NN}$) is much easier to meet than for inversion homozygotes to be favoured over inversion heterozygotes ($W_{II} > W_{NI}$

$W_{II} > W_{NI}$): inversions are therefore much likely to be maintained at intermediate frequencies on autosomes than to fix. See Figure S1 for a graphical representation of these results and the Appendix, section 9, for derivation details.

We used these equilibrium frequencies as functions of *m* to compute the expected equilibrium frequency of inversions that can occur in the genome (F_{equ}, Figure 2d and S9). To do so, we sum, for all $\mathrm{m} \in \{0, ..., n\}$ the equilibrium frequency of inversions ($F_{equ,m}$) weighted by their occurrence probability (P_m). For inversions on the Y chromosome, we therefore have:

$$\begin{aligned}[F_{equ}] &= \sum_{m=0}^{n} P_m \times F_{equ,m} \\ &= \sum_{m=0}^{n} \binom{m}{n} q^m (1-q)^{n-m} F_{equ,m} \\ &= \sum_{m=0}^{\lfloor nq \rfloor} \binom{m}{n} q^m (1-q)^{n-m}.\end{aligned} \qquad \text{(Equation 4)}$$

For inversions on autosomes, we obtain, using the approximate values for β_1 and β_2 derived earlier in the case q<<1 and s<<1:

$$[F_{equ}] = \sum_{m=0}^{\lfloor nq \frac{h}{1-h} \rfloor} \binom{m}{n} q^m (1-q)^{n-m} + \sum_{m=\lfloor nq \frac{h}{1-h} \rfloor + 1}^{\lfloor nq \rfloor} \binom{m}{n} q^m (1-q)^{n-m} \frac{W_{NI} - W_{NN}}{2W_{NI} - W_{NN} - W_{II}}$$

(Equation 5)

The expected equilibrium frequency of less-loaded inversions (Figure 2d) is therefore the expected equilibrium frequency of all inversions divided by the probability of occurrence of less-loaded inversions:

$$\begin{aligned}[F_{equ}^{less-loaded}] &= \frac{[F_{equ}]}{\sum_{m=0}^{\lfloor nq \rfloor} P_m} \\ &= \frac{[F_{equ}]}{\sum_{m=0}^{\lfloor nq \rfloor} \binom{m}{n} q^m (1-q)^{n-m}}\end{aligned} \qquad \text{(Equation 6)}$$

Individual-based simulations

We used SliM V3.2 (*68*) to simulate the evolution of a single panmictic population of *N*=1000 or *N*=10,000 individuals in a Wright-Fisher model. To assess the fate of inversions under various conditions (Figure 3), we simulated individuals with two pairs of 10Mb chromosomes on which mutations occurred at a rate *u*, with *u* ranging from 10^{-6} to 10^{-9} per bp, their dominance coefficient *h* ranged from 0 to 0.5 (0, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5) and their selection coefficient *s* from -0.5 to 0 (0, -0.001, -0.01, -0.1, -0.25, -0.5). The main and supplementary figures show the results for $u=10^{-8}$ and $u=10^{-9}$ unless otherwise stated. We also simulated populations where each occurring mutation had its selection coefficient *s* drawn from a gamma distribution with a shape of 0.2, and its dominance coefficient *h* randomly sampled among 0, 0.001, 0.01, 0.1, 0.25, 0.5 with uniform probabilities (Figure

S17). We considered recombination rates of 10^{-6} and 10^{-5} per bp, which gave similar results. Results are presented for analyses in which the recombination rate was 10^{-6}. Among the 5,000,000 sites on chromosome 1, a single segregating locus was subject to balancing selection, for which several situations were considered: (i) the locus had two alleles, only one of which was permanently heterozygous, mimicking a classical XY (or ZW) determining system (Figure 3), (ii) the locus had two permanently heterozygous alleles, mimicking, for instance, the situation encountered at most fungal mating-type loci, iii) the locus had three (or more) permanently heterozygous alleles, mimicking, for instance, the situation encountered in plant self-incompatibility systems and mushroom (*Agaromycotina*) mating-type loci.

For each parameter combination (*u*, *h*, *s*, *N*, heterozygosity rule at the locus with a permanently heterozygous allele), a simulation was run for 15,000 generations, to allow the population to reach an equilibrium for the number of segregating mutations. Figure S18 shows that each population had reached equilibrium by the end of the burn-in period. The population state was saved at the end of this initialization phase. These saved states (one for each parameter combination) were repeatedly used as initial states for studying the dynamics of recombination modifiers. Recombination modifiers mimicking inversions of 500kb, 1000kb, 2000kb and 5000kb were then introduced on chromosome 1 around the locus under balancing selection (X-linked or Y-linked inversions) or on chromosome 2 (autosomal inversions). For each parameter combination (*h*, *s*, *u*, *N*, heterozygosity rule, size of the region affected by the recombination modification and position on the genome), we ran 10,000 independent simulations starting with the introduction of a single recombination modifier in the same saved initial population. These inversion-mimicking, recombination modifier mutations were introduced on a single, randomly selected chromosome and, when heterozygous, they suppressed recombination across the region in which they reside (i.e., as a *cis*-recombination modifier). We monitored the frequency of these inversion-mimicking mutations during 10,000 generations, during which all evolutionary processes (such as point mutation, recombination and mating) remained unchanged, e.g. mutations were still appearing on inversions following their formation. Under the same assumptions and parameters, we also studied the dynamics of recombination modifiers suppressing recombination also when homozygous and not only when heterozygous, again across a fragment in which they reside (Figure S14).

To study more specifically the formation of evolutionary strata on sex chromosomes (Figures 4 and 5), we also simulated the evolution of two 100 Mb chromosomes, one of which carried an XY sex-determining locus at the 50 Mb position, over 115,000 generations (including an initial burn-in of 15,000 generations): individuals could be either XX or XY and could only mate with individuals of a different genotype at this locus. We simulated randomly mating populations of N=1000 and N=10,000 individuals. Point mutations appeared at a rate of $u=10^{-9}$ per bp, and their individual selection coefficients were determined by sampling a gamma distribution with a mean of -0.03 and with a shape of 0.2; these parameter values were set according to observations in humans (*22*, *69*). For each new mutation, a dominance coefficient was chosen from the following values, considered to have uniform probabilities: 0, 0.001, 0.01, 0.1, 0.25, 0.5. At the beginning of each simulation, we randomly sampled *k* genomic positions over the two chromosomes that could be used as inversion breakpoints, with *k* being 10, 100, 1000 or 10,000. After the 15,000 generations of the burn-in period allowing populations to reach an equilibrium in terms of the number of segregating mutations, we introduced each generation

j inversions in the population, *j* being sampled from a Poisson distribution of parameter λ, with $\lambda = N * k * u_i$, u_i being the inversion rate. In order to keep the simulation time tractable, we used inversion rates allowing one inversion to occur on average each generation in the population (i.e. $N * k * u_i = 1$). For each inversion, the first breakpoint was randomly chosen among the *k* positions, and the second among the potential breakpoint positions less than 20Mb apart on the same chromosome (considering therefore only a subset of the *k* positions for the second breakpoints). Two independent inversions could use the same breakpoints, allowing in particular inversion reversion restoring recombination. We assumed that inversions subsequently partially overlapped by another inversion or that captured a smaller inversion could not reverse, i.e. the restoration of recombination by further inversions, even if using the same breakpoints, was prevented. For *N*=1000, for each values of *k* (i.e. 10, 100, 1000, 10000), we ran 10 simulations (Figure 5). For *N*=10,000, because of computing limitation (each simulation taking two or three weeks to run), we only ran one simulation per number of breakpoints.

Simulations were parallelized with GNU Parallel (*70*).

Acknowledgment: We thank Ricardo Rodriguez de la Vega, Fanny Hartmann, Jacqui Shykoff, Sylvain Billiard, Janis Antonovics and Olivier Tenaillon for comments on a draft version of the manuscript. ET thanks Denis Roze for insightful discussions. ET and AV acknowledge support from the chaire program « Mathematical modelling and biodiversity » (Ecole Polytechnique, Museum National d'Histoire Naturelle, Veolia Environnement, Fondation X). PJ thanks the authors of SLiM for their outstanding software and manual.

Funding: This work was supported by the European Research Council (ERC) EvolSexChrom (832352) grant and a Louis D. Foundation (Institut de France) prize to TG.

Author contributions: Original ideas, PJ and TG; deterministic model conception, PJ and ET with input from AV; simulations and data analyses, PJ; interpretation, PJ, TG and ET; manuscript writing, PJ and TG; editing, ET and AV; project management and funding, TG.

Competing interests: The authors have no competing interests to declare.

Data and materials availability: The SLiM and R scripts used for this study are available from GitHub (https://github.com/PaulYannJay/SexChromosomeTheory). No new data were generated for this study. Details concerning the mathematical modeling are available in the appendix.

List of supplementary materials:

Figures S1-19

Appendix (supplementary methods)

References

1. J. K. Abbott, A. K. Nordén, B. Hansson, Sex chromosome evolution: historical insights and future perspectives. *Proc. Biol. Sci.* **284** (2017), doi:10.1098/rspb.2016.2806.

2. D. Charlesworth, Plant contributions to our understanding of sex chromosome evolution. *New Phytol*. **208**, 52–65 (2015).

3. T. Schwander, R. Libbrecht, L. Keller, Supergenes and complex phenotypes. *Curr Biol*. **24**, R288-294 (2014).

4. R. Bergero, D. Charlesworth, The evolution of restricted recombination in sex chromosomes. *Trends in Ecology & Evolution*. **24**, 94–102 (2009).

5. F. E. Hartmann, M. Duhamel, F. Carpentier, M. E. Hood, M. Foulongne Oriol, P. Silar, F. Malagnac, P. Grognet, T. Giraud, Recombination suppression and evolutionary strata around mating-type loci in fungi: documenting patterns and understanding evolutionary and mechanistic causes. *New Phytologist*. **229**, 2470–2491 (2021).

6. S. Ponnikas, H. Sigeman, J. K. Abbott, B. Hansson, Why do sex chromosomes stop recombining? *Trends Genet*. **34**, 492–503 (2018).

7. A. E. Wright, R. Dean, F. Zimmer, J. E. Mank, How to make a sex chromosome. *Nature Communications*. **7**, 12087 (2016).

8. F. Ruzicka, L. Dutoit, P. Czuppon, C. Y. Jordan, X.-Y. Li, C. Olito, A. Runemark, E. I. Svensson, H. P. Yazdi, T. Connallon, The search for sexually antagonistic genes: Practical insights from studies of local adaptation and statistical genomics. *Evolution Letters*. **4**, 398–415 (2020).

9. W. R. Rice, The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution*. **41**, 911–914 (1987).

10. E. Cavoto, S. Neuenschwander, J. Goudet, N. Perrin, Sex-antagonistic genes, XY recombination and feminized Y chromosomes. *Journal of Evolutionary Biology*. **31**, 416–427 (2018).

11. J. E. Ironside, No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome evolution. *BioEssays*. **32**, 718–726 (2010).

12. L. Beukeboom, N. Perrin, *The Evolution of Sex Determination* (Oxford University Press, Oxford, New York, 2014).

13. P. Jay, M. Chouteau, A. Whibley, H. Bastide, H. Parrinello, V. Llaurens, M. Joron, Mutation load at a mimicry supergene sheds new light on the evolution of inversion polymorphisms. *Nat. Genet.* **53**, 288–293 (2021).

14. Z. Yan, S. H. Martin, D. Gotzek, S. V. Arsenault, P. Duchen, Q. Helleu, O. Riba-Grognuz, B. G. Hunt, N. Salamin, D. Shoemaker, K. G. Ross, L. Keller, Evolution of a supergene that regulates a trans-species social polymorphism. *Nature Ecology & Evolution*. **4**, 240–249 (2020).

15. A. L. Bazzicalupo, F. Carpentier, S. P. Otto, T. Giraud, Little evidence of antagonistic selection in the evolutionary strata of fungal mating-type chromosomes (*Microbotryum lychnidis-dioicae*). *G3 Genes|Genomes|Genetics*. **9**, 1987–1998 (2019).

16. S. Branco, H. Badouin, R. C. R. de la Vega, J. Gouzy, F. Carpentier, G. Aguileta, S. Siguenza, J.-T. Brandenburg, M. A. Coelho, M. E. Hood, T. Giraud, Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism. *PNAS*. **114**, 7067–7072 (2017).

17. F. Úbeda, M. M. Patten, G. Wild, On the origin of sex chromosomes from meiotic drive. *Proc Biol Sci*. **282**, 20141932 (2015).

18. B. Charlesworth, J. A. Coyne, N. H. Barton, The relative rates of evolution of sex chromosomes and autosomes. *The American Naturalist*. **130**, 113–146 (1987).

19. J. Antonovics, J. Y. Abrams, Intratetrad mating and the evolution of linkage relationships. *Evolution*. **58**, 702–709 (2004).

20. B. Charlesworth, J. D. Wall, Inbreeding, heterozygote advantage and the evolution of neo–X and neo–Y sex chromosomes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. **266**, 51–56 (1999).

21. D. L. Jeffries, J. F. Gerchen, M. Scharmann, J. R. Pannell, A neutral model for the loss of recombination on sex chromosomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **376**, 20200096 (2021).

22. A. Eyre-Walker, P. D. Keightley, The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* **8**, 610–618 (2007).

23. B. Charlesworth, Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*. **10**, 195–205 (2009).

24. B. Charlesworth, D. Charlesworth, M. T. Morgan, Genetic loads and estimates of mutation rates in highly inbred plant populations. *Nature*. **347**, 380–382 (1990).

25. H. W. Deng, M. Lynch, Inbreeding depression and inferred deleterious-mutation parameters in Daphnia. *Genetics*. **147**, 147–155 (1997).

26. D. Charlesworth, B. Charlesworth, Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics*. **18**, 237–268 (1987).

27. G. Bertorelle, F. Raffini, M. Bosse, C. Bortoluzzi, A. Iannucci, E. Trucchi, H. E. Morales, C. van Oosterhout, Genetic load: genomic estimates and applications in non-model animals. *Nature Reviews Genetics*, 1–12 (2022).

28. M. Nei, K.-I. Kojima, H. E. Schaffer, Frequency changes of new inversions in populations under mutation-selection equilibria. *Genetics*. **57**, 741–750 (1967).

29. T. Ohta, Associative overdominance caused by linked detrimental mutations. *Genetics Research*. **18**, 277–286 (1971).

30. D. Bachtrog, Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat Rev Genet*. **14**, 113–124 (2013).

31. M. Nei, Accumulation of nonfunctional genes on sheltered chromosomes. *The American Naturalist*. **104**, 311–322 (1970).

32. O. Blaser, S. Neuenschwander, N. Perrin, A. E. A. Gardner, E. T. Day, Sex-chromosome turnovers: the hot-potato model. *The American Naturalist*. **183**, 140–146 (2014).

33. C. Grossen, S. Neuenschwander, N. Perrin, The evolution of xy recombination: sexually antagonistic selection versus deleterious mutation load. *Evolution*. **66**, 3155–3166 (2012).

34. T. Connallon, C. Olito, Natural selection and the distribution of chromosomal inversion lengths. *Molecular Ecology* (2021), doi:10.1111/mec.16091.

35. C. Olito, J. K. Abbott, The evolution of suppressed recombination between sex chromosomes by chromosomal inversions. *bioRxiv* (2020), doi:10.1101/2020.03.23.003558.

36. R. A. Neher, B. I. Shraiman, Fluctuations of fitness distributions and the rate of Muller's ratchet. *Genetics*. **191**, 1283–1293 (2012).

37. S. Glémin, How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution*. **57**, 2678–2687 (2003).

38. S. Takayama, A. Isogai, Self-incompatibility in plants. *Annu Rev Plant Biol*. **56**, 467–489 (2005).

39. C. Küpper, M. Stocks, J. E. Risse, N. dos Remedios, L. L. Farrell, S. B. McRae, T. C. Morgan, N. Karlionova, P. Pinchuk, Y. I. Verkuil, A. S. Kitaysky, J. C. Wingfield, T. Piersma, K. Zeng, J. Slate, M. Blaxter, D. B. Lank, T. Burke, A supergene determines highly divergent male reproductive morphs in the ruff. *Nature Genetics*. **48**, 79–83 (2016).

40. J. Wang, Y. Wurm, M. Nipitwattanaphon, O. Riba-Grognuz, Y.-C. Huang, D. Shoemaker, L. Keller, A Y-like social chromosome causes alternative colony organization in fire ants. *Nature*. **493**, 664–668 (2013).

41. L. Stauber, T. Badet, A. Feurtey, S. Prospero, D. Croll, Emergence and diversification of a highly invasive chestnut pathogen lineage across southeastern Europe. *eLife*. **10**, e56279 (2021).

42. T. L. Kubisiak, M. G. Milgroom, Markers linked to vegetative incompatibility (vic) genes and a region of high heterogeneity and reduced recombination near the mating type locus (MAT) in *Cryphonectria parasitica*. *Fungal Genet Biol*. **43**, 453–463 (2006).

43. F. Boideau, G. Richard, O. Coriton, V. Huteau, C. Belser, G. Deniot, F. Eber, C. Falentin, J. Ferreira de Carvalho, M. Gilet, M. Lodé-Taburel, L. Maillet, J. Morice, G. Trotoux, J.-M. Aury, A.-M. Chèvre, M. Rousseau-Gueutin, Epigenomic and structural events preclude recombination in *Brassica napus*. *New Phytologist* (2021), doi:10.1111/nph.18004.

44. M. Kimura, T. Ohta, The average number of generations until fixation of a mutant gene in a finite population. *Genetics*. **61**, 763–771 (1969).

45. D. Porubsky, W. Höps, H. Ashraf, P. Hsieh, B. Rodriguez-Martin, F. Yilmaz, J. Ebler, P. Hallast, F. A. M. Maggiolini, W. T. Harvey, B. Henning, P. A. Audano, D. S. Gordon, P. Ebert, P. Hasenfeld, E. Benito, Q. Zhu, H. G. S. V. Consortium (HGSVC), C. Lee, F. Antonacci, M. Steinrücken, C. R. Beck, A. D. Sanders, T. Marschall, E. E. Eichler, J. O. Korbel, Haplotype-resolved inversion landscape reveals hotspots of mutational recurrence associated with genomic disorders. *bioRxiv* (2021), doi:10.1101/2021.12.20.472354.

46. T. Lenormand, D. Roze, Y recombination arrest and degeneration in the absence of sexual dimorphism. *Science*. **375**, 663–666 (2022).

47. M. Kirkpatrick, N. Barton, Chromosome inversions, local adaptation and speciation. *Genetics*. **173**, 419–434 (2006).

48. T. Connallon, C. Olito, L. Dutoit, H. Papoli, F. Ruzicka, L. Yong, Local adaptation and the evolution of inversions on sex chromosomes and autosomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **373**, 20170423 (2018).

49. M. Kirkpatrick, How and why chromosome inversions evolve. *PLoS Biol.* **8**, e1000501 (2010).

50. B. T. Lahn, D. C. Page, Four evolutionary strata on the human X chromosome. *Science*. **286**, 964–967 (1999).

51. S. J. Hanson, K. P. Byrne, K. H. Wolfe, Mating-type switching by chromosomal inversion in methylotrophic yeasts suggests an origin for the three-locus *Saccharomyces cerevisiae* system. *Proc Natl Acad Sci U S A*. **111**, E4851-4858 (2014).

52. L. Cui, H. Neoh, A. Iwamoto, K. Hiramatsu, Coordinated phenotype switching with large-scale chromosome flip-flop inversion observed in bacteria. *PNAS*. **109**, E1647–E1656 (2012).

53. H. Badouin, M. E. Hood, J. Gouzy, G. Aguileta, S. Siguenza, M. H. Perlin, C. A. Cuomo, C. Fairhead, A. Branca, T. Giraud, Chaos of rearrangements in the mating-type chromosomes of the anther-smut fungus *Microbotryum lychnidis-dioicae*. *Genetics*. **200**, 1275–1284 (2015).

54. S. B. Carey, J. Jenkins, J. T. Lovell, F. Maumus, A. Sreedasyam, A. C. Payton, S. Shu, G. P. Tiley, N. Fernandez-Pozo, A. Healey, K. Barry, C. Chen, M. Wang, A. Lipzen, C. Daum, C. A. Saski, J. C. McBreen, R. E. Conrad, L. M. Kollar, S. Olsson, S. Huttunen, J. B. Landis, J. G. Burleigh, N. J. Wickett, M. G. Johnson, S. A. Rensing, J. Grimwood, J. Schmutz, S. F. McDaniel, Gene-rich UV sex chromosomes harbor conserved regulators of sexual development. *Sci Adv*. **7**, eabh2488 (2021).

55. M. B. Couger, S. W. Roy, N. Anderson, L. Gozashti, S. Pirro, L. S. Millward, M. Kim, D. Kilburn, K. J. Liu, T. M. Wilson, C. W. Epps, L. Dizney, L. A. Ruedas, P. Campbell, Sex chromosome transformation and the origin of a male-specific X chromosome in the creeping vole. *Science*. **372**, 592–600 (2021).

56. B. M. Skinner, C. A. Sargent, C. Churcher, T. Hunt, J. Herrero, J. E. Loveland, M. Dunn, S. Louzada, B. Fu, W. Chow, J. Gilbert, S. Austin-Guest, K. Beal, D. Carvalho-Silva, W. Cheng, D. Gordon, D. Grafham, M. Hardy, J. Harley, H. Hauser, P. Howden, K. Howe, K. Lachani, P. J. I. Ellis, D. Kelly, G. Kerry, J. Kerwin, B. L. Ng, G. Threadgold, T. Wileman, J. M. D. Wood, F. Yang, J. Harrow, N. A. Affara, C. Tyler-Smith, The pig X and Y Chromosomes: structure, sequence, and evolution. *Genome Res*. **26**, 130–139 (2016).

57. D. W. Bellott, J. F. Hughes, H. Skaletsky, L. G. Brown, T. Pyntikova, T.-J. Cho, N. Koutseva, S. Zaghlul, T. Graves, S. Rock, C. Kremitzki, R. S. Fulton, S. Dugan, Y. Ding, D. Morton, Z. Khan, L. Lewis, C. Buhay, Q. Wang, J. Watt, M. Holder, S. Lee, L. Nazareth, J. Alföldi, S. Rozen, D. M. Muzny, W. C. Warren, R. A. Gibbs, R. K. Wilson, D. C. Page, Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature*. **508**, 494–499 (2014).

58. C. Lemaitre, M. D. V. Braga, C. Gautier, M.-F. Sagot, E. Tannier, G. A. B. Marais, Footprints of inversions at present and past pseudoautosomal boundaries in human sex chromosomes. *Genome Biol Evol*. **1**, 56–66 (2009).

59. N. Osada, Genetic diversity in humans and non-human primates and its evolutionary consequences. *Genes Genet Syst*. **90**, 133–145 (2015).

60. B. M. Henn, L. R. Botigué, C. D. Bustamante, A. G. Clark, S. Gravel, Estimating the mutation load in human genomes. *Nat Rev Genet*. **16**, 333–343 (2015).

61. C. Giner-Delgado, S. Villatoro, J. Lerga-Jaso, M. Gayà-Vidal, M. Oliva, D. Castellano, L. Pantano, B. D. Bitarello, D. Izquierdo, I. Noguera, I. Olalde, A. Delprat, A. Blancher, C. Lalueza-Fox, T. Esko, P. F. O'Reilly, A. M. Andrés, L. Ferretti, M. Puig, M. Cáceres, Evolutionary and functional impact of common polymorphic inversions in the human genome. *Nature Communications*. **10**, 4222 (2019).

62. M. Wellenreuther, L. Bernatchez, Eco-evolutionary genomics of chromosomal inversions. *Trends Ecol. Evol.* **33**, 427–440 (2018).

63. M. V. Chibalina, D. A. Filatov, Plant Y chromosome degeneration is retarded by haploid purifying selection. *Curr Biol*. **21**, 1475–1479 (2011).

64. D. A. Filatov, Homomorphic plant sex chromosomes are coming of age. *Mol Ecol*. **24**, 3217–3219 (2015).

65. S. Coelho, J. Gueno, A. Lipinska, J. Cock, J. Umen, UV chromosomes and haploid sexual systems. *Trends Plant Sci*. **23**, 794–807 (2018).

66. J. R. Chasnov, Mutation-selection balance, dominance and the maintenance of sex. *Genetics*. **156**, 1419–1425 (2000).

67. A. Agrawal, M. Whitlock, Mutation load: the fitness of individuals in populations where deleterious alleles are abundant. *Annual Review of Ecology Evolution and Systematics*. **43**, 115–135 (2012).

68. B. C. Haller, P. W. Messer, SLiM 3: Forward genetic simulations beyond the wright–fisher model. *Mol. Biol. Evol.* **36**, 632–637 (2019).

69. B. Y. Kim, C. D. Huber, K. E. Lohmueller, Inference of the distribution of selection coefficients for new nonsynonymous mutations using large samples. *Genetics*. **206**, 345–361 (2017).

70. O. Tange, Gnu parallel-the command-line power tool. *login: The USENIX Magazine* (2011), pp. 42–47.