Revisiting the shelter theory: can deleterious mutations alone explain the evolution of sex chromosomes?

Authors: Thomas Lenormand[1]*, Denis Roze[2,3]

Affiliations:

[1] CEFE, Univ Montpellier, CNRS, Univ Paul Valéry Montpellier 3, EPHE, IRD, Montpellier, France.

[2] CNRS, IRL 3614, Roscoff, France.

[3] Sorbonne Université, Station Biologique de Roscoff, France.

*Correspondence to: thomas.lenormand@cefe.cnrs.fr

Short title: Revisiting the shelter theory

Abstract

Y and W chromosomes often stop recombining and degenerate. Different theories have been proposed to explain these observations. The classic theory supposes that recombination arrest is favored because it allows to permanently link sex-antagonistic alleles to the sex in which they are favorable. The regulatory theory supposes that recombination arrest is initially caused by lucky inversions on the Y or W carrying fewer deleterious mutations than average, which then become selectively stabilized by the evolution of nascent dosage compensation creating sex-antagonistic regulatory effects. Last, the shelter theory also supposes that lucky Y (or W) inversions do occur, but that they are maintained in the long term because of a lack of genetic variation to restore recombination. The last two theories are comparable in the sense that they do not require sexually dimorphic traits, and are based primarily on the flux of deleterious mutations. In this paper, we focus on the shelter theory and show that contrarily to previous claims, it can hardly explain the evolution of sex chromosomes. We show that it is not robust to extremely low rates of recombination restoration, that it would cause population extinction rather than Y degeneration, and that it is orders of magnitudes less efficient than the mechanism described in the regulatory theory. We also argue that the assumption of a lack of genetic variation to restore recombination is unplausible given known mechanisms of recombination variation, and given the extremely small rates of restoration that are sufficient to make the shelter theory inoperant.

Introduction

Maladaptation is ubiquitous. Populations constantly adapt, against a diversity of processes limiting the optimization of phenotypes [1–3]. In many cases, maladaptation is the result of conflicting selective pressures - due to trade-offs between traits [4] or genetic conflicts [5] - or stems from the effect of other processes than natural selection acting on populations, such as genetic drift [6] or gene flow [7]. Maladaptation may be strong enough to lead to the extinction of populations or even species [3,8–10]. Within the genome, a conspicuous case of maladaptation seems to include degenerate Y or W sex chromosomes. Beyond the fact that the degeneration of sex chromosomes directly involves an accumulation of deleterious mutations, it does seem to have a negative effect on lifespan [11–13]. How could such a degeneration evolve at the scale of a whole chromosome without being opposed by natural selection?

Several theories have been put forward to explain this, often separating the problem into several subproblems: Why does recombination stop between sex chromosomes? Why does degeneration occur on a nonrecombining part of a genome? Why is recombination suppression maintained in the long term, once degeneration has already caused substantial degradation of the Y or W? Why isn't recombination eventually reestablished to limit maladaptation? These questions have been intensely investigated, and have led to different theoretical scenarios for sex chromosome evolution. (1) In the *classic* scenario [14–20], suppression of recombination is selectively favored due to the occurrence of loci with sex antagonistic (SA) effects on sex chromosomes. In this case, a non-recombining Y benefits from a selective advantage, by permanently combining male determining and male-beneficial alleles. This selective advantage ensures that suppression of recombination is selectively maintained. Then, deleterious mutations accumulate due to selective interference in the absence of recombination. (2) In the *regulatory* scenario [21,22], 'lucky inversions' capture by chance a portion of Y with a smaller than average load of deleterious mutations (the process can involve any Y variation suppressing recombination, not only inversions). This creates an immediate (but fortuitous) selective advantage to such a non-recombining Y, which then may spread to fixation. Once recombination is suppressed, cis regulators of gene expression may randomly diverge between the X and Y, for genes located on the non-recombining portion of the Y. This regulatory instability may lead to the evolution of early dosage compensation (through the joint evolution of cis and trans acting factors), concomitant with Y early silencing and degeneration. The emergence of this dosage compensation builds up SA regulatory effects, preventing the long-term reestablishment of recombination. (3) The *shelter* scenario [23] involves the same initial step as the regulatory scenario (fixation of 'lucky inversions'), but proposes that degeneration occurs through selective interference, as in the *classic* scenario. This scenario therefore only relies on deleterious mutations, and does not involve selection against restoration of X-

Y recombination (beyond the fortuitous initial advantage of the lucky inversion). The term “mutation shelter” refers to the idea that deleterious mutations could easily shelter and accumulate on the Y chromosome, after these fortuitous recombination suppression events. Additional and more specific scenarios have been proposed, but will not be discussed here [see 24,25]. In particular, the neutral accumulation of sequence divergence near the sex-determining locus may locally contribute to recombination suppression [26]. Tight linkage on the Y may also be favored by sex-differences in the intensity of selection (but not in the direction of selection as considered in SA theory) against deleterious mutations [27], or by overdominance (or associative overdominance) in partially inbred populations [28,29].

In the classic and shelter scenarios, it is not totally clear why maladaptation is maintained in the long term. Eventually, the accumulation of deleterious mutations on the Y may offset the initial fitness of the non-recombining Y. At this point, there may be no selective advantage to recombination suppression, and recombination may be reestablished. In particular, this is a pressing problem in the shelter scenario, where the initial selective advantage of the nonrecombining Y is likely to be modest. Indeed, the initial advantage of a lucky inversion spanning a given segment of chromosome is drawn from the distribution of marginal fitnesses of the different copies of this segment in the population at mutation-selection balance, and is unlikely to exceed a few percents. This small initial advantage is therefore likely to be quickly offset by a modest process of degeneration [22]. The long-term maintenance of recombination suppression is also an issue in the classic scenario, although less pressing, as SA loci can provide substantial fitness benefits to the non-recombing Y and can evolve repeatedly. In this case, recombination arrest can remain beneficial for a longer period before the benefit of SA loci is entirely offset by degeneration, potentially allowing populations to evolve gene silencing / dosage compensation limiting maladaptation caused by Y degeneration.

In the shelter scenario, as deleterious mutations accumulate on the Y on many genes, it becomes less and less likely that the current non-recombing Y has a fitness higher than the fitness of a Y that would be freshly recombined with the X and carrying (almost) perfectly functional genes. For instance, it is straightforward to see that the fitness of a Y with fewer than average deleterious mutations may be initially above average, but soon below average after starting to accumulate deleterious mutations. Hence, it is difficult to explain how recombination suppression can be maintained in the long term in the shelter scenario, unless invoking regulatory effects or some specific and ad hoc constraints. Regulatory effects include the evolution of Y silencing and dosage compensation, mitigating the deleterious effects of mutations accumulated on the Y. However, if these regulatory processes are considered, they have to be included from the start, and the processes described in the regulatory scenario will likely take place. Alternatively, one may suppose that recombination cannot be

reestablished once it has been suppressed. One possibility may be that it is difficult to reestablish synteny once multiple nested inversions have occurred. This is the scenario proposed by Jay et al. [23], assuming that recombination can never be reestablished once a nested or overlapping inversion occurs on the Y, placing this absolute constraint as the core explanation for the evolution of degenerated sex chromosomes. Although rarely investigated [30–32], a similar difficulty occurs in the classic scenario, where the possibility that recombination may be reestablished is generally not considered. In this paper, we revisit this question and argue that explanations relying on the constraint that recombination cannot be reestablished on the Y suffer from a series of serious theoretical difficulties, making the shelter theory unlikely to account for the evolution of sex chromosomes. We then also explain why the mechanistic constraints on recombination restoration are very unlikely to be nearly as absolute as assumed in the shelter scenario described in [23].

Results

Modelling sex chromosome evolution

We use the general model of sex chromosome evolution that we previously introduced to explore the regulatory scenario [19,20, that should be consulted for more details]. This model considers a sex chromosome pair with a large number of genes (here 500) subject to deleterious mutations occurring in their coding sequence. The expression of each of these genes is controlled by a cis-regulator and two trans-regulators (one only expressed in males and the other only expressed in females). These regulators determine quantitative traits that control the total and allele-specific level of expression of each gene. Total gene expression is supposed to be under stabilizing selection for all genes. Fitness is determined multiplicatively across loci by the effect of deleterious mutations (whose dominance depends on the relative strength of cis-regulators in heterozygotes, with a baseline dominance in absence of *cis*-regulatory variation equal to 0.25) and by the departure from optimal expression at each locus. Cis and trans regulators mutate at a fixed rate per generation (with gaussian variation in trait values). For simplicity, the sex-determining locus (SDL) is located at one extremity of the chromosome. Recombination variation is modeled by introducing mutations suppressing recombination in a region around the SDL. These mutations can be thought as being inversions, and we will refer to them as such, although other types of mechanisms are possible. For instance, the removal of recombination hotspots, or the addition of a recombination suppressor sequence would work too. Specifically, as described in [14], we assume that inversions occur on the Y at a rate U_{inv} per chromosome per generation (we use $U_{inv} = 10^{-5}$). We only consider inversions that include the sex locus. Other inversions are not confined to males and can fix in the population, which does not lead to recombination suppression on the Y (e.g., homozygous inversions recombine normally). We denote *z*

133 (between 0 and 1) the non-recombining fraction of the Y. This variable is also used to measure the
134 endpoint of each inversion on the chromosome. When $z = 0$, X and Y chromosomes recombine freely,
135 but otherwise X-Y recombination only occurs within the chromosomal segment [z, 1] (the sex-
136 determining locus being located at position 0). When $z = 1$, the X and Y do not recombine at all. When
137 a new inversion occurs, its size is drawn as a uniform fraction of the non-recombining part of the Y.
138 Specifically, on a Y where recombination is already stopped between 0 and z_i, after a new inversion i+1
139 the non-recombining region will extend to $z_{i+1} = z_i + (1 - z_i)u$, where u is a uniform deviate between
140 0 and 1. To allow for the possibility that recombination may be reestablished, we assume that
141 reversions can also occur (at a rate U_{rev} per chromosome per generation), reverting the last inversion
142 on the non-recombining part of the Y. We investigate the dynamics of Y evolution in this model by
143 supposing that reversion rates are much smaller than inversion rates, with $U_{rev} = 10^{-6}, 10^{-7}, 10^{-8}$,
144 10^{-9}, i.e., from 1 to 4 orders of magnitude lower than the rate of occurrence of inversions. This model
145 can easily compare the shelter and regulatory scenario, everything else being equal: indeed, the shelter
146 scenario can be directly studied by setting mutations rates on regulators to zero. At the start of a
147 simulation, each individual carries a pair of fully recombining chromosomes, with the SDL located at
148 one extremity. The simulations assume a constant population size (10^4 offpring individuals are drawn
149 each generation, irrespectively of the average absolute fitness of male and female individuals in the
150 previous generation). However, the population was considered extinct when the average fitness of
151 males became a thousand times lower than the average fitness of females.

152 The shelter scenario does not lead to long-term recombination suppression

153 We previously showed that the regulatory scenario leads to the stacking of inversions on the Y that
154 persist indefinitely [22]. These inversions become stabilized by the early evolution of dosage
155 compensation, which automatically generates sex-antagonistic regulatory effects. In this case,
156 reversions occur but they are not selectively favored. Indeed, the reestablishment of X-Y
157 recombination on a given stratum causes X cis-regulators to move to the Y, creating recombinant (low
158 fitness) Y that cause a departure from optimal gene expression in males. This is a case of an evolved
159 selective constraint. The ultimate cause of long-term recombination suppression is not the absence of
160 genetic variation for recombination once an inversion has fixed (inherent constraint), but that it is
161 selectively unfavorable to re-establish it.

162 To evaluate whether the shelter scenario could lead to long-term recombination suppression, we
163 investigated cases where the rate of reversion was one to four orders of magnitude lower than the
164 rate of inversion. We ran replicated simulations lasting four million generations. With our standard
165 parameters (Table 1), approximately 0.76 inversions fix per million generations (which is well within
166 the observed rate of fixation of inversions in Drosophila [33]). Typical outcomes are illustrated on Fig

1 (taken from runs with $U_{rev} = 10^{-8}$). The majority of inversions that reach fixation remain relatively short-lived (Fig 2A) and each of them occurs one at a time (i.e. they are reverted before a second one fixes). When they fix, Y inversions start degenerating, which reduces male fitness and eventually offsets their initial fitness advantage. At this point, they become deleterious, and it is just a matter of time before a reversion occurs that would be selectively favored (Fig 2B shows the marginal fitness at birth and at the time of reversion for fixed inversions, for different reversion rates). This is the example illustrated on Fig 1A. In a few cases, another inversion fixes on top of the first one, before the first has reversed (this corresponds to the example illustrated on Fig 1B). When the rate of reversion is very low, several inversions may stack on the first one. In these cases, the life-span of the first inversion is prolonged because it can be reversed only after the second one (or third one etc.) is reversed. This is due to the rather stringent hypothesis of our model that reversions can only occur on the last stratum present on the Y. This assumption protects the first inversion from reversion (until all other strata have reverted), and considerably reduces effective reversion rates (as it becomes zero for all but the last stratum on a Y). Considering reversions that could fully reestablish recombination on the Y would greatly reduce this (potentially unrealistic) effect. Whether or not stacking occurs, however, all inversions become reversed at some point (or the population goes extinct as we discuss below). There is no long-term maintenance of recombination suppression. Typically, Y chromosomes transiently carry one or two strata (Fig 3) for a relatively short time (Fig 2A). We can therefore conclude that the shelter scenario cannot explain the long-term maintenance of recombination suppression for rates of reversion up to 4 orders of magnitude lower than rates of inversion. This conclusion is very conservative, as our model of reversion does not allow the reversion of inversions associated to more recent inversions.

The shelter scenario does not lead to degeneration

When reversion rates are low, some inversions can fix and persist in the population for some time, especially when secondary inversions also occur prolonging the lifespan of the first one. Usually, this does not correspond to a large fraction of the Y, but on rare occasions, this can be significant (in particular, when several fixed inversions are stacked, which reduces the rate of reversion of all but the last one of them, as explained above). For example, in the example illustrated on Fig 1B, about a third of the Y stopped recombining for nearly a million generations. Even in these extreme cases, however, degeneration remains moderate: even with extremely low rates of reversion (10^{-9}), almost no gene accumulates deleterious mutations up to s_{max} (the maximal value that would correspond to a loss of function for a gene, fixed to s_{max} = 0.3 in our simulations). Yet, we use a relatively high rate of deleterious mutation per gene (U_g = 2.10^{-4}), a distribution of fitness effects of mutations with a relatively high mean (s_{mean} = 0.05) and a large proportion of small effect mutations (the distribution of

effects is exponential). In the vast majority of cases, almost no degeneration is detectable (Fig 4A). The reason for this lack of degeneration is that many weakly deleterious mutations accumulate on all genes present on a fixed inversion. Collectively, their impact on the marginal fitness of the inversion starts to be strong long before any gene in particular becomes fully nonfunctional. Hence, inversions become selectively disfavored (and therefore selectively eliminated as soon as a reversion arises), long before they exhibit significant degeneration. We can therefore conclude that the shelter scenario cannot explain strong degeneration for rates of reversion up to 4 orders of magnitude lower than rates of inversion.

The shelter scenario would rather cause extinction than Y degeneration

It may be argued that reversion rates are even smaller than the one we considered, making the shelter scenario a possibility, at least theoretically. There is a strong argument against this possibility. When an inversion fixes and starts accumulating deleterious mutations, it starts depressing male fitness. Initially, a lucky inversion is selectively favored because it captures a fraction of the Y carrying fewer deleterious mutations (or deleterious mutations with smaller effects) compared to the average Y population. Once it is fixed, deleterious mutations start accumulating within the inversion due to selective interference. As explained above, a first threshold is reached when the accumulation of deleterious mutations depresses the marginal fitness of this portion of the Y below the average marginal fitness of the homologous portion of the X in the population. Reversions then become selectively favored, and it is just a matter of time before one occurs and eliminates the inversion. If the reversion rate is extremely low, this can indeed take a long time. However, a second threshold will be reached relatively quickly, corresponding to the non-viability or sterility of males, and hence the extinction of the population/species. This cannot happen in our model as we assume a constant population size, i.e., the absolute number of individuals in the population does not depend on the fitness of individuals (soft selection), but the simulations show a crash in male fitness relative to female fitness when inversions are maintained for a sufficiently long time (Fig 4B). It is not easy to determine the male fitness threshold that would lead to population extinction in nature. We used a relatively conservative threshold equal to 10^{-3} (meaning that the fitness of males is three orders of magnitude lower than the fitness of females). Such a distortion of male vs. female fitness would be particularly conspicuous in natural conditions. In the few cases where some degeneration occurs (e.g. when a large inversion unfortunately fixes), the population reaches this limit quickly and becomes extinct. Reversions can rescue the population and prevent extinction, but if they are too rare, they do not occur quickly enough to prevent it. For instance, with very low rates of reversions ($U_{rev} = 10^{-9}$), this threshold was often reached in our simulations (extinction occurred in 70% of cases, 14 replicates out of 20 within the first 4 million generations of evolution). This estimate is conservative, since considering that

235 the number of males in the population may be much lower than assumed under our soft selection
236 regime would lead to an even faster accumulation of deleterious mutations on the Y (due to stronger
237 drift). Hence, a theory based on maladaptation alone cannot explain both degeneration and the
238 persistence of the population/species. For degeneration and persistence to occur, Y silencing and
239 dosage compensation must also evolve. However, considering these regulatory changes from the start
240 would lead to the regulatory scenario, in which there is no need to suppose any constraint on the rate
241 of reversion in the first place. We can therefore conclude that the shelter scenario cannot explain the
242 evolution of sex chromosomes, even when considering extremely low rates of reversion.

243 The shelter scenario is slow, partial and transient compared to the regulatory scenario

244 We saw the difficulties associated with the shelter scenario. Here, we examine whether alternatives
245 fare better. To do so, we compared the evolution of Y chromosomes with the same low reversion rates
246 used above, under the shelter and regulatory scenarios. We find that everything else being equal, the
247 regulatory theory leads to much faster Y recombination suppression and degeneration. In fact, low
248 reversion rates are favorable to any theory aiming at explaining the maintenance of recombination
249 arrest, not only to the shelter theory. Indeed, low reversion rates gives more time for degeneration (in
250 the classic and shelter scenarios) or dosage compensation and degeneration (in the regulatory
251 scenario) to happen. With very low reversion rates, almost any fixed inversion in the regulatory
252 scenario has time to develop nascent dosage compensation, generating sex-antagonistic regulatory
253 effects that effectively disfavor recombination [20]. Fig 3 shows that many more strata accumulate on
254 the Y under the regulatory scenario, while Fig 4C shows that most of these strata are long lived and
255 fully degenerated (while almost none is degenerated for the same parameter values under the shelter
256 scenario, Fig 4A). Fig 4D shows that this degeneration is not associated to a large drop in male to female
257 fitness ratio, while this drop is considerable under the shelter scenario. Overall, for the parameter
258 values considered here, we see that after 2 million generations a large fraction of the Y has stopped
259 recombining and degenerated under the regulatory theory with the lowest reversion rates (for high
260 reversion rates the process is still ongoing with only few small stabilized and degenerated strata, Fig.
261 5). Nothing nearly comparable occurs under the shelter theory, where at most, in a few cases with
262 extremely low rates of reversion, some recombination suppression evolves and drives the population
263 to extinction without leading to significant Y degeneration (Fig 5). We conclude that even if reversion
264 rates were extremely low, so that the shelter scenario could produce partial and transient
265 degeneration and recombination suppression, the regulatory scenario is orders of magnitude more
266 efficient at producing complete and permanent degeneration and recombination suppression.

Discussion

The different theories mentioned above place a different emphasis on different ultimate causes of recombination suppression on sex chromosomes. The classical theory gives a predominant role to sex-antagonistic loci involved in sexual dimorphism. The shelter theory emphasizes the lack of variation to restore recombination once it has been suppressed through lucky inversions. The regulatory theory emphasizes the build-up of regulatory sex-antagonistic effects associated with the emergence of dosage compensation on lucky inversions. As explained above, there are theoretical difficulties to not consider regulatory variation from the start in any theory of sex-chromosome evolution. Among the most pressing, we can list (1) the crash of male fitness occurring quicker than degeneration, (2) the slow rate of degeneration in large populations, (3) the slow rate of degeneration on small strata with only a few genes. Beyond these difficulties, the values of some of the most important parameters involved in each of these theories remain largely unknown. What is the density of polymorphic sex-antagonistic loci within genomes? What are the rates of recombination suppression and restoration? What are the rates of regulatory changes, especially sex-limited ones?

The constraint on reversions

Regarding the shelter scenario, which is the main topic of this paper, perhaps the most pressing issue is to evaluate the degree to which recombination restoration is effectively constrained. The case for the existence of a constraint has been made by Jay et al [23] based on their model of inversions and reversions. In this model, a number k of possible inversion breakpoints are distributed randomly along the chromosome. When an inversion occurs, the first breakpoint is drawn at random and the second is then drawn among all those which are within 1/5 of the chromosome of the first. For reversion to be possible (and recombination to be restored), a second inversion involving exactly the same breakpoints (and thus restoring the exact initial gene order) must arise before another inversion overlaps or falls within the first (in which case reversion is assumed to be no longer possible). When the number of possible breakpoints is small (k = 10), Jay et al [21] observed that all inversions are eventually reversed and recombination suppression is thus not maintained: in this case, most inversions presumably occur between adjacent breakpoints and thus cannot be blocked by other inversions. For k >=100, however, reversions are never observed. Indeed, in this case it becomes extremely unlikely (in fact almost impossible) that a reversion will occur: the probability that a second inversion has exactly the same breakpoints as the first is of the order $1/k^2$, while the probability that a second inversion overlaps (or is contained within) the first is roughly of the order 0.1 (given that the average size of an inversion is about 1/10 of the chromosome). So, it is extremely unlikely that a reversion will happen before an inversion overlapping the first one has occurred. This model thus imposes very strict conditions for recombination to be reestablished. However, one may wonder

301 whether this model of inversion and reversion is realistic. In particular, does it really take an inversion
302 that exactly reverses the first one for recombination to be restored?

303 Crossing-over within inversions

304 Inversions are frequent in genomes and are well known to reduce recombination rates in
305 heterokaryotes. However, this reduction in recombination is not necessarily because inversions inhibit
306 homologous pairing. If inversions are sufficiently large, pairing can occur, and inversions form loops
307 allowing for a local alignment of the two homologous chromosomes. Inversion loops can directly
308 inhibit chiasma formation, especially near the breakpoints of the inversion [34,35, but see 36].
309 However, the suppression of recombination is also strongly mediated by the fact that an odd number
310 of crossovers within the inversion loop leads to the production of unbalanced chromosomes. For
311 paracentric inversions (i.e., inversions not carrying the centromere), one crossover in the loop leads to
312 an acentric recombined fragment (which lacks a centromere and will therefore be lost) and centric
313 fragments missing an entire part of the original chromosome (to the right or left of the crossover) that
314 are likely to be strongly deleterious (unless the missing part is very short, for example if the inversion
315 is close to a telomere) [37]. With pericentric inversions (i.e., inversions carrying the centromere), the
316 recombined chromosomes all carry a single centromere, but show a missing and a duplicated part
317 compared to the original chromosome (corresponding roughly to the left or right parts of the
318 chromosomes compared to the position of the crossover). These unbalanced chromosomes are likely
319 to be strongly deleterious too, although they may not be eliminated immediately (unlike the case of
320 paracentric inversions where recombinant fragments are immediately eliminated due to the lack of a
321 centromere or the occurrence of a dicentric bridge) [37]. Even in the case of an inversion located on
322 the Y and degenerated, these recombination events would not allow to specifically delete
323 nonfunctional parts of the Y as the deletions would involve full chromosome arms, not only the
324 inverted part. For these reasons, recombination is highly suppressed within inversions. However,
325 recombination can occur when the number of crossovers falling within the inversion is even. Rates of
326 double crossing over are low, but much larger (10^{-3} to 10^{-5}) than the reversion rates that we considered,
327 especially for relatively large inversions where crossover interference is weaker [38–41]. Gene
328 conversion events can also locally lead to the elimination of deleterious mutations carried on
329 inversions [34,41]. In the long term, these events, even if they remain rare and orders of magnitude
330 lower than standard recombination rates, can have important effects on patterns of variation within
331 inversions [41,42]. In the long-term rare recombination events within inversions are therefore likely to
332 occur, at least with double crossing-over. It therefore seems unrealistic to assume that recombination
333 can only occur if the exact initial gene order is restored by a second inversion with exactly the same
334 breakpoints. The accumulation of deleterious mutations in Y inversions could in principle be purged by

335 rare double crossovers occurring within inversions, which would be sufficient to prevent long-term
336 degeneration of the Y (unless SA effects are present within inversions, which prevents the creation of
337 such recombined Y with a high fitness).

338 Nested and overlapping inversions

339 Jay et al's model also assumes that once a secondary inversion takes place within or overlapping with
340 a first one, then recombination is irreversibly suppressed. This seems unplausible too. As long as there
341 are homologous sequences on the two chromosomes, recombination could occur, even if they are not
342 perfectly colinear (and gene conversion as well). This is illustrated on Fig 6 for an overlapping inversion
343 that includes the SDL and on Fig 7 for nested inversions that include the SDL. In both cases, the
344 secondary inversion has a homologous region on the X and can pair at meiosis, allowing recombination
345 to occur around the SDL. If a crossover occurs within this region, a new Y can be produced which may
346 not carry the exact complement compared to the original chromosome (with either deleted or
347 duplicated positions). In each case, such a crossover will eliminate parts of the former Y on the first
348 stratum, i.e., regions that may have already accumulated a load of deleterious mutations. This
349 recombined Y could be particularly favorable (compared to the degenerated Y), even if some
350 sequences are duplicated or missing compared to the X (see Figs 6, 7). Furthermore, the recombined
351 Y can be further and easily "improved" in further steps since it can now recombine easily with the X
352 around the SDL after this first recombination event. In particular, a second crossover near the SDL can
353 further improve collinearity with the X and eliminate further degenerated parts of the Y from the first
354 stratum that are still present (see Figs 6, 7). Overall, it does not seem that recombination suppression
355 could be maintained indefinitely if secondary inversions are involved. Quite on the contrary, a
356 secondary nested or overlapping inversion is potentially a very efficient mechanism to restore
357 recombination around the SDL, and not something that would irreversibly suppress recombination as
358 assumed in Jay et al. (2022).

359 Broadening the scope for recombination restoration

360 Restoration of recombination on the Y can occur for a variety of reasons, beyond the two mechanisms
361 just mentioned (double crossovers within inversions, and nested/overlapping secondary inversions).
362 Another important way to restore recombination around the SDL is simply to move the SDL outside
363 the inversion where it finds itself. Such an event can occur for instance by recruiting a new master
364 switch gene for sex-determination, or following a duplication of the existing SDL into a new location.
365 Gene duplication are frequent events in eukaryotic genomes [43]. Rates in the range 10^{-5} – 10^{-7} per
366 gene per generation have been reported in animals [44–46], i.e. a much higher rate than the rate of
367 reversions that we investigated. If the SDL moves to another (recombining) location on the sex
368 chromosome, a new Y recently rederived from the X (and fully recombining) can evolve. Examples of

such events are often reported [47–49], indicating that they may be common. The SDL may also move to another autosomal pair, leading to the evolution of a new pair of fully recombining sex chromosomes. Examples of such turnovers abound [50], indicating that they are also common. Other possibilities for restoring recombination on the Y involve environmental sex reversal when recombination is sex-dependent [30] or inversions occurring on the X, facing inversion on the Y [22]. Hence, assuming that recombination cannot be restored on the Y requires that none of these mechanisms can occur even at very small rates, which seems particularly unplausible.

Conclusion

Deleterious mutations can definitely play a role as a proximal mechanism for recombination suppression on Y or W chromosomes. “Lucky” inversions can arise and fix relatively frequently in populations by capturing a portion of the Y carrying fewer or milder deleterious mutations than the average Y in the population. This process does not require the occurrence of sex-antagonistic loci to initially promote recombination suppression. This has been proposed as a possible first step in the regulatory theory [22]. In that theory, regulatory sex-antagonistic effects build up on these lucky inversions, because recombination suppression allows X and Y cis-regulators to diverge, in turn leading to the emergence of dosage compensation. This process ensures that recombination suppression is maintained in the long term, even when reversions occur at a high rate, since recombination suppression stays selectively favored throughout. The shelter theory [23] proposes the same initial step involving lucky inversions, but explains the long-term maintenance of recombination suppression by assuming that recombination cannot be reestablished due to constraints on the relevant genetic variation. We have shown here that this theory faces strong difficulties. Mechanistically, recombination can be restored by many well described mechanisms, and even extremely low restoration rates are sufficient to prevent Y degeneration. Theoretically, deleterious mutations alone cannot explain the evolution of sex-chromosomes without leading to extreme male-female fitness imbalance or population extinction, which inevitably occurs when rates of recombination restoration are assumed to be extremely low.

Acknowledgments

We thank Aline Muyle for useful comments. This work is supported by the ANR Genasex and CisTransEvol.

Code availability

The simulation code is available at:

T. Lenormand, D. Roze, Zenodo (2021), doi:10.5281/zenodo.5504423.

References

1. Barton NH, Partridge L. Limits to natural selection. BioEssays. 2000;22: 1075–1084.

2. Crespi BJ. The evolution of maladaptation. Heredity. 2000;84: 623–629.

3. Brady SP, Bolnick DI, Angert AL, Gonzalez A, Barrett RDH, Crispo E, et al. Causes of maladaptation. Evolutionary Applications. 2019;12: 1229–1242. doi:10.1111/eva.12844

4. Roff D a., Fairbairn DJ. The evolution of trade-offs: where are we? Journal of Evolutionary Biology. 2007;20: 433–447. doi:10.1111/j.1420-9101.2006.01255.x

5. Burt A, Trivers R. Genes in Conflict: The Biology of Selfish Genetic Elements. Boston: Harvard University Press; 2006.

6. Lenormand T, Roze D, Rousset F. Stochasticity in evolution. Trends in Ecology & Evolution. 2009;24: 157–165. doi:10.1016/j.tree.2008.09.014

7. Lenormand T. Gene flow and the limits to natural selection. Trends in Ecology and Evolution. 2002;17: 183–189. doi:10.1016/S0169-5347(02)02497-7

8. Gordo I, Charlesworth B. The degeneration of asexual haploid populations and the speed of Muller's ratchet. Genetics. 2000;154: 1379–1387.

9. Lynch M, Conery J, Bürger R. Mutation accumulation and the extinction of small populations. The American Naturalist. 1995;146: 489–518.

10. Lande R. Risk of population extinction from fixation of new deleterious mutations. Evolution. 1994;48: 1460–1469.

11. Cayuela H, Lemaître J-F, Léna J-P, Ronget V, Martínez-Solano I, Muths E, et al. Sex-related differences in aging rate are associated with sex chromosome system in amphibians. Evolution. 2022;76: 346–356. doi:10.1111/evo.14410

12. Marais GAB, Lemaître J-F. Sex chromosomes, sex ratios and sex gaps in longevity in plants. Philosophical Transactions of the Royal Society B: Biological Sciences. 2022;377: 20210219. doi:10.1098/rstb.2021.0219

13. Xirocostas ZA, Everingham SE, Moles AT. The sex with the reduced sex chromosome dies earlier: a comparison across the tree of life. Biology Letters. 2020;16: 20190867. doi:10.1098/rsbl.2019.0867

14. Rice WR. Sex chromosomes and the evolution of sexual dimorphism. Evolution. 1984;38: 735–742.

15. Rice W. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. Evolution. 1987;41: 911–914.

16. Bull JJ. Evolution of sex determining mechanisms. Menlo Park, CA: Benjamin Cummings; 1983.

17. Charlesworth B. The Evolution of Sex-Chromosomes. Science. 1991;251: 1030–1033.

436 18. Charlesworth D. Evolution of recombination rates between sex chromosomes. Phil Trans R Soc
437 B. 2017;372: 20160456. doi:10.1098/rstb.2016.0456

438 19. Charlesworth B. Model for evolution of Y chromosomes and dosage compensation. Proceedings
439 of the National Academy of Sciences of the United States of America. 1978;75: 5618–5622.

440 20. Bachtrog D. Y-chromosome evolution: Emerging insights into processes of Y-chromosome
441 degeneration. Nature Reviews Genetics. 2013;14: 113–124. doi:10.1038/nrg3366

442 21. Lenormand T, Fyon F, Sun E, Roze D. Sex chromosome degeneration by regulatory evolution.
443 Current Biology. 2020;30: 3001-3006.e5. doi:10.1016/j.cub.2020.05.052

444 22. Lenormand T, Roze D. Y recombination arrest and degeneration in the absence of sexual
445 dimorphism. Science. 2022;375: 663–666. doi:10.1101/2021.05.18.444606

446 23. Jay P, Tezenas E, Véber A, Giraud T. Sheltering of deleterious mutations explains the stepwise
447 extension of recombination suppression on sex chromosomes and other supergenes. PLOS
448 Biology. 2022;20: e3001698. doi:10.1371/journal.pbio.3001698

449 24. Charlesworth D. When and how do sex-linked regions become sex chromosomes? Evolution.
450 2021;75: 569–581. doi:10.1111/evo.14196

451 25. Ponnikas S, Sigeman H, Abbott JK, Hansson B. Why do sex chromosomes stop recombining?
452 Trends in Genetics. 2018;34: 492–503. doi:10.1016/j.tig.2018.04.001

453 26. Jeffries DL, Gerchen JF, Scharmann M, Pannell JR. A neutral model for the loss of recombination
454 on sex chromosomes. Philosophical Transactions of the Royal Society B: Biological Sciences.
455 2021;376. doi:10.1098/rstb.2020.0096

456 27. Lenormand T. The evolution of sex dimorphism in recombination. Genetics. 2003;163: 811–822.

457 28. Charlesworth B, Wall JD. Inbreeding, heterozygote advantage and the evolution of neo–X and
458 neo–Y sex chromosomes. Proc R Soc Lond B. 1999;266: 51–56. doi:10.1098/rspb.1999.0603

459 29. de Waal Malefijt M, Charlesworth B. A model for the evolution of translocation heterozygosity.
460 Heredity. 1979;43: 315–331. doi:10.1038/hdy.1979.85

461 30. Grossen C, Neuenschwander S, Perrin N. The evolution of XY recombination: Sexually
462 antagonistic selection versus deleterious mutation load. Evolution. 2012;66: 3155–3166.
463 doi:10.1111/j.1558-5646.2012.01661.x

464 31. Blaser O, Grossen C, Neuenschwander S, Perrin N. Sex-Chromosome Turnovers Induced by
465 Deleterious Mutation Load. Evolution. 2013;67: 635–645. doi:10.1111/j.1558-
466 5646.2012.01810.x

467 32. Blaser O, Neuenschwander S, Perrin N. Sex-Chromosome Turnovers: The Hot-Potato Model. The
468 American Naturalist. 2014;183: 140–146. doi:10.1086/674026

469 33. Ranz JM, Casals F, Ruiz A. How Malleable is the Eukaryotic Genome? Extreme Rate of
470 Chromosomal Rearrangement in the Genus Drosophila. Genome Res. 2001;11: 230–239.
471 doi:10.1101/gr.162901

472 34. Navarro A, Betrán E, Barbadilla A, Ruiz A. Recombination and gene flux caused by gene
473 conversion and crossing over in inversion heterokaryotypes. Genetics. 1997;146: 695–709.
474 doi:10.1093/genetics/146.2.695

475 35. Koury SA. Predicting recombination suppression outside chromosomal inversions in Drosophila
476 melanogaster using crossover interference theory. Heredity (Edinb). 2023. doi:10.1038/s41437-
477 023-00593-x

478 36. Pegueroles C, Ordóñez V, Mestres F, Pascual M. Recombination and selection in the maintenance
479 of the adaptive value of inversions. Journal of Evolutionary Biology. 2010;23: 2709–2717.
480 doi:10.1111/j.1420-9101.2010.02136.x

481 37. Wellenreuther M, Bernatchez L. Eco-Evolutionary Genomics of Chromosomal Inversions. Trends
482 in Ecology & Evolution. 2018;33: 427–440. doi:10.1016/j.tree.2018.04.002

483 38. Stevison LS, Hoehn KB, Noor MAF. Effects of Inversions on Within- and Between-Species
484 Recombination and Divergence. Genome Biology and Evolution. 2011;3: 830–841.
485 doi:10.1093/gbe/evr081

486 39. Levine RP. Crossing Over and Inversions in Coadapted Systems. The American Naturalist. 1956
487 [cited 6 Feb 2023]. doi:10.1086/281905

488 40. Ishii T, Nagai A, Hirono T, Nakamura K. Hibernation of mosquitoes in rock caves on Miyato island.
489 Sci Rep Tôhoku Univ Ser IV. 1964;30: 159–165.

490 41. Korunes KL, Noor MAF. Pervasive gene conversion in chromosomal inversion heterozygotes. Mol
491 Ecol. 2019;28: 1302–1315. doi:10.1111/mec.14921

492 42. Andolfatto P, Depaulis F, Navarro A. Inversion polymorphisms and nucleotide variability in
493 Drosophila. Genetics Research. 2001;77: 1–8. doi:10.1017/S0016672301004955

494 43. Li W-H, Gu Z, Cavalcanti ARO, Nekrutenko A. Detection of gene duplications and block
495 duplications in eukaryotic genomes. J Struct Funct Genomics. 2003;3: 27–34.

496 44. Watanabe Y, Takahashi A, Itoh M, Takano-Shimizu T. Molecular Spectrum of Spontaneous de
497 Novo Mutations in Male and Female Germline Cells of Drosophila melanogaster. Genetics.
498 2009;181: 1035–1043. doi:10.1534/genetics.108.093385

499 45. Turner DJ, Miretti M, Rajan D, Fiegler H, Carter NP, Blayney ML, et al. The rates of de novo meiotic
500 deletions and duplications causing several genomic disorders in the male germline. Nat Genet.
501 2008;40: 90–95. doi:10.1038/ng.2007.40

502 46. Lipinski KJ, Farslow JC, Fitzpatrick KA, Lynch M, Katju V, Bergthorsson U. High Spontaneous Rate
503 of Gene Duplication in Caenorhabditis elegans. Curr Biol. 2011;21: 306–310.
504 doi:10.1016/j.cub.2011.01.026

505 47. Charlesworth D, Bergero R, Graham C, Gardner J, Keegan K. How did the guppy Y chromosome
506 evolve? PLOS Genetics. 2021;17: e1009704. doi:10.1371/journal.pgen.1009704

507 48. Sharma A, Heinze SD, Wu Y, Kohlbrenner T, Morilla I, Brunner C, et al. Male sex in houseflies is
508 determined by Mdmd, a paralog of the generic splice factor gene CWC22. Science. 2017;356:
509 642–645. doi:10.1126/science.aam5498

510 49. Meisel RP, Olafson PU, Adhikari K, Guerrero FD, Konganti K, Benoit JB. Sex Chromosome
511 Evolution in Muscid Flies. G3 Genes|Genomes|Genetics. 2020;10: 1341–1352.
512 doi:10.1534/g3.119.400923

513 50. Furman BLS, Metzger DCH, Darolti I, Wright AE, Sandkam BA, Almeida P, et al. Sex Chromosome
514 Evolution: So Many Exceptions to the Rules. Genome Biology and Evolution. 2020;12: 750–763.
515 doi:10.1093/gbe/evaa081

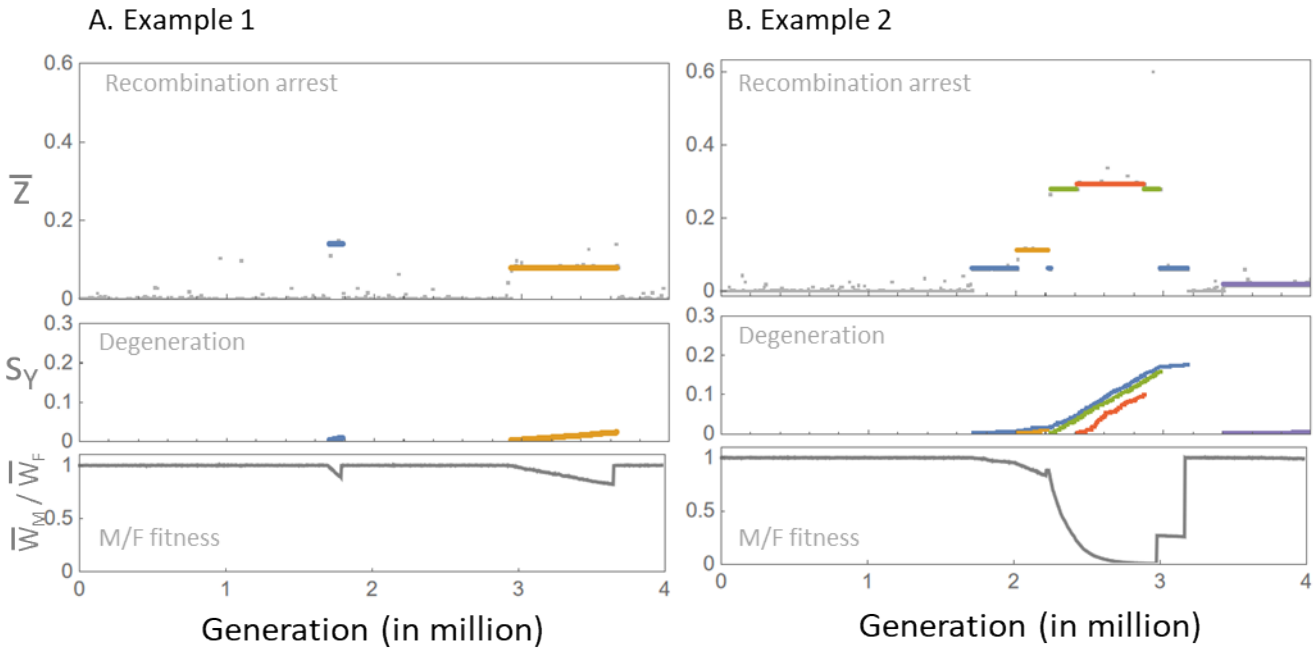

Fig 1. Examples of inversion-reversion dynamics. Each example is illustrated with three panels. The top panel shows the average non recombining fraction of the Y in the population ($\bar{z}$) through time (x axis in million generations). Color lines correspond to fixed inversions (i.e., all inversions reaching a frequency of 1 during the simulation); different fixed inversions have a different color. The color line extends between the time of occurrence of the inversion and the time when it becomes extinct. The middle panel shows the (per gene) average cumulated fitness effect of deleterious mutations on these inversions through time (same color code as in top panel). The bottom panel shows the ratio of the average fitness of males / females in the population through time.

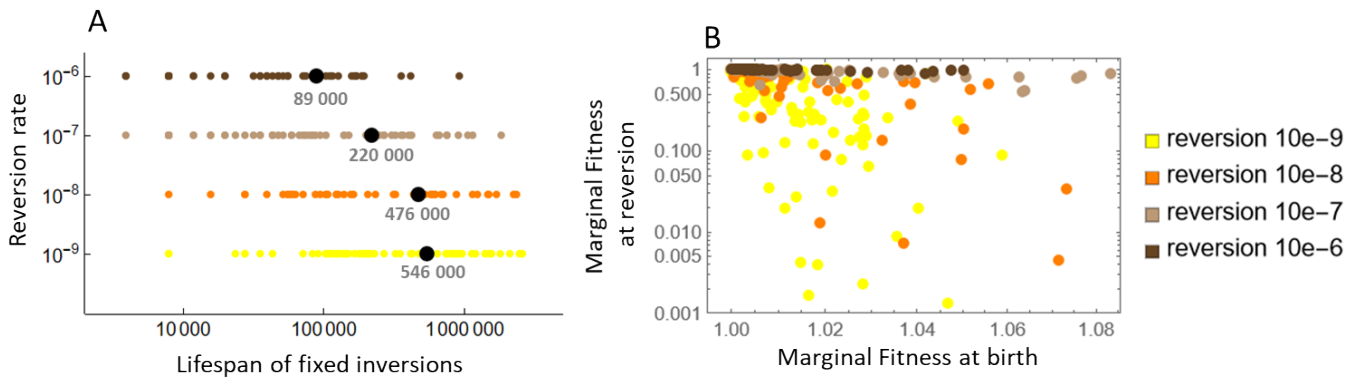

Fig 2. Characteristics of inversions in the shelter scenario. Panel A indicates the lifespan of fixed inversions (taken on 20 replicates) for different reversion rates (on y-axis). Time was cut at the end of the simulation (after 4 million generations) or if the population became extinct (by reaching a male/female fitness ratio < 0.001). Panel B shows the marginal fitness at birth of inversions (x-axis) versus marginal fitness at last recorded time (y-axis, log scale). The latter most often corresponds to a reversion, but in a few cases it corresponds to the end of the simulation, or population extinction. Color codes correspond to different reversion rates as indicated in the legend.

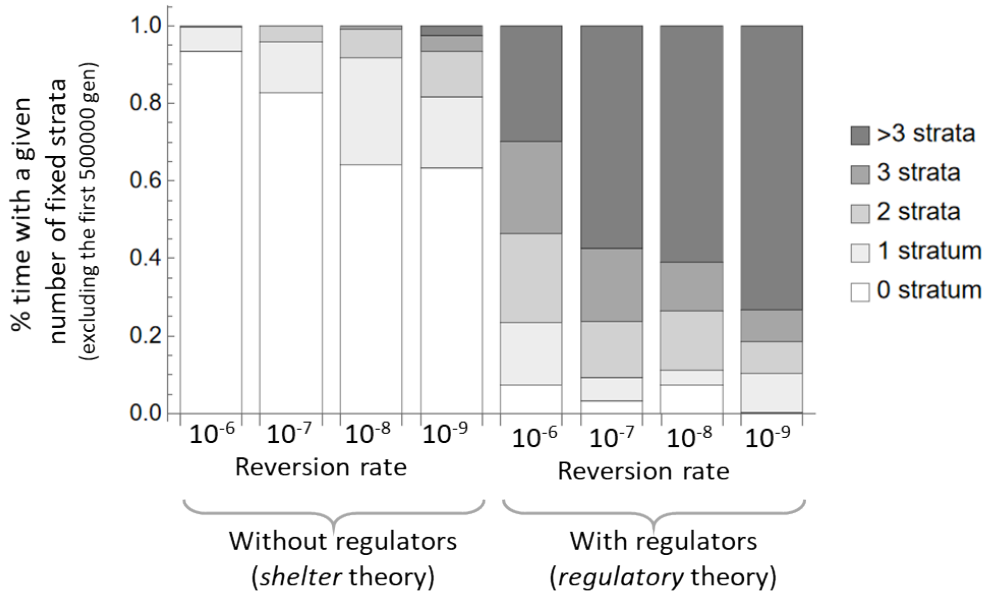

Fig 3. Fraction of the time during which the Y carries a given number of fixed strata. During a replicate simulation, there are times during which an inversion is fixed, perhaps including several strata, and times during which no fixed inversion is present (see examples on Fig 1). The bar chart gives this %time, across all replicates when a given number of strata are fixed in the population, for different reversion rates (without regulator evolution as in the shelter scenario, four bars on the left, or with regulator evolution as in the regulatory scenario, four bars on the right). This % is computed excluding the first 500 000 generation (to cut the initial phase influenced by the initial condition where the chromosome starts fully recombining and without any fixed inversion). The gray level corresponds to the number of fixed strata present as indicated in the legend on the right. For instance, in the shelter scenario with reversion rate equal to 10^{-6}, there is no fixed inversion 93.5% of the time, and one fixed inversion 6.3% of the time. In contrast, in the regulatory scenario with the same reversion rate, there is no fixed inversion 7.3% of the time, one fixed inversion 16% of the time, 2 fixed inversions 23% of the time, 3 fixed inversions 23.7% of the time and more than 3 fixed inversions 30% of the time (over the 3.5 million last generations of a simulation).

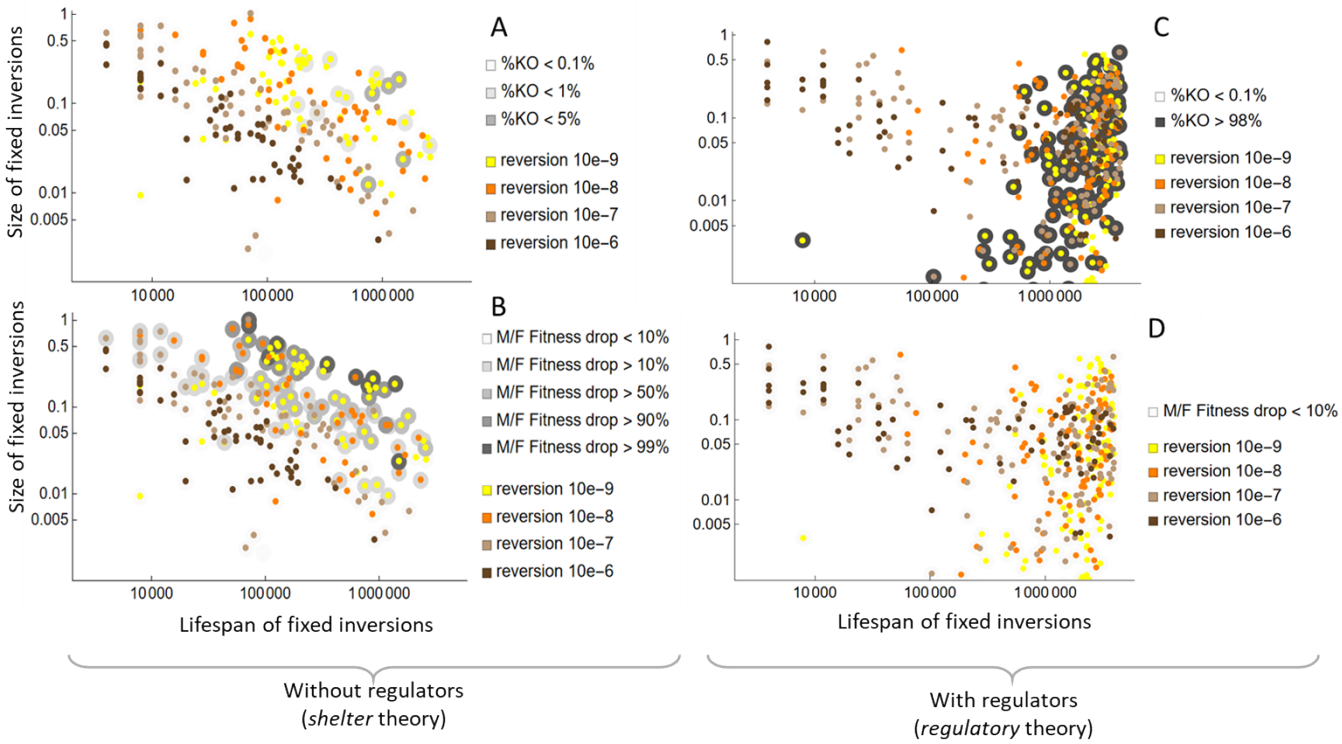

Fig 4. Detailed characteristics of fixed inversions in the shelter (panels A, B) and regulatory (panels C, D) scenarios. In all panels, the x-axis gives the lifespan of inversions (as defined in Fig 2A) and the y-axis the size of the inversion (both in log scale). Each dot represents a different fixed inversion that occurred across all replicates. The color of the dot indicates the reversion rate of the simulation during which the inversion was observed (as given in the legend). On panels A and C, a gray disk is added around each inversion. The gray level indicates the % of gene knocked out (%KO) on that inversion at the last time when the inversion is observed (i.e., just before it is reversed, at the end of the simulation at 4 million generations, or at population extinction). On panel A (shelter scenario), the gray level is light as this %KO never exceeds 5%. On panel C (regulatory scenario), this gray level is darker as this %KO reaches very high values (being either close to zero or above 98%). On panel B and D, a gray disk is also added around each inversion, this time representing the drop in the male / female fitness ratio caused by this inversion between the first and last time it is observed. Noting r(t) this male / female ratio at time t, this drop is computed as $r(t_1)/r(t_2)$ between times t_1 and t_2. When several inversions are simultaneously present in a given time interval, the log(drop) is portioned proportionally to the relative size of each inversion $s_1/(s_1+s_2)$, i.e. with two inversions of size s_1 and s_2, the drop accrued to the first inversion is Exp(s1/(s1+s2) Log $r(t_1)/r(t_2)$). For instance, with two inversions of equal size, each is assigned the square root of the fitness drop on the interval (so that the product of the fitness drop of each inversion gives the overall fitness drop). With this correction, the fitness drop associated to an inversion is more representative of what is happening on this inversion (rather than being caused by the presence of another inversion). On panel B (shelter scenario), these fitness drops can reach large

values (more than 99% reduction in the male/female fitness ratio), but they remain very low in the regulatory scenario (never exceeding 10% on panel D).

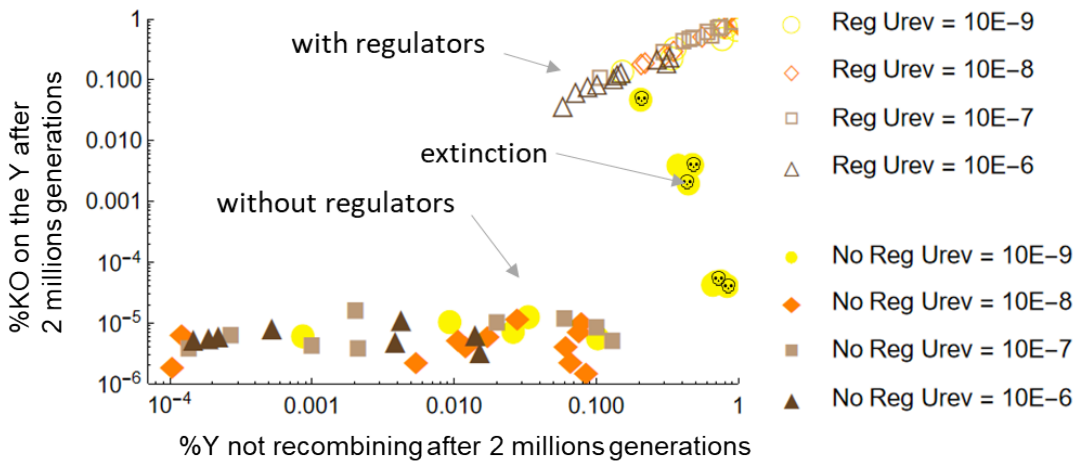

Fig 5. Overall evolution of the Y chromosomes in the shelter and regulatory scenarios. The x-axis (in log scale) gives the fraction of the Y that is non recombining (averaged over all males in the population) after two million generations. The y-axis (in log scale) gives the fraction of knock-out genes on the Y (averaged over all males in the population; a knock out is defined as a gene having accumulated deleterious mutations up to s_{max} = 0.3), after two million generations. Each dot represents a replicated population. Open symbols: regulators evolve (regulatory scenario); filled symbol: regulators do not evolve (shelter scenario). Color codes indicates different rates of reversion (U_{rev}). The rate of inversion is 10^{-5} in all cases. In a few cases (with U_{rev} = 10^{-9} in the shelter scenario, filled yellow disks) the population became extinct before 2 million generations. In these cases (marked with a skull), the x and y axes values are taken at the time of extinction.

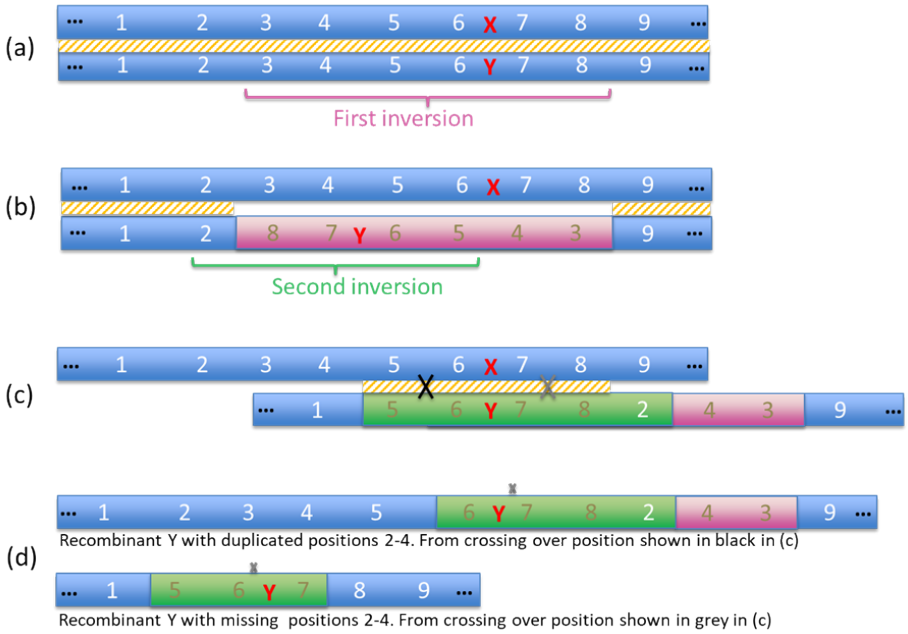

Fig. 6. Restoration of recombination on the Y with overlapping inversions. (a) depicts the fully recombining XY pair with the SDL in red. The orange dashing indicates where homologous pairing is possible. A first inversion occurs (purple) on the Y between positions 3 and 8 leading to the situation in (b). Pairing and recombination do not occur around the SDL. This first stratum on the Y chromosome can start degenerating in the absence of recombination. This is shown by the brown color of the position numbers. A second overlapping inversion occurs (green) on the Y between positions 2 and 5. The resulting Y in (c) can now pair with the X between positions 5 and 8. Crossing-over on the left of the SDL (black cross) can generate recombinant Y with duplicated positions 2-4 (note that this Y recovers functional copies in positions 3-5, and has two functional copies in position 2). Crossing-over on the right of the SDL (gray cross) can generate recombinant Y with missing positions 2-4 (note that the lack of these positions may not reduce fitness a lot if they were degenerated). In both cases, a second crossing-over with the X occurring near the SDL (shown with a small cross) can restore full collinearity with the X and get rid of other degenerated copies (on the right and left position of the Y, respectively).

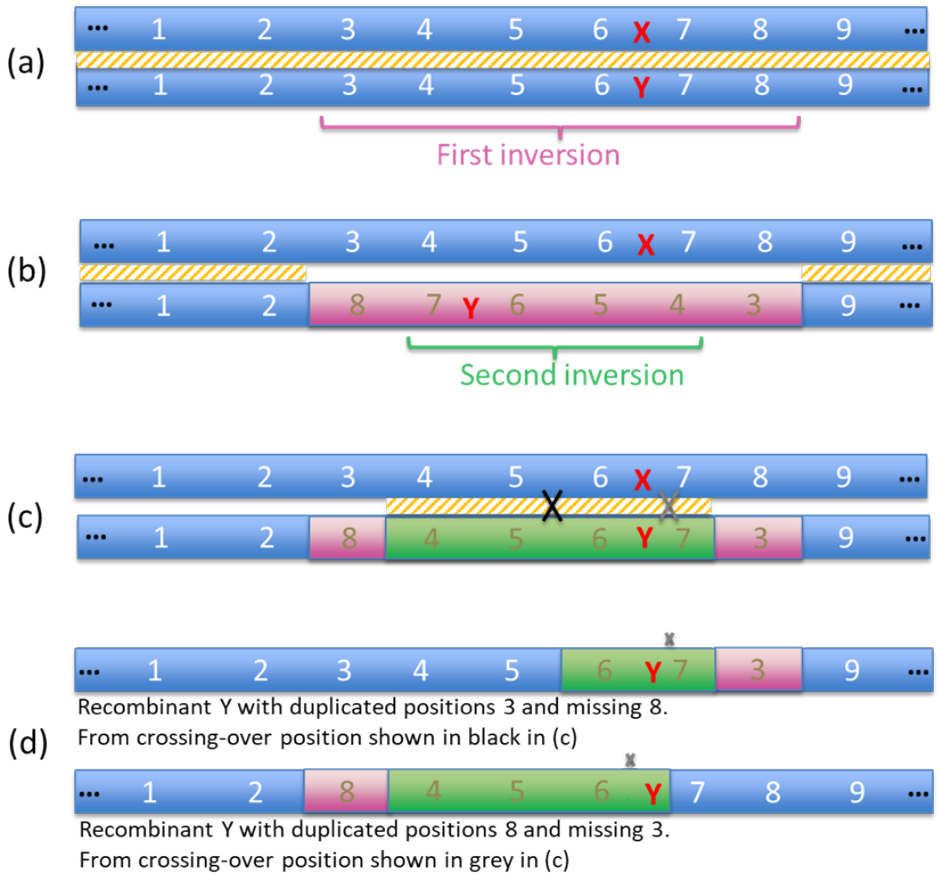

Fig. 7. Restoration of recombination on the Y with nested inversions. (a) depicts the fully recombining XY pair with the SDL in red. The orange dashing indicates where homologous pairing is possible. A first inversion occurs (purple) on the Y between positions 3 and 8 leading to the situation in (b). Pairing and recombination do not occur around the SDL. This first stratum on the Y chromosome can start degenerating in the absence of recombination. This is shown by the brown color of the position numbers. A second nested inversion occurs (green) on the Y between positions 4 and 7. The resulting Y in (c) can now pair with the X between positions 4 and 7. Crossing-over on the left of the SDL (black cross) can generate a recombinant Y with duplicated positions 3 and missing 8. Note that this Y can have a higher fitness by recovering functional copies at positions 3-5. Position 8 is missing, but this may not be consequential since it was a degenerated position. Crossing-over on the right of the SDL (gray cross) can generate a recombinant Y with duplicated position 8 and missing 3. Again, the fitness of this Y may increase, since it recovers functional copies at positions 7-8, while losing a degenerated copy at position 3. In both cases, a second crossover with the X occurring near the sex-determining locus (shown with a small cross) can reconstitute a Y chromosome fully colinear with the X.

633

Parameters	**Values**
Population size	10000
Standard deviation of mutational effects on cis and trans regulatory traits	0.2
Number of gene, cis-regulator, male-limited and female-limited trans-regulators	500
Average effect of deleterious mutations (their distribution is exponential)	0.05
Baseline dominance level of deleterious mutations	0.25
Maximum fitness effect on a gene, corresponding to a knock-out (s_{max})	0.3
Intensity of stabilizing selection	0.1
Mutation rates on genes	0.0002
Mutation rates on cis-regulators (set to zero in shelter theory)	0.0002
Mutation rates on trans-regulators (set to zero in shelter theory)	0.0001
Recombination rate between consecutive genes (R_g)	0.0005
Recombination rate between cis-regulator and its gene (R_c)	0.00005
Rate of inversion mutation	0.00001
Rate of reversion mutations (10 to 10^4 lower than rate of inversions)	variable

634 **Table 1. Parameter values used in simulation.**

635