Sex Chromosome Degeneration by Regulatory Evolution

Thomas Lenormand, Frédéric Fyon, Eric Sun, Denis Roze

Sex chromosome degeneration by regulatory evolution

Thomas Lenormand[1,2,§], Fredric Fyon[1], Eric Sun[2], Denis Roze[3,4]

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

2 Radcliffe Institute, Harvard University, Cambridge, MA, USA

3 CNRS, UMI 3614, Roscoff, France

4 Sorbonne Université, Station Biologique de Roscoff, France

§ Corresponding author: thomas.lenormand@cefe.cnrs.fr

Keywords: sex chromosome; population genetics theory; *cis*-regulators; degeneration; selective interference; dosage compensation

Highlights

- A new theory for Y degeneration, not based on selective interference
- Initiated by X and Y *cis*-regulator divergence after the arrest of recombination
- Works faster and in larger populations than current theory
- Works on small non-recombining region on the Y

eTOC Blurb

Y chromosomes are often degenerate. This remarkable feature of the genome of plants and animals has been intensely investigated. Current theory proposes that this degeneration occurs from "selective interference" after the arrest of recombination between the X and Y chromosome, meaning that natural selection tends to be inefficient in the absence of recombination. This theory has been in place for more than 40 years. In this paper, we propose a

new theory (termed degeneration by regulatory evolution) based on the instability of diploid expression between non-recombining chromosomes. We show that this theory can explain fast Y degeneration (especially when selective interference and regulatory evolution co-occur) and simultaneous Y silencing and dosage compensation. It is open to tests using genomic and transcriptomic methods that have been recently developed in many plant and animal species.

Abstract

In many species, the Y (or W) sex chromosome is degenerate. Current theory proposes that this degeneration follows the arrest of recombination and results from the accumulation of deleterious mutations due to selective interference--the inefficacy of natural selection on non-recombining genomic regions. This theory requires very few assumptions, but does not robustly predict fast erosion of the Y (or W) in large populations or the stepwise degeneration of several small non-recombining strata. We propose a new mechanism for Y/W erosion that works over faster timescales, in large populations, and for small non-recombining regions (down to a single sex-linked gene). The mechanism is based on the instability and divergence of *cis*-regulatory sequences in non-recombining genome regions, which become selectively haploidized to mask deleterious mutations on coding sequences. This haploidization is asymmetric, because *cis*-regulators on the X cannot be silenced (otherwise there would be no expression in females). This process causes rapid Y/W degeneration and simultaneous evolution of dosage compensation, provided that autosomal *trans*-regulatory sequences with sex-limited effects are available to compensate for *cis*-regulatory divergence. Although this “degeneration by regulatory evolution” does not require selective interference, both processes may act in concert to further accelerate Y degeneration.

Main

The contemporary theory for the evolution of sex chromosome crystallized in the 1970s [1–3] and applies to both XX/XY and ZZ/ZW sex determination systems, which share important convergent similarities [4]. Both cases involve a chromosome that is heterozygous (the Y or W) and present in only one sex. Although a broad range of situations has been described, in many

cases, most of the chromosome has stopped recombining and has degenerated considerably. Current theory, which we term “Degeneration by selective interference” (DSI) has been substantially refined since the 70s’, but its core idea--degeneration caused by selective interference--has remained unchanged [3,5–8]. The theory about selective interference, as well as its empirical evaluation, has also been largely developed since the 70s’, well beyond the case of Y degeneration [9–11]. DSI involves a sequence of steps that occur after the arrest of recombination between the Y and X chromosomes (all our arguments also apply to Z/W chromosome system): (a) degeneration of Y-linked genes by 'selective interference' (also known as the 'Hill-Robertson effect'), due to processes such as Muller's Ratchet, hitchhiking, and background selection [8], (b) facultatively, adaptive silencing of Y-linked genes, and (c) evolution of dosage compensation. A variant of this theory proposes that the accumulation of deleterious alleles in regulatory sequences by selective interference leads to reduced Y gene expression [12]. Y-linked alleles, being partially hidden would then further accumulate deleterious mutations and degenerate [13].

In this paper, we propose a new “Degeneration by Regulatory Evolution” (DRE) theory to explain Y chromosome degeneration. The main differences from the DSI model are that our theory does not require selective interference, and that steps a-c occur simultaneously after recombination suppression. We previously showed that, for autosomal genes, a ‘*cis*-regulator runaway’ process occurs that leads stronger *cis*-regulators to become associated with chromosomes with fewer deleterious mutations [14]. This favors the stronger *cis*-regulatory alleles, provided they are tightly linked to their coding gene. We also showed that *cis*-regulators diverge in asexuals, where diploid expression is unstable and quickly becomes “haploidized” [15]. DRE theory involves such divergence of *cis*-regulators, but with an asymmetry between the X and Y chromosomes (preventing the suppression of gene expression on the X). We investigate DRE using individual-based stochastic simulations of a population of N_{pop} diploid individuals, with *XY* males and *XX* females. In order to capture the essence of the mechanism, we first study a minimal system with only four loci (Fig 1): a gene *G*, its *cis*-regulator *C*, and two *trans*-regulators *T* (we will then extend the model to the case of a non-recombining region comprising a larger number of genes). *Trans*-regulators, such as transcription factors, are not closely linked to their target gene,

and influence expression on both homologs, whereas c*is*-regulators, such as enhancers, control the expression of the closely linked gene, and influence only the copy carried on the same chromosome as themselves [16]. We assume that *G* and *C* are present on both sex chromosomes, and that they recombine (in females only) at a rate R_c. We include *trans*-regulators in order to examine whether, and over what time-scale, dosage compensation will evolve when expression of the Y-linked allele decreases. With these *trans*-regulators, overall expression can be maintained, i.e. dosage compensated, even if *cis*-regulators change and diverge between the X and Y. Dosage compensation cannot evolve if the *trans*-regulators act in the same way in both males and females, or equally on the X and the Y. Hence we do not consider all potential trans-regulators, but only those that could influence dosage compensation. Specifically, we focus on the simplest symmetrical case with one trans-regulator expressed in males T_m, and one in females T_f. Both cases have been described empirically (*C. elegans* dosage compensation works by halving X expression in females, whereas in *Drosophila*, it works by doubling X expression in males [17]). For simplicity, we assume that these *T* loci are autosomal and that they recombine freely with each other and with the *G* and *C* loci.

This initial model is later extended to n_L CGT_mT_f quadruplets of genes (where n_L = 1, 50, 500). In these models, we assume that the *C* and *G* loci are uniformly spaced on the sex chromosomes, with two adjacent genes *G* recombining at a rate R_g in females, and where each *C* locus is assumed to be closer to the *G* gene it regulates than to the next *G* gene (ten times closer in the simulations,); again, recombination is assumed not to occur in males, representing a non-recombining region including the sex-determining locus, while all *T* factors are again assumed to recombine freely with the sex-determining region. In this model, each *CG* pair is influenced by its own T_mT_f pair, which represents the lowest degree of pleiotropy of these *trans*-acting factors, but involves a very high number of *trans*-regulators. We also considered a model where only one T_mT_f pair controls all the *G* and *C* loci, representing the other extreme case where *trans*-regulators are maximally pleiotropic and influence many (here all n_L) genes.

Deleterious mutations occur within genes *G* at a rate U_G per gene. Their fitness effect *s* is drawn from an exponential distribution with mean s_{mean}. The effects of multiple mutations in the same gene are assumed to be additive, but with a maximum effect per gene, s_{max} (which may be

interpreted as the fitness effect of a full gene knock-out). Their dominance depends on the strength of their associated *cis*-regulator (see methods). The effects of alleles at the *cis*- (C) and *trans*-regulators (T_m, T_f) are modelled as quantitative traits denoted by c, t_m, t_f, respectively, and control the level of expression Q of the gene, which is under stabilizing selection with intensity I (methods). The different events of the life cycle occur in the following order: diploid selection, meiosis with recombination, mutation, and syngamy. Simulations are initialized with no polymorphism present, and the optimal gene expression level (no deleterious allele, all c and t_m, t_f alleles fixed to 1). After a burn-in phase where the chromosome evolves with recombination in both sexes, we stop XY recombination in males to create a sex-linked region and follow the frequencies and effects of deleterious mutations on the X and Y, as well as the evolution of the regulatory genes. These outputs are averaged over different numbers of replicates depending on the variance in the process under different parameter values (methods, Table S1). At regular intervals, we compute $P_{halfsilent}$, the probability (across replicates) that Y allele-specific expression $\phi_{Y,i}$ decreased by two fold from the initial value of 0.5 to $\phi_{Y,i} < 0.25$. A complementary approach to quantify partial silencing is to measure the dominance coefficient of deleterious alleles on the Y, measured by $h_{Y,i} = \phi_{Y,i}^{-\ln h/\ln 2}$ (see Methods). We also compute P_{silent}, the probability that $\phi_{Y,i}$ becomes close to zero (below 0.01), so that alleles on the Y become nearly entirely recessive. The quantity $P_{halfdead}$ then refers to the probability that, by a given number of generations after the Y-linked region stopped recombining, deleterious mutations on the Y gene copy have reduced fitness by an amount $s_{max}/2$, and P_{dead} that they reduced fitness by an amount s_{max}, indicating that the gene has entirely degenerated on the Y.

Fig **2** illustrates the process with n_L =1 (one gene, one *cis*-regulator, one male limited and one female limited *trans*-regulator). The system does not generate any male-female or X-Y asymmetry before recombination arrest (methods, Fig S5). After recombination arrest, the gene carried by the Y degenerates: it becomes progressively recessive, as h_Y changes from h = 0.25 to zero, and accumulates deleterious mutations (the overall fitness effect of mutations present on the Y copy increases up to s_{max}), despite there being very limited selective interference (at most only occurring between the gene and its *cis*-regulator). Silencing occurs first, and the accumulation of deleterious mutations follows later in the process (the curve representing $P_{halfsilent}$ is ahead of the

one showing the accumulation of deleterious mutation on the Y, as measured by $P_{halfdead}$). Degeneration also occurs with full dosage compensation, and overall expression never departs from the optimum in either sex (Fig **2b**). Compensation typically involves, at least initially, a mixture of upregulation of X gene copies in males and downregulation in females (methods, Fig. S7).

What is the underlying cause of this asymmetrical degeneration? Once X and Y stop recombining, diploid expression becomes unstable. *Cis*-regulators on the X and Y can diverge, eventually leading to the haploidization of expression in males. This is not prevented by stabilizing selection on expression levels as long as *trans*-regulators can coevolve to maintain near optimal total expression in both sexes. When the strength of *cis*-regulators on the Y starts decreasing, the process is accelerated by a "haploidization" positive feedback loop. Indeed, weak Y *cis*-regulators become associated with coding sequences carrying more deleterious mutations, as they cause a reduction in dominance. They are then selected to weaken further in order to mask those deleterious mutations, which leads to the accumulation of even more deleterious mutations, and so forth. By contrast, the other 'haploidized' situation (where the X is silenced in males) is reversible, as X chromosomes with weak *cis*-regulators and a higher load of deleterious mutations cannot fix, as they become homozygous and selected against in females when too frequent (unlike partially silenced Y genes, which can spread as they stay heterozygous in males). Therefore, the regulatory system has only one stable equilibrium, where the Y is silenced and degenerate.

Selective interference plays no role in this process, explaining why degeneration occurs even for a single Y-linked gene. However, the process is stochastic, as it is initiated by a random departure from diploid expression with a sufficiently weak Y *cis*-regulator to trigger the "haploidization" feedback loop. In individual simulation replicates, degeneration is indeed very abrupt, but occurs at varying time points (Fig **2a**). Because of this stochasticity, the process is slowed down in larger populations: it is ~10 times slower in ~10 times larger populations (Fig **2d**). With the same parameters but without mutation in the *cis*-regulator (U_c = 0, Fig **2c**), as in DSI theory, degeneration does not occur, as expected, since there is neither selective interference nor *cis*-regulatory divergence. Degeneration does not occur either in the absence of mutation in the

trans-regulators (U_t = 0, Fig **2c**) and for the same parameter values, since *cis* and *trans*-regulators have to coevolve to maintain total expression levels: if *trans*-regulators cannot evolve, the divergence of *cis*-regulators is prevented and Y degeneration cannot occur. However, if the intensity of stabilizing selection on expression levels is weak enough, degeneration evolves but is not dosage compensated (Fig **2c**). Otherwise, the intensity of stabilizing selection on dosage only plays a marginal role in DRE (Fig **2c**). Control simulations without mutation in the coding gene (U_G = 0, Fig **2c**) show that *cis*-regulatory divergence and Y silencing can occur even in the absence of deleterious mutations, but as expected, this silencing is slower, not being accelerated by the "haploidization" feedback loop, and reversible (methods, Fig S6).

As expected, DRE and DSI combine when more than a single Y-linked gene is considered. The effect is strong: Fig **3** shows that a 50-fold or 500-fold increase in the number of loci results in degeneration being 5-fold and 10-fold faster, respectively: a larger non-recombining Y-linked region degenerates faster than a small one. Without mutation in regulators (i.e. with only DSI), degeneration occurs but is 23-36 fold slower with 50 loci (depending on the control used for the comparison, methods, Fig 4). With 500 loci, however, the comparison with and without regulators is problematic, as very quickly, a modest accumulation of deleterious mutations on the many Y-linked genes causes an important reduction in male average fitness, so that male fitness reaches unrealistically low values (of the order 10^{-17}). Even if a proportion of genes affecting male fitness may be under soft selection, it seems unlikely that a population with such a low male fitness would survive.

The drop in male fitness is less dramatic with regulatory evolution, as those mutations become progressively more recessive as the Y degenerates. There is nevertheless a transient drop in male average fitness, which can be quite large (e.g. 3% and 85% reduction for a 50-gene and 500-gene Y-linked region, respectively, Fig **3**). Data on divergence between sex chromosomes indicate that Y degeneration is often sequential in chiasmate species, with several regions of various sizes, termed "strata" [18], having stopped recombining at different time points. This high transitory fitness drop may prevent large strata from occurring in small populations and may bias towards scenarios involving multiple small strata as, for example, in humans [19]. Comparatively, these

scenarios involving small strata are more difficult to explain with DSI, as selective interference is weak on small non-recombining regions.

Last, degeneration is initially slower but faster overall when there is only one T_m and T_f controlling all *cis*-regulators. Despite being highly pleiotropically constrained, having only two autosomal *trans*-regulators precipitates degeneration: the tipping point where it is worth fully silencing the Y is quickly reached when many very weakly deleterious mutations have accumulated on the Y. This is consistent with the observation that dosage compensation can occur locally on a gene-by-gene basis or by chromosome-wide *trans*-acting effects [17]. Intermediate cases involving *cis*- and *trans*-regulators with regional effects [as e.g. in Drosophila 20] may be worth investigating, but are likely to behave similarly, as long as *trans*-regulators only target X genes that have a copy on the Y non-recombining region.

Simulations for lower values of mutation rates and strength of selection against deleterious alleles are shown on Fig. S8 (for n_L = 1 and n_L = 50, methods). Unsurprisingly, reducing mutation rates slows degeneration, while the effect of s is more complicated. However, the acceleration of degeneration caused by regulatory evolution still holds. Fig. S8 also shows that a scaling argument from diffusion theory indicates that larger populations with weaker mutation and selection should also behave similarly, albeit on a longer timescale.

The DRE theory proposes a different view of sex-chromosome evolution compared to the DSI theory that has been developed over the past 40 years. In both cases, degeneration starts after the arrest of recombination in a genome region completely linked to the sex-determining locus. In both cases, degeneration is slower in larger populations, but this is considerably less so in the DRE model. However, there are important differences. With very few exceptions [21,22], DSI was developed without explicitly modelling regulatory evolution. With DSI, regulatory evolution (Y silencing and dosage compensation) is supposed to occur only after deleterious mutations have accumulated on the Y [3,23–25], although it has also been proposed that silencing may result directly from the accumulation of deleterious mutations in regulatory regions [13]. This is certainly possible, as selective interference applies to all functional sequences and may contribute to the fixation of many kinds of deleterious mutations, including those maintaining adequate expression levels. DRE is based on a reverse causality: regulatory evolution initiates the

degeneration process. Contrary to the standard model where compensation is needed because degeneration damages genes' function in males, compensation evolves here from the very beginning of the process, by reducing the proportion of transcripts from Y-linked relative to X-linked alleles and maintaining an almost constant overall level of expression in both sexes. However, compensation may not occur when expression levels are under weak stabilizing selection (Fig 2c). Whether compensation occurs on a gene-by-gene basis or chromosome wide depends on the availability of the corresponding *trans*-acting factor, but both can occur in DRE, and, surprisingly, at approximately the same rate (Fig 3).

The specific mode of dosage compensation depends on the type of *trans*-regulators. Here we considered *trans*-regulators with sex-limited expression, which can mimic several well-known dosage compensation systems (female-limited factors corresponding to *C. elegans* or mammal systems, while male-limited factors would be more similar to a *Drosophila*-like situation [17], methods). The symmetric DRE model that we used often led to a Drosophila-like compensation, where the X in males was eventually expressed twice as much compared to the situation at the recombination arrest (Fig **2b**). However, this is certainly dependent on the relative mutation rates on the different types of *trans*-regulators. The theory should be extended to examine the diversity of dosage compensation mechanisms, including sex-of-origin effects [26].

Overall, we have presented an alternative theory for the degeneration of sex chromosomes. Although many underlying parameters are still poorly known, this theory could be tested quantitatively, as it works faster than current theory, on smaller non-recombining regions, and does not require small population sizes or recurrent beneficial mutations causing hitchhiking effects. It does not exclude selective interference, which will necessarily co-occur as long as many genes stop recombining simultaneously. However, a hallmark of DRE is that regulatory changes occur very early. This is consistent with recent studies showing that dosage compensation evolves early on [27,28] and that Y transcriptional downregulation accompanies degeneration of protein-coding genes from the start [29–31].

Acknowledgments

We thank G. Marais, B and D. Charlesworth, S. P. Otto and two anonymous reviewers for insightful comments. We thank MBB cluster from Labex CEMEB, Harvard Odyssey cluster and CNRS ABiMs cluster. We thank J Wakeley for helping access Odyssey cluster. TL and DR acknowledge Radcliffe Institute, CNRS, and grant GenAsex ANR-17-CE02-0016-01 for support.

Author contributions

Original idea TL, FF; Model conception TL, DR, FF; Code DR, TL, ES; Simulations TL, ES; Data analyses TL; Interpretation TL, DR; First draft TL, FF; Editing TL, DR, ES; Reviews TL, DR; Project management and funding TL.

Declaration of interest

No conflict of interest

Fig 1

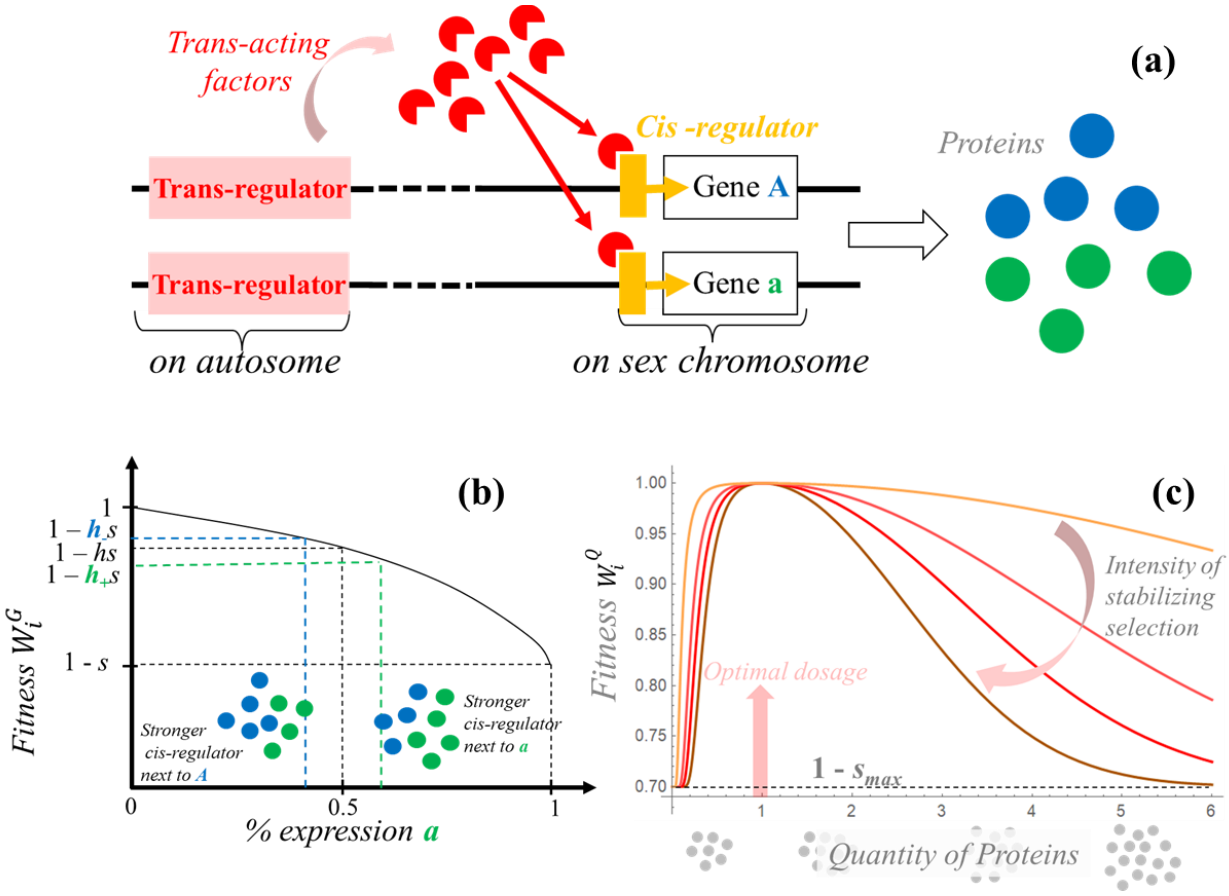

Fig 1. Model presentation. (**a**) The simplest genetic model involves autosomal *trans*-regulators expressed either in males or females, and a sex-linked *cis*-regulator controlling the expression of a coding gene. A deleterious allele **a** at the coding gene is expressed at the same level as the wild type **A** allele if their associated *cis*-regulators have equal strength. (**b**) If there is *cis*-regulatory variation, the deleterious allele may be over or under-expressed, depending on whether it is associated with the stronger or weaker *cis*-regulator. This can be considered as determining the dominance of the **a** allele. The black curve shows the fitness of an Aa heterozygote (y-axis) for varying allele-specific expression levels (x-axis). (**c**) Selection also acts on the total amount of protein produced (x-axis), with stabilizing selection around an optimal amount. The maximal fitness effect of a departure from the optimal amount is s_{max}, the same as the maximal fitness effect of a deleterious mutation in the coding gene.

Fig 2

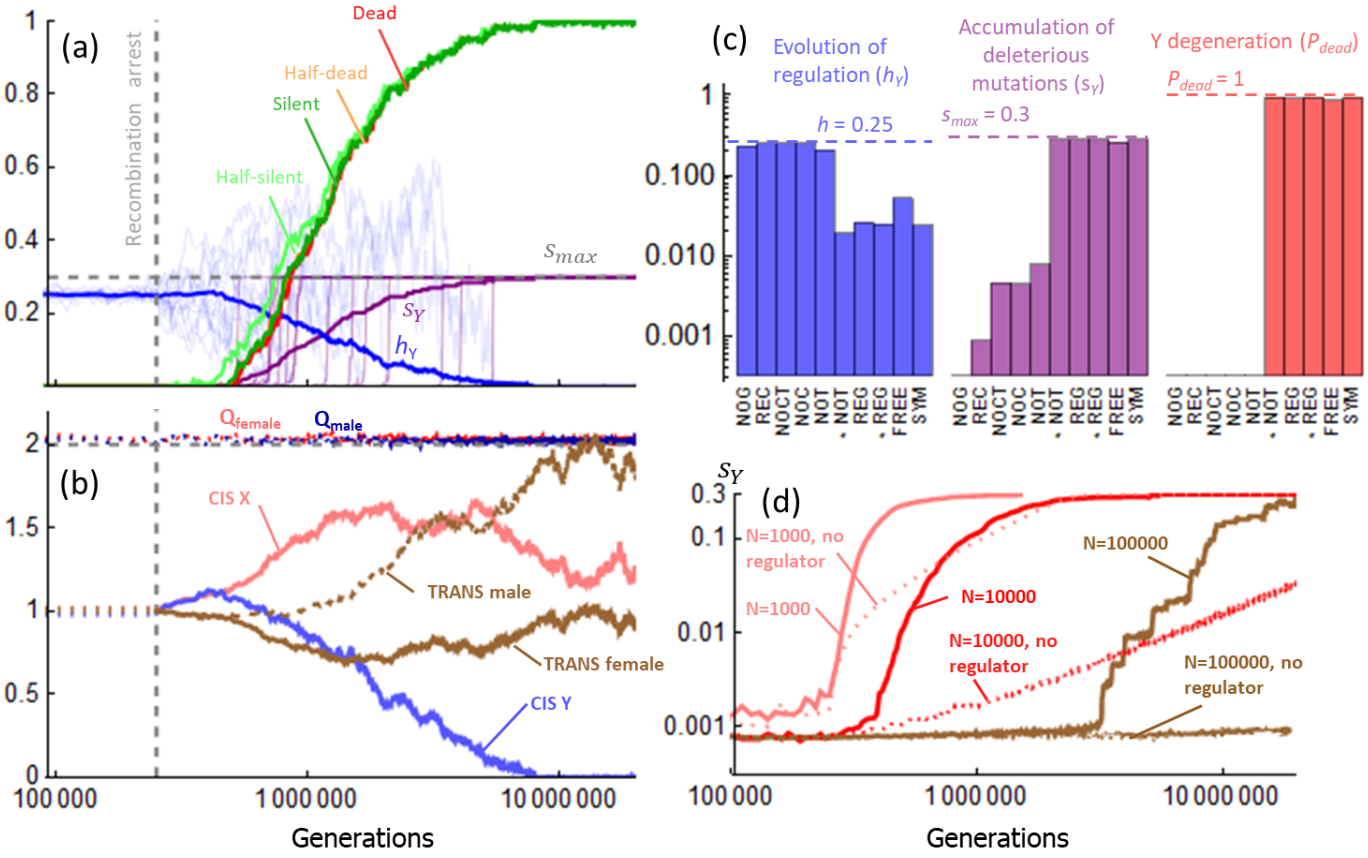

Fig 2. Y degeneration by regulatory evolution. **(a)** x-axis: time in number of generations, in log-scale. Recombination stops between the X and Y chromosome at generation 250000 (vertical dashed gray line); y-axis: The probability that a coding gene on the Y is expressed at less than half the level of the X copy (half-silent, light green curve), entirely silenced (dark green curve), has accumulated deleterious mutations reducing fitness effect half as much as a loss-of-function mutation (half-dead, orange curve), or as much as a full loss-of-function mutation (dead, red curve). Curves are averages over 100 replicates of the process. The model was run with n_L = 1 (one gene, one *cis*-regulator, one female and one male limited *trans*-regulator). Parameter values are: population size $N = 10^5$; Mutation rates U_G, U_C = 2 x 10^{-4}; U_T = 10^{-4} ; recombination rate between the gene and the *cis*-regulator R_C = 5 x 10^{-5} (in both sexes during the burn-in period, and in females thereafter). The mean effect of each deleterious mutation in the coding gene s_{mean} = 0.05 (and loss-of-function effect s_{max} = 0.3, dashed gray line). The dominance coefficient of

deleterious mutations (when both alleles are equally expressed) is h = 0.25, and the intensity of stabilizing selection on dosage I = 0.1. The figure also shows the dominance of deleterious mutations carried on the Y (h_Y, blue curve) and the average fitness effect of alleles on the Y (s_Y, purple curve). Some individual trajectories are indicated for h_Y and s_Y (same color code, thin lines) to show that degeneration occurs abruptly in each replicate. **(b)** Time-variation of regulatory traits corresponding to the case illustrated in panel a. x-axis: time in number of generations (in log-scale); y-axis: regulator trait values. Pink: X *cis*-regulator strength; Blue: Y *cis*-regulator strength; Brown: *trans*-regulator strength (plain: female limited, dashed: male-limited). Optimal dosage is 2 (dashed gray line). Total expression in males and females is indicated by the dotted curves (male value in blue, female in red). Regulatory trait values before recombination arrest are not stable due to runaway evolution[14], but are rescaled at 1 at generation 250000 for fair comparisons across parameter values. They are represented at this rescaled value on the figure, to avoid overloading the figure (see methods and Fig S5 for further details). **(c)** Values of s_Y, h_Y and P_{dead} at generation 3 x 10^6, for simulations like in panel (a) except for some parameter values. The star indicates weaker stabilizing selection (I =0.01) compared to I = 0.1 in panel (a). NOG (U_G = 0); REC (no arrest of recombination); NOCT (U_C = U_T = 0); NOT (U_T = 0); REG (with mutations on gene and regulators, as in panel a); FREE (R_c = 0.5); SYM (symmetrized stabilizing selection function, method). **(d)** Accumulation of deleterious mutations on the Y for different population sizes (10^3, 10^4 like in panel a, 10^5). x-axis: number of generations in log-scale. y-axis: s_Y. Plain lines: with evolving regulators. Dotted lines: without regulator evolution (NOCT simulations, U_C = U_T = 0).

Fig 3

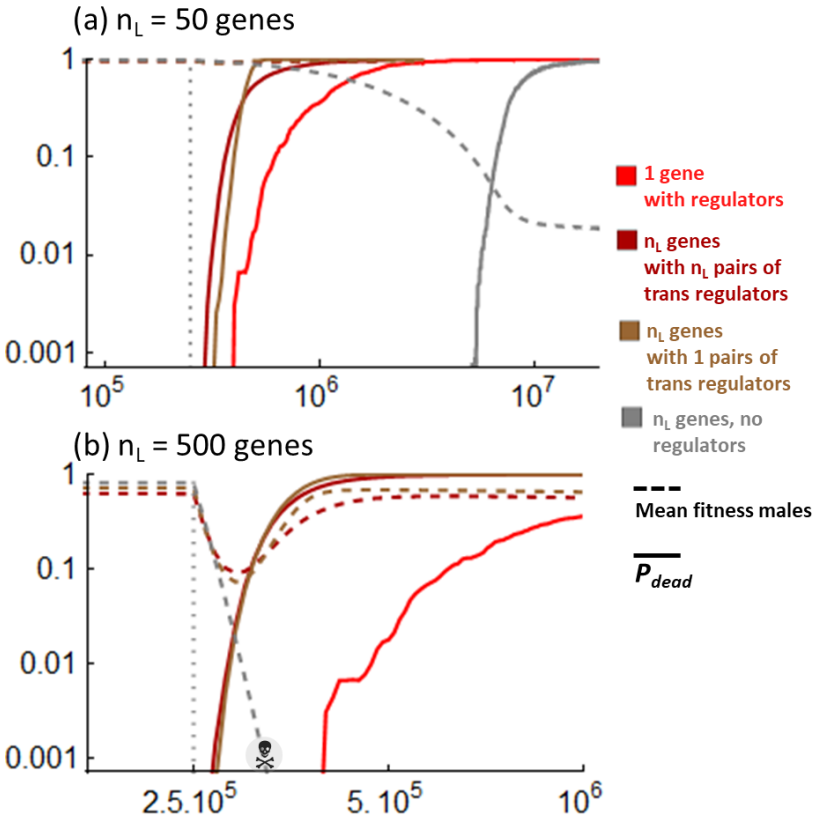

Fig 3. DRE with many loci. Y degeneration in non-recombining regions with n_L loci. **(a)** n_L = 50. **(b)** n_L = 500. Coding genes are positioned at regular interval on the X/Y chromosome, each at a recombination rate $R_G = 5 \times 10^{-4}$ (map length of the non-recombining region is therefore 2.5 or 25 cM, respectively). Recombination occurs in both sexes during the burn-in period (ending at 2.5 x 10^5 generations, vertical dotted line), and in females thereafter. *x*-axis: number of generation in log-scale. *y*-axis: degeneration of the Y, as measured by P_{dead}, in log-scale (plain lines) or male mean fitness (dashed lines). Red: the n_L = 1 case given for comparison. Dark red: n_L genes with n_L male-limited *trans*-acting factors and n_L female-limited trans acting factors. Brown: n_L genes with 1 male-limited *trans*-acting factor and 1 female-limited trans acting factor influencing all *cis*-regulators. Gray: n_L genes without regulatory evolution (NOCT simulations, $U_C = U_T = 0$). Other parameters are as in Fig 1a. Note the different time range in panels (a) and (b). In the absence of regulatory evolution (NOCT simulations), with n_L = 500, male average fitness drops quickly to a vanishingly small number (to the order 10^{-17} after a million generation) that should lead to

population extinction, as indicated by the skull symbol. This drop occurs quickly, even before any appreciable mutation accumulation on genes (i.e. it is already down to the inverse of population size when s_Y reaches ~0.01).

Fig 4

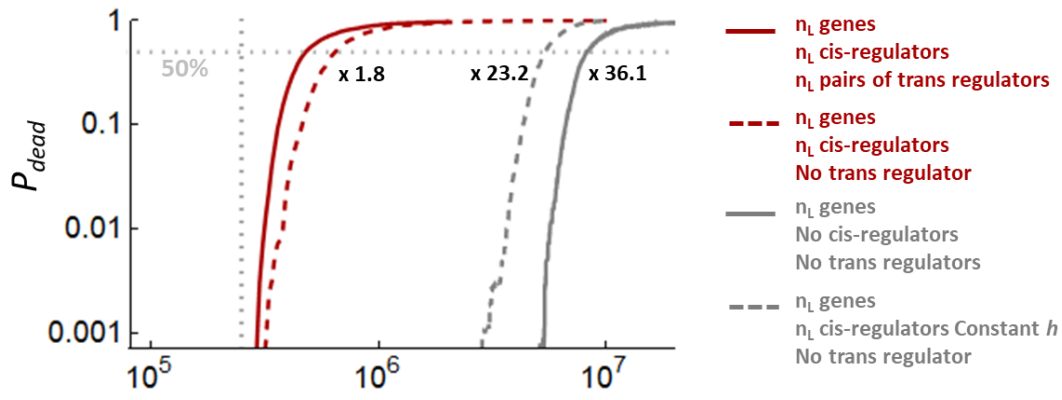

Fig 4. Comparing DRE and DSI. Y degeneration in non-recombining regions with 50 loci. Coding genes are positioned at regular interval on the X/Y chromosome, each at a recombination rate R_G = 5 x 10^{-4} (map length of the non-recombining region is therefore 2.5 cM). Recombination occurs in both sexes during the burn-in period (ending at 2.5 x 10^5 generations, vertical dotted line), and in females thereafter. *x*-axis: number of generations in log-scale. *y*-axis: degeneration of the Y, as measured by P_{dead}, in log-scale. **Dark red**: 50 genes, with 50 *cis*-regulators, 50 male-limited *trans*-acting factors and 50 female-limited trans acting factors. **Dark red, dashed**: 50 genes, with 50 *cis*-regulators, no mutation in *trans*-regulators (NOT simulations, U_T = 0). **Gray, dashed**: 50 genes, with 50 *cis*-regulators, no *trans*-regulators (NOT simulations, U_T = 0), but dominance of the effect of mutations on genes maintained at h = 0.25. **Gray**: 50 genes without regulatory evolution (NOCT simulations, U_C = U_T = 0). Other parameters are as in Fig 2a.

Star Methods

RESOURCE AVAILABILITY

Lead Contact

Further information should be directed to and will be fulfilled by the Lead Contact, Thomas Lenormand (thomas.lenormand@cefe.cnrs.fr).

Materials Availability

Not applicable

Data and Code Availability

The code used in this study is available on Github:

https://github.com/denisroze/Sex_chromosome_degeneration

QUANTIFICATION AND STATISTICAL ANALYSIS

Not applicable

METHOD DETAILS

Model

Mutation and selection

Mutations in *cis* and *trans* regulators are assumed to occur at rates U_c and U_t, respectively, and add a Gaussian deviate to allelic values for these traits ($c + dc \sim N(0, \sigma_c)$, $t + dt \sim N(0, \sigma_t)$). Negative trait values are counted as zero. These values are used to compute the total and allele-specific expression values for each coding gene *G*. In the following expressions, 1 and 2 denote the two copies of a gene (and its *cis*-regulator) present in an individual (male or female), where by convention (and without loss of generality) 1 denotes the gene copy carrying deleterious mutations with the strongest overall effect. Mutations present on copy 1 of gene *i* would decrease fitness by $s_{1,i}$ in a homozygous individual, while mutations present on copy 2 would decrease fitness by $s_{2,i}$ ($s_{1,i} > s_{2,i}$). The overall fitness effect of deleterious mutations in gene *i* depends on the *cis*-regulator strengths associated with alleles 1 and 2, denoted $c_{1,i}$ and $c_{2,i}$. We assume that

the fraction of the protein expressed from allele 1 is $\phi_{1,i} = c_{1,i}/(c_{1,i} + c_{2,i})$, and that the resulting effective dominance coefficient of allele 1 is given by

$$h_{1,i} = {\phi_{1,i}}^{-\ln(h)/\ln(2)} \qquad \textbf{(Eq. 1)}.$$

where *h* is a parameter measuring the dominance of the fitness effect of deleterious mutations in a heterozygote when both alleles are equally expressed ($\phi_{1,i} = 1/2$ leads to $h_{1,i} = h$, see below for a discussion of this and other model assumptions)[14], which will be fixed to 0.25. According to Eq. 1, if allele 1 is relatively more expressed ($\phi_{1,i} > 1/2$), it is assumed to have a larger fitness effect ($h_{1,i} > h$). The fitness effect resulting from the presence of deleterious mutations in gene *i* is then

$$W_i^G = 1 - s_{2,i} - h_{1,i}\left(s_{1,i} - s_{2,i}\right) \qquad \textbf{(Eq. 2)}.$$

Together, Eq. 1 and 2 link the relative rate of expression $\phi_{1,i}$ of the most deleterious allele present at locus *i* to the contribution of gene G_i to fitness, illustrated by Fig 1b.

In addition to the fitness consequences of carrying deleterious mutations in the coding gene, we also assume that the overall expression level of coding genes is under stabilizing selection with an optimum value Q_{opt}. In males, the total expression level Q_i equals $(c_{1,i} + c_{2,i})\bar{t}_{m,i}$, where $\bar{t}_{m,i}$is the average strength of the *trans*-regulators expressed in males, which assumes that both *cis*- and *trans*-regulators are essential for proper expression (neither can be zero). Symmetrically, it is $(c_{1,i} + c_{2,i})\bar{t}_{f,i}$ in females. We assume that $\ln(Q_i)$ is under Gaussian stabilizing selection around $\ln(Q_{opt})$ (with Q_{opt} = 2). We use log-scale to ensure that, irrespective of the intensity of stabilizing selection, the fitness effect of complete regulatory silencing (Q_i = 0) would be s_{max}, the maximum permitted fitness effect of deleterious alleles on the coding gene (Fig **1c**), which we assume to be the same as the effect of a gene knock-out. Denoting by *I* the intensity of stabilizing selection on the expression level, the fitness resulting from the departure from optimal dosage W_i^Q is

$$W_i^Q = 1 - s_{max}\left(1 - e^{-I\left(\ln Q_i - \ln Q_{opt}\right)^2}\right) \qquad \textbf{(Eq. 3)}.$$

This function is equivalent to assuming that fold-changes in expression levels are under symmetric stabilizing selection, while selection on expression levels Q_i is asymmetric (Fig **1c**) — in order to investigate whether this asymmetry influenced the results, we also considered a symmetrized

version of this function (see below). Finally, the overall fitness of an individual was computed as the product over all genes *i* of $W_i^G W_i^Q$.

Dominance

The relationship between relative expression and dominance, as given by Eq. 1 is based on several considerations. First, it is reasonable to assume that there is a smooth and monotonic curve relating the relative expression of a deleterious mutation and its dominance. This curve also must include three fixed points with fitness 1, 1-*hs*, 1-*s* at 0%, 50%, and 100% relative expression, respectively. Given these constraints, the curve relating relative expression and fitness is perhaps not exactly given by Eq. 1, but with a very similar shape and curvature. In addition, Eq. 2 assumes that the identity of the precise mutation in a given gene does not matter in computing fitness and dominance. For instance, a heterozygous individual carrying two different deleterious mutations at a given gene, each with an effect *s*, will have a fitness equal to 1-*s*, i.e. the same fitness than an individual homozygous for the same deleterious mutation with effect *s*. Finally the value of *h* = 0.25 used throughout is based on observations made in a diversity of systems [see survey in 32].

Effect of mutations

We suppose that all mutations occurring on the gene are deleterious. We used an average effect of 0.05 based on direct estimates [see Table 1 in 33]. We used a simple one-parameter exponential distribution to model the diversity of selective effects. Observed distributions of selective effects are similar, especially in absence of beneficial mutations [33]. The model is not considering compensatory mutations in the genes. However, it includes compensatory mutations on regulators, modeled as quantitative traits. In addition, mutations on *cis*-regulators, by partially silencing the effect of deleterious mutations on the genes, also act as compensatory mutations for mutations in the genes. DRE is based on the fact that they cause a long-lasting compensatory effect, on all future deleterious mutations that will occur on the gene (when a gene becomes partially silenced, all future deleterious on that gene will experience a lower dominance). Although more work is needed to investigate this point in detail, it is unlikely that compensatory mutations in the gene will qualitatively alter the results, as long as they are less frequent than deleterious mutations. Once the expression of a gene on the Y has started decreasing (through

the evolution of *cis* and *trans* regulators), selection becomes less effective on the Y copy of the gene, leading to an accumulation of deleterious alleles, which further favors reduced expression. This accumulation may proceed more slowly due to the occasional fixation of a reverse mutation, but should occur anyway as long as the rate of reverse mutations is weak. Eventually, when a gene becomes nearly completely silenced, back mutations on the gene become irrelevant, as they will have no effect (being also silenced).

Cis-regulators runaway

For autosomal genes, a 'cis-regulator runaway' process occurs that leads to stronger *cis*-regulators to become associated with better purged chromosomes with fewer deleterious mutations [14]. This favors stronger *cis*-regulatory alleles, provided they are tightly linked to their coding gene. This process occurs during the burn-in phase of all simulations presented here. Consequently, when recombination stops, cis and trans trait values differed from their initial values (arbitrarily set to one for all regulatory traits) to an extent that differed among simulations with different parameter values (notably population sizes, intensity of stabilizing selection), but not between the sexes or sex chromosomes (Figure S5a). To avoid introducing a bias from this burn-in phase effect, we rescaled all regulatory trait values at the arrest of recombination, by dividing them by their average over the whole population. This allows for a better comparison across parameter values and is not biologically consequential as the absolute value of these traits is arbitrary. Fig S5a shows this runaway for the case illustrated in Fig. 2. This trait variation and scaling is not represented on Fig 2 to avoid overloading the figure. While runaway before the arrest of recombination plays no important role, as it does not introduce a sex or sex chromosome bias, we investigated whether it played a role after recombination arrest. At this point, *cis*-regulators on the Y become more tightly linked to their gene, which should cause a higher runaway rate on the Y. We expect this effect to be minor as the rate of runaway shows a plateau towards lower recombination rates [14]. This small asymmetry can be seen immediately following recombination arrest: Y *cis*-regulator increases slightly faster than X *cis*-regulator (Fig S5a). This effect is slowing down degeneration as it is initiated by a slight Y relative silencing (which is less likely to occur if Y *cis*-regulators are on average slightly stronger than X *cis*-regulators). To confirm this interpretation, we used simulation with larger recombination rate between the *cis*-regulator

and the gene. We found indeed that it further delayed degeneration (see results with R_c artificially set to 0.5, Fig 2c, "FREE" simulations). In any case, after this slight initial asymmetry in runaway rates, *cis*-regulatory divergence between the X and Y takes over.

Initial conditions

Simulations were set up to ensure that no X-Y or male-female asymmetry occurred during the burn-in period. As explained in the previous paragraph on *cis*-regulators runaway, there is an inherent instability of regulatory systems based on *cis* and trans control of gene expression. To allow for possible dosage compensation while X and Y *cis*-regulators diverge, it is necessary to introduce *trans*-regulators with male-limited or female-limited expression. Introducing only male or female-limited *trans*-regulators would introduce an initial asymmetry in the runaway process that may cause male-female differences before the arrest of recombination. For instance, if *cis*-regulators runaway can only be compensated by a female-limited *trans*-regulator, it will lead to a situation where expression is near optimal dosage in females (since the *trans*-regulator with female-limited expression ensures that expression level stays optimal in females), but would overshoot in males (since no *trans*-regulators compensate for *cis*-regulator runaway in males). This initial male overexpression would greatly facilitate Y silencing immediately after recombination arrest, when X and Y *cis*-regulators can start diverging. Indeed, Y silencing would be directly selected for to correct for this male overexpression. This silencing would reflect this initial unbalanced expression, rather than a mechanism that inherently breaks X and Y symmetry after recombination arrest. The reciprocal initial condition (with only *trans*-regulator with male-limited expression) would on the contrary oppose Y silencing (as overexpression in females would rather tend to decrease X *cis*-regulator strength, which would lead to stronger Y relative expression in males). To avoid these initial conditions effects, where male female symmetry is broken before recombination arrest, we considered the symmetrical case with both male and female limited *trans*-regulators. However, this introduce twice as many *trans*- versus *cis* mutations, as there are two *trans*- for one *cis*-regulator. To avoid introducing a higher mutational variance to trans versus *cis*-effects, we therefore halved mutation rates on the two *trans*-regulator loci ($U_C = 2\ U_T$). The absence of *cis*- versus trans trait bias during the burn-in period is

shown on Fig S5b. Overall, we used initial conditions that did not introduce male versus female biases and that did not introduce *cis*- versus *trans*-regulators trait biases.

Regulatory system instability

As soon as recombination stops between the X and Y, *cis*-regulators can start diverging on the X and Y. This divergence can be compensated for by the corresponding evolution of *trans*-regulators expressed in males and in females. Hence, X and Y regulatory systems can drift apart even in absence of deleterious mutations in the coding sequences, while maintaining optimal expression levels in both sexes. In the absence of coding sequence degeneration, the regulatory system can however return to X and Y coexpression after a period where only the X or Y was expressed in males. This is not possible when the Y has degenerated as an increased expression of a dysfunctional gene copy cannot be favored back. In addition, the divergence of the regulatory system in absence of deleterious mutations in the coding sequence is slower than when the 'haploidization' feedback loop occurs. These two effects are illustrated on Fig S4 (see also Fig. 2c, NOG results).

Small population sizes

If population size become small (N_{pop} = 1000, Fig **2d**), Y degeneration can still occur, without DSI or DRE, simply because effective population size of the Y becomes small enough (and smaller than that of the X) to quickly fix deleterious mutations [see [6] for details about this regime]. The asymmetry between the X and the Y results from the lower effective population size of the latter (there are 3 X for one Y in a population with balanced sex-ratio). With small population size, deleterious mutations can also accumulate on the X, for the same reason, although at a much smaller rate (Fig S5d). This may cause population extinction if the non-recombination region is large enough. This mutational meltdown is mitigated if compensatory mutations are included in the model, which is not considered here.

Symmetrized stabilizing selection

To model stabilizing selection on expression levels, we used Gaussian selection on the log of total expression. We used this log to ensure that, irrespective of the intensity of stabilizing selection, the fitness effect of complete regulatory silencing would be s_{max}, the maximum permitted effect

of deleterious alleles on the coding gene. This fitness function introduces an asymmetry on the natural scale for total expression (Fig. 1c). This asymmetry is not biological implausible in most cases. In particular, decreasing expression to zero is likely more deleterious than increasing expression by the same amount. Quite generally, the cost of producing a protein (or the opportunity cost it represents) is likely to be lower than the loss-of-function caused by the shortage of that protein. However, even if biologically plausible, we investigated whether the asymmetry of this fitness function influenced our results. To do so, we used a symmetrized version of the fitness function.

$$W_i^Q = 1 - s_{max}\left(1 - e^{-I\left(\ln Q_i - \ln Q_{opt}\right)^2}\right) \; if \; Q_i < Q_{opt}$$
$$W_i^Q = 1 - s_{max}\left(1 - e^{-I\left(\ln(2Q_{opt} - Q_i) - \ln Q_{opt}\right)^2}\right) \; if \; Q_i > Q_{opt} \qquad \textbf{(Eq. 4)}.$$

This fitness function is identical to Eq. 3 when $Q_i < Q_{opt}$, but is reflected around $Q_i = Q_{opt}$ for $Q_i > Q_{opt}$. It is not defined for $Q_i > 2Q_{opt}$, which makes it not biologically very plausible. However it is entirely symmetrical around $Q_i = Q_{opt}$, and with identical effects than the function we use on the left of Q_{opt}, i.e. for expression levels corresponding to silencing ($Q_i < Q_{opt}$). With this fitness function, Y degeneration occurs almost identically (Fig 2c, Fig S6), indicating that the specific form of the stabilizing selection function plays a limited role in the process. It also indicates that the fitness effects corresponding to large overexpression plays minimal role. This is expected as the population almost never evolves to such extreme trait value, due to the stabilizing selection around Q_{opt}.

Modes of dosage compensation

Many different mechanisms of dosage compensation have been described in a diversity of organisms. It is possible to relate these mechanisms to the model we use. Once the Y is fully silenced ($c_{Y,i} = 0$), and assuming that the population stays at Q_{opt} in both males and females (which is a very good approximation unless stabilizing is very weak), we have

$$Q_{opt} = c_{X,i}\bar{t}_m = 2c_{X,i}\bar{t}_f \qquad \textbf{(Eq. 5)}.$$

We therefore have $\bar{t}_m = 2\bar{t}_f$, which defines dosage compensation (Fig S7). Hence the way dosage compensation works, i.e. the triplet $(\bar{t}_m, \bar{t}_f, c_{X,i})$ can be described by a single parameter. We can choose e.g. to use $\bar{t}_m$ for this description. Compared to the initial system with $\bar{t}_m = \bar{t}_f = c_{X,i} = c_{Y,i} = 1$, a final compensation characterized by $\bar{t}_m$ =1 (i.e. $\bar{t}_m = 1, \bar{t}_f = 0.5, c_{X,i} = 2, c_{Y,i} = 0$) would correspond to the *Caenorhabditis elegans* case, where the X is inherently expressed twice as much ($c_{X,i} = 2$) to obtain optimal expression in males, while a female-limited *trans*-regulator halves expression ($\bar{t}_f = 0.5$) to recover optimal expression in females. This is also very similar to the mammal case where a female-limited *trans*-regulator halves expression by randomly silencing one X (rather than halving expression of each X like in *C. elegans*). The case $\bar{t}_m$ = 2 (i.e. $\bar{t}_m = 2, \bar{t}_f = 1, c_{X,i} = 1, c_{Y,i} = 0$) would correspond to the *Drosophila* case, where a male-limited *trans*-acting factor doubles X expression to obtain optimal expression in males (nothing being changed in females, $\bar{t}_f = 1, c_{X,i} = 1$). Several other compensation mechanisms may occur, such as *trans*-acting factors specifically targeting the X or Y in males, based on parent-of-origin imprints [26,but see 34]: indeed, maternal or paternal imprints can easily identify proto-X and proto-Y in males. The current model should be extended to examine this diversity of dosage compensation mechanisms including the possibility of mixture of mechanisms [35]. This extension may also require introducing different types of constraints to stabilize the runaway of *cis*-regulators, in the period preceding the arrest of recombination, so that the type of compensation could be evaluated against a stable regulatory system.

Number of selected loci

The quantitative comparison between DRE and DSI is not straightforward as most models of DRI do not involve regulatory loci. At first sight, it might be possible to 'convert' *cis*-regulators into genes exposed to deleterious mutations to make a comparison that would use the same number of selected loci, mutation load and the same genetic map. This is however not trivial, as selective effects on *cis*-regulators are not directly comparable to those on genes. Regulatory traits are modelled as traits under stabilizing selection, i.e. with deleterious and beneficial mutations, and

influencing the dominance of mutations on the genes. On Fig. 4, we present two better control cases. The first corresponds to simulations in the absence of mutation in *trans*-regulators. This is slowing down degeneration by making *cis*-regulatory divergence more difficult, but the regulatory feedback loop is still present (dashed red curve). However, the effect of *cis*-divergence on silencing is still present, and this is therefore not representing DSI well. A second control considers the same situation, but where the link between regulatory trait values and dominance of mutation on the genes is artificially removed (setting *h* constant and equal to 0.25 in HFIX simulations). In this case, regulatory traits are just quantitative traits with no effect on genes, and the simulation therefore has the same number of loci, load and map compared to a full DRE simulation with the regulatory effects removed (gray dashed line). In this case, we observe that degeneration occurs faster than when all regulators are removed, as expected as there are more interfering loci. However, the dynamics are still much slower than when the regulatory feedback loop is occurring. Specifically, it is 23 times slower to reach the time where 50% of genes are degenerate, compared to DRE.

Scaling and parameter effects

Fig S8 shows additional simulation results for lower values of mutations rates and strength of selection against deleterious mutations, and test a scaling argument from diffusion theory.

Mutation rates

With smaller mutation rates, all processes are slower. Fig S8a shows results for mutation rates 10 times lower. Results show that the dynamics with only one gene are the same, just ten times slower. This scaling holds true for simulations with or without regulators; therefore, all our conclusions remain the same. Results with 50 genes show that degeneration is slowed down by a factor ≈20 when mutation rates are 10 times lower. This scaling also holds for simulations with or without regulators, and thus also leaves our main conclusions unchanged (in particular, regarding the fact that regulatory evolution may considerably accelerate degeneration). The difference in scaling between the 1 and 50 genes cases may be due to the fact that the speed of degeneration caused by selective interference increases faster than linearly with the mutation rate. Since the one-gene case is less impacted by lowering the mutation rate than the 50 locus case, we conclude

overall that DRE is robust to our assumptions. If anything, it should play a greater role at lower mutation rates, compared with selective interference.

Effect of deleterious mutations

Fig S8b shows the results of simulations with lower mean effect s_{mean} of deleterious mutations. Interpreting the effect of s_{mean} is not straightforward, as the fitness effect of deleterious mutations has a non-monotonic effect on the rate of fitness decline in Muller's ratchet. With 10 times lower s_{mean}, we find that when regulators evolve, degeneration occurs twice as fast with only one selected gene, but occurs at about the same rate with 50 genes. Simulations without regulatory evolution show a different scaling (degeneration is 10 times faster and 5 times faster with 1 and 50 loci, respectively). This difference in scaling may be expected as with regulator evolution, dominance of the mutations carried on the Y declines, and tends to zero, which makes the dynamics less dependent on s_{mean}. In any case, as with the simulations with lower mutation rates, degeneration always occurs much more rapidly in simulations with regulatory evolution than in simulations without. Our main conclusion is thus robust to our choice of parameters, even if quantitative differences can arise with different parameter values. This is to be expected in all cases, as the simulations with regulatory evolution also necessarily include the effect of selective interference. It only adds an extra process that accelerates degeneration. Degeneration by regulation can occur 'alone' (as in the 1 locus simulations), but it will necessarily combine and add to selective interference in other cases.

Scalings

From diffusion theory, one expects that allele frequency dynamics should be roughly insensitive to population size, as long as the product of population size N_{pop} with the other parameters (mutations rates, strength of selection, etc.) remains constant, and when time is measured in units of N_{pop} generations (provided that N_{pop} is sufficiently large, while mutation and selection are sufficiently weak). This scaling argument is interesting, as it indicates that the behaviour of very large populations (that could not be simulated in a reasonable amount of time) can be deduced from simulations of smaller populations [e.g. 36]. In order to test if such a scaling may hold in our model, we compared simulations with N_{pop} = 1000 and our default parameter values (REG

simulations) with simulations with N_{pop} = 10000, and where the values of the parameters σ_C, σ_T, s, I, U_g, U_c, U_{tm}, U_{tf}, R_g and R_c were divided by 10, so that the product between N_{pop} and these parameters stayed constant (SCAL simulations). The results show that SCAL simulations are comparable but about 100 times slower on the accumulation of deleterious mutations on the Y (10x caused by the scaling of time expressed in units of N_{pop} generations, and 10x caused by the scaling of the variable measured N_{pop} s_Y). This scaling works for SCAL simulations with 1 or 50 genes, suggesting that our results may be extrapolated to other parameter combinations, provided the products of deterministic parameters with N_{pop} are kept constant.

KEY RESOURCES TABLE

Software and Algorithms		
C++ Simulations	This paper	https://github.com/denisroze/Sex_chromosome_degeneration
Mersenne Twister random number generator	M. Matsumoto and T. Nishimura (1998). Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. ACM Trans. on Modeling and Computer Simulation 8:3-30	http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/VERSIONS/C-LANG/MersenneTwister.h

Supplementary Fig S5

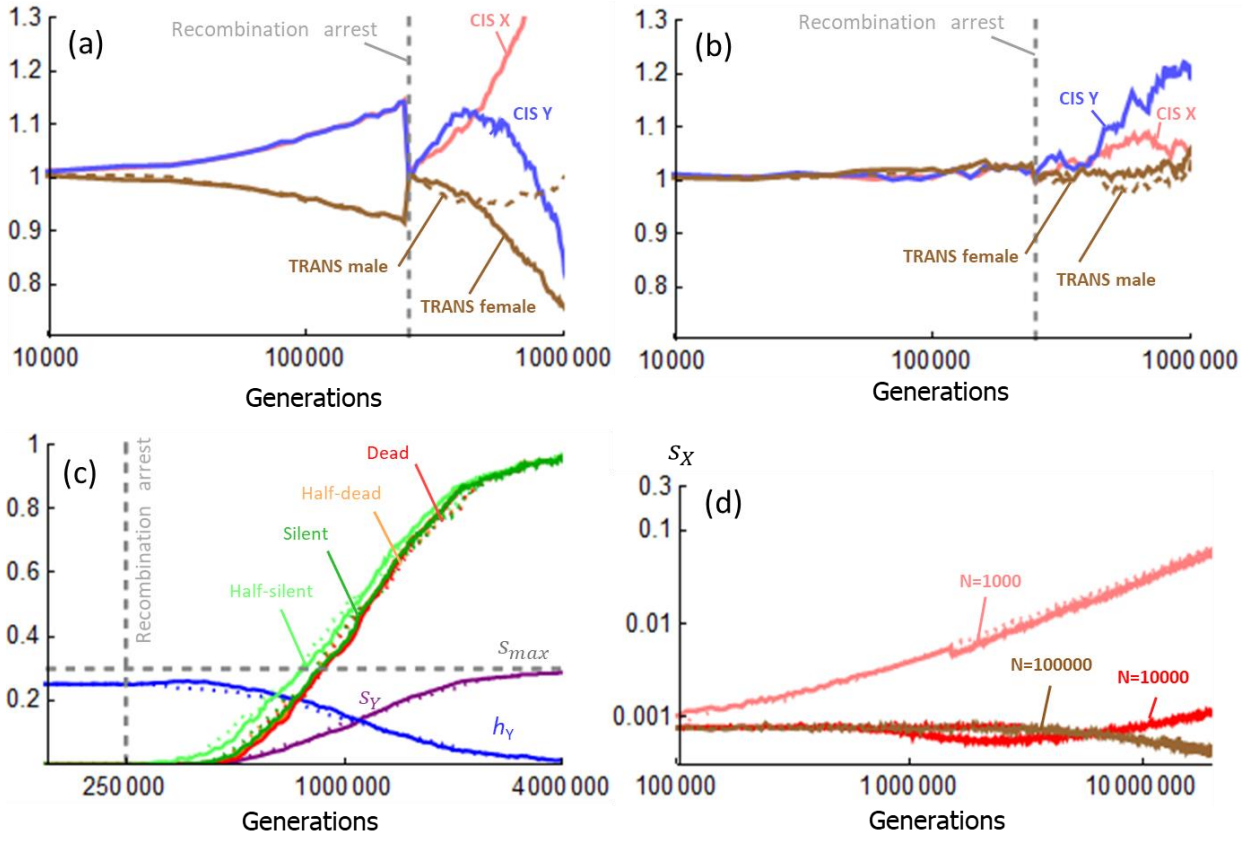

Fig S5. Control simulations. (a) *Cis*-regulator runaway during the burn-in phase. Variation of regulatory traits corresponding to the case illustrated in Fig 2, zooming on the burn-in phase. Regulatory trait values are rescaled to one at the arrest of recombination (Generation 250 000) for a fair comparison across simulations with different parameter values. Before recombination arrest, *cis*-regulator strength tends to increase while *trans*-regulator strength decreases

(achieving the same overall expression level). These variations are however symmetric for the *cis*-regulators on the X and Y, and for female limited or male limited *trans*-regulators. Pink: X *cis*-regulator strength; Blue: Y *cis*-regulator strength; Brown: *trans*-regulator strength (plain: female limited, dashed: male-limited). **(b)** *Cis*- and *trans*-regulator evolution in absence of deleterious mutations in the gene. Variation of regulatory traits corresponding to the case illustrated in (a), except that $U_G = 0$. Pink: X *cis*-regulator strength; Blue: Y *cis*-regulator strength; Brown: *trans*-regulator strength plain: female limited, dashed: male-limited. **(c)** Y degeneration under regulatory evolution. Same as Figure 2a. Plain line correspond to simulations using Eq. 2 for stabilizing selection, while dotted lines use the symmetrized version given in Eq. 4. **(d)** Accumulation of deleterious mutations on the X for different population sizes (10^3, 10^4, 10^5). Same parameter values than in Fig 2. *x*-axis number of generations in log-scale. *y*-axis: s_X. Plain lines: with evolving regulators. Dotted lines: without regulator evolution (NOCT simulations, $U_C = U_T = 0$), but both are confounded at each population size.

650

Supplementary Fig S6

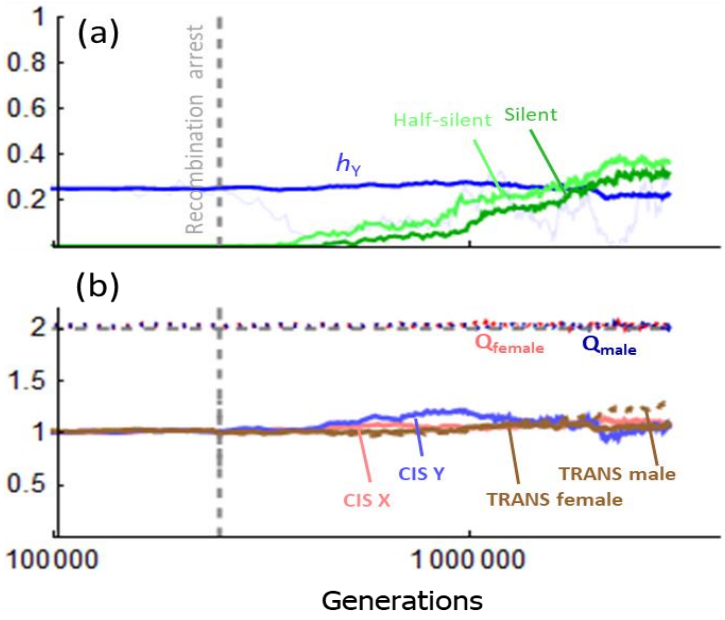

Fig S6. Regulatory system instability. Regulatory system instability in absence of deleterious mutation in the coding sequence. Same parameter values than in Fig 2, except that $U_G = 0$. **(a)** x-axis: time in number of generations, in log-scale. Recombination stops between the X and Y chromosome at generation 250000 (vertical dashed gray line); y-axis: The probability that a coding gene on the Y is expressed at less than half the level of the X copy (half-silent, light green curve), entirely silenced (dark green curve). The figure also shows the dominance of deleterious mutations carried on the Y (h_Y, blue curve). A particular replicate is added (thin blue line) to show how it can reach $h_Y = 0$ and can then rebound with $h_Y > 0$. **(b)** Time-variation of regulatory traits corresponding to the case illustrated in panel a. x-axis: time in number of generations in log-scale; y-axis: regulator trait values. Pink: X *cis*-regulator strength; Blue: Y *cis*-regulator strength; Brown: *trans*-regulator strength plain: female limited, dashed: male-limited. Optimal dosage is 2 (dashed gray line). Total expression in males and females is indicated by the dotted curves (male value in blue, female in red). Note that panel (b) shows averages across replicates. Even if there is no bias in the average regulatory trait values, there is always a chance that at a given time, in a given replicate, the Y is silent. The divergence of *cis*-regulators corresponds here to a random walk with the only constraint being that the overall level of expression is maintained.

Supplementary Fig S7

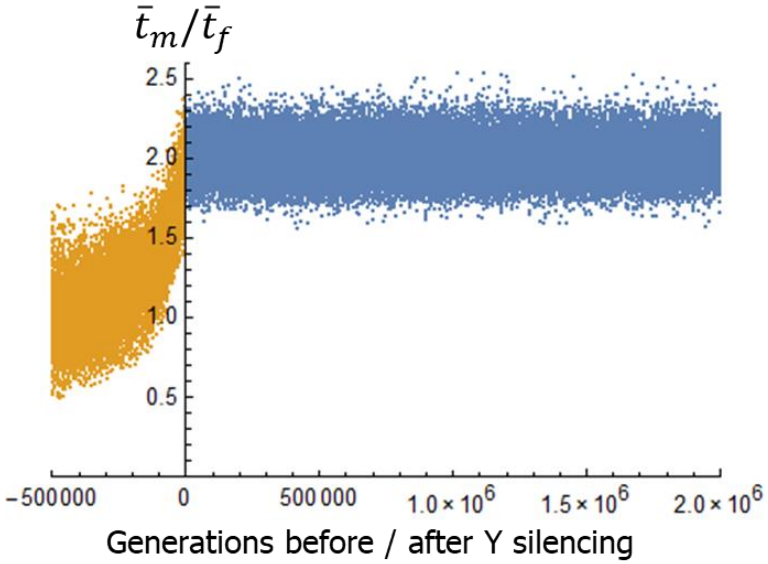

Fig S7. Evolution of dosage compensation. Dosage compensation is measured by the $\bar{t}_m/\bar{t}_f$ ratio (methods). Simulation parameters corresponding to Fig 2a. The data of all replicates are aggregated. *x*-axis: number of generations after Y silencing (defined as the time at which c_Y reaches 10^{-4}). Yellow: period before Y silencing. Blue: period after Y silencing.

Supplementary Fig S8

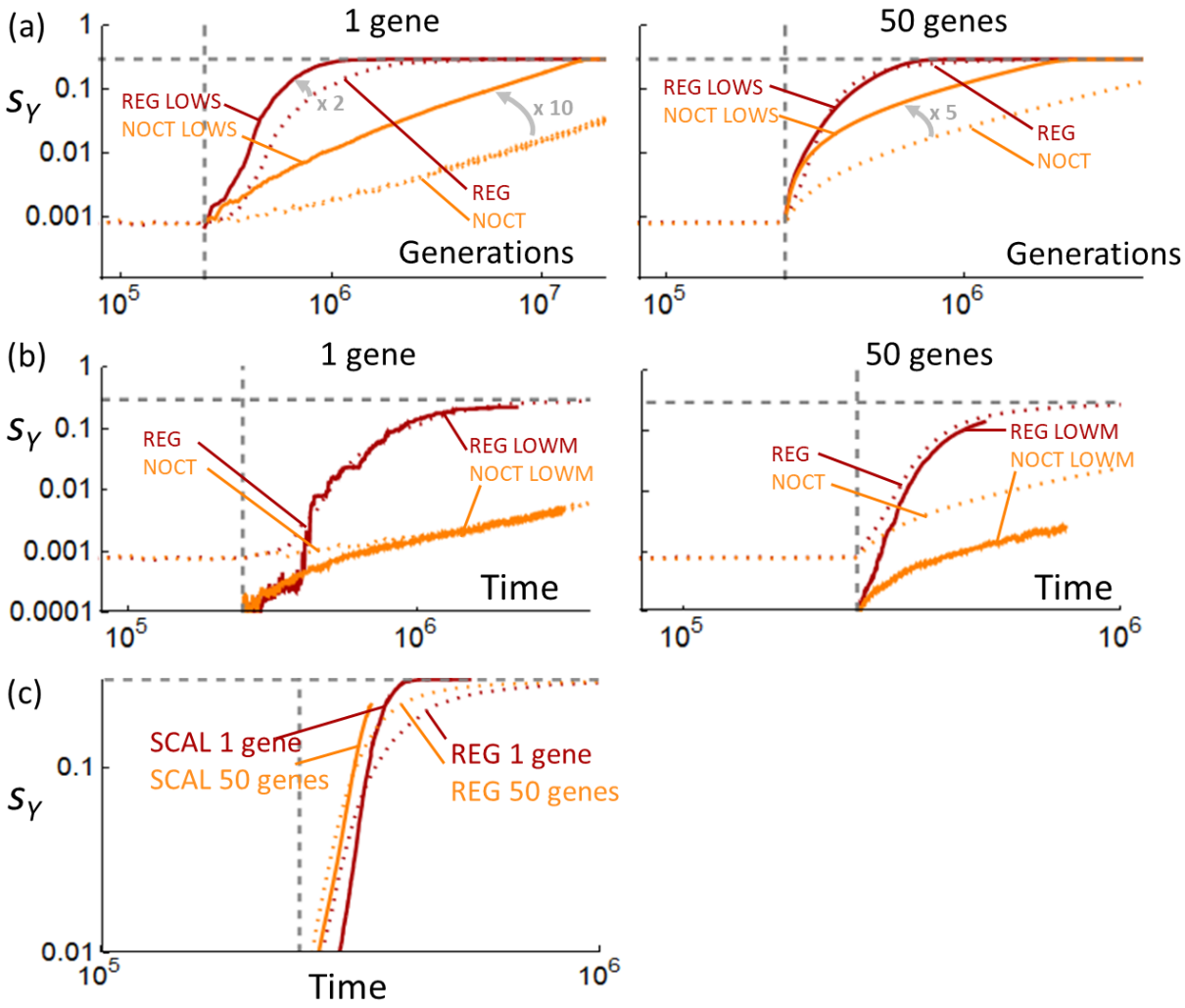

Fig S8 Scaling and parameter values. (a). Simulations with lower mutation rates. y-axis: accumulation of deleterious mutations on the non-recombining region of the Y (s_Y). Left panel, 1 gene simulations; right panel, 50 genes simulations. x-axis: time in generations for REG simulations, but in units of 10 (left panel) or 20 (right panel) generations for LOWM simulations (in reality LOWM simulations are 10 or 20 times slower). N_{pop} =10000 for all simulations. Red: REG simulations as in Fig 2. Orange: NOCT simulations (i.e. simulation without mutations in regulators). LOWM simulations are simulations where mutations rates are divided by 10. **(b)**. Simulations with lower average effect of deleterious mutations in the genes. y-axis: accumulation of deleterious mutations on the non-recombining region of the Y (s_Y). Left panel, 1 gene

687 simulations; right panel, 50 genes simulations. x-axis: time in generations. Npop =10000 for all
688 simulations. Red: REG simulations as in Fig 2. Orange: NOCT simulations (i.e. simulation without
689 mutations in regulators). LOWS simulations are simulations where s_{mean} is divided by 10. Gray
690 numbers indicate the time scaling required for curves to superpose. **(c)** Comparison of simulations
691 in which the product of deterministic parameters and population size stays constant. y-axis:
692 accumulation of deleterious mutations on the non-recombining region of the Y (s_Y). x-axis: time
693 in generations for REG simulations, but in units of 100 generations for SCAL simulations (in reality
694 SCAL simulations are 100 times slower). Red: 1 gene, Orange : 50 genes. Dotted lines: REG
695 simulations, like in Fig 2d with N_{pop} = 1000. Plain line, SCAL simulations with all parameters divided
696 by 10 compared to the reference simulation in Fig 2a, except N_{pop} = 10000 and s_{max} = 0.3.

697

Supplementary table 1

Simulation	Replicates	Figure	Parameter values
REG	330	2a, 2b, 3a, 3b	Default
REG*	200	2c	I=0.01
NOG	100	2c, S3, S4	$U_g = 0$
REC	100	2c	$R_c = 0.5$
NOCT	150	2c, 2d	$U_c = U_t = 0$
NOC	100	2c	$U_c = 0$
NOT	100	2c	$U_t = 0$
NOT*	100	2c	$U_t = 0$; I=0.01
SYM	200	2c, S5	*Symmetrized stabilizing selection fitness function*
SCAL	94	S8	*Default parameters/10 except Npop and* s_{max}
LOWM	40	S8	*All mutation rates /10*
LOWS	40	S8	s_{mean} */10*
NOCTLOWM	40	S8	*Ug/10;* $U_c = U_t = 0$
NOCTLOWS	40	S8	s_{mean} */10;* $U_c = U_t = 0$
REG 1000	300	2d	N_{pop} *= 1000*
REG 100000	190	2d	N_{pop} *= 100000*
NOCT 1000	200	2d	N_{pop} *= 1000;* $U_c = U_t = 0$
NOCT 100000	120	2d	N_{pop} *= 100000;* $U_c = U_t = 0$
REG 50 locus	40	3a	$n_L = 50$
SING 50 loci	48	3a	$n_L = 50$; single male- and single female-limited *trans*-regulator
NOT 50 loci	25	S7	$n_L = 50$; $U_t = 0$
HFIX 50 loci	25	S7	$n_L = 50$; $U_t = 0$; dominance always maintained to h=0.25
NOCT 50 loci	60	3a	$n_L = 50$; $U_c = U_t = 0$
SCAL 50 loci	30	S8	$n_L = 50$; *Default parameters/10 except Npop and* s_{max}
LOWM 50 loci	20	S8	$n_L = 50$; *All mutation rates /10*
LOWS 50 loci	20	S8	$n_L = 50$; s_{mean} */10*

NOCTLOWM 50 loci	11	S8	n_L = 50; *Ug/10;* $U_c = U_t = 0$
NOCTLOWS 50 loci	15	S8	n_L = 50; s_{mean} */10;* $U_c = U_t = 0$
REG 500 locus	20	3b	n_L = 500
SING 500 locus	40	3b	n_L=500; single male- and single female-limited *trans*-regulator
NOCT 500 locus	20	3b	n_L = 500; $U_c = U_t = 0$

Table 1. Simulation plan and parameter values. List of simulations with their acronym, their parameter values and numbers of replicates. Default parameters: Population size (N): 10000; Standard deviation of mutational effects on regulators $\sigma_T = \sigma_C = 0.2$; Number of quadruplets gene, *cis*-regulator, male-limited *trans*-regulator, female limited *trans*-regulator (n_L): 1; Number of generations during burn-in phase: 250000; Average effect of deleterious mutations (s): 0.05; fitness effect of a gene knock-out (s_{max}): 0.3; Intensity of stabilizing selection (I): 0.1. Mutation rates $U_g = U_c = 2\,U_{tm} = 2\,U_{tf} = 0.0002$; Recombination rate between consecutive genes ($\underline{R_g}$): 0.0005; Recombination rate between *cis*-regulator and its gene (R_c): 0.00005.

References

1. Bull, J.J. (1983). Evolution of sex determining mechanisms (Menlo Park, CA: Benjamin Cummings).

2. Abbott, J.K., Nordén, A.K., and Hansson, B. (2017). Sex chromosome evolution: historical insights and future perspectives. Proc. R. Soc. B Biol. Sci. *284*, 1–9.

3. Charlesworth, B. (1978). Model for evolution of Y chromosomes and dosage compensation. Proc. Natl. Acad. Sci. U. S. A. *75*, 5618–5622.

4. Ellegren, H. (2011). Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. Nat Rev Genet *12*, 157–166.

5. Beukeboom, L.W., and Perrin, N. (2014). The Evolution of Sex Determination (Oxford University Press).

6. Charlesworth, B., and Charlesworth, D. (2000). The degeneration of Y chromosomes. Philos. Trans. R. Soc. B Biol. Sci. *355*, 1563–1572.

7. Charlesworth, B. (1996). The evolution of chromosomal sex determination and dosage compensation. Curr. Biol. *6*, 149–162.

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

9. Good, B.H., Walczak, A.M., Neher, R.A., and Desai, M.M. (2014). Genetic Diversity in the Interference Selection Limit. PLoS Genet. *10*, e1004222.

10. Charlesworth, B., and Campos, J.L. (2014). The Relations Between Recombination Rate and Patterns of Molecular Variation and Evolution in Drosophila. Annu. Rev. Genet. *48*, 383–403.

11. Neher, R.A. (2013). Genetic Draft, Selective Interference, and Population Genetics of Rapid Adaptation. Annu. Rev. Ecol. Evol. Syst. *44*, 195–215.

12. Bachtrog, D. (2006). Expression Profile of a Degenerating Neo-Y Chromosome in Drosophila. Curr. Biol. *16*, 1694–1699.

13. Zhou, Q., and Bachtrog, D. (2012). Chromosome-Wide Gene Silencing Initiates Y Degeneration in *Drosophila*. Curr. Biol. *22*, 522–525.

14. Fyon, F., Cailleau, A., and Lenormand, T. (2015). Enhancer runaway and the evolution of diploid gene expression. PLoS Genet. *11*, e1005665.

15. Fyon, F., and Lenormand, T. (2018). Cis-regulator runaway and divergence in asexuals. Evolution.

16. Wray, G.A., Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M. V, and Romano, L. a (2003). The evolution of transcriptional regulation in eukaryotes. Mol. Biol. Evol. *20*, 1377–419.

17. Mank, J.E. (2013). Sex chromosome dosage compensation: Definitely not for everyone. Trends Genet. *29*, 677–683.

18. Lahn, B.T., and Page, D.C. (1999). Four evolutionary strata on the human X chromosome. Science.

19. Pandey, R.S., Wilson Sayres, M.A., and Azad, R.K. (2013). Detecting Evolutionary Strata on the Human X Chromosome in the Absence of Gametologous Y-Linked Sequences. Genome Biol. Evol. *5*, 1863–1871.

20. Ellison, C., and Bachtrog, D. (2019). Contingency in the convergent evolution of a regulatory network : Dosage compensation in Drosophila. PLoS Biol. *17*, e3000094.

21. Orr, H.A., and Kim, Y. (1998). An adaptive hypothesis for the evolution of the Y chromosome. Genetics *150*, 1693–1698.

22. Engelstädter, J. (2008). Muller's Ratchet and the Degeneration of Y Chromosomes: A Simulation Study. Genetics *180*, 957–967.

23. Ercan, S. (2015). Mechanisms of X chromosome dosage compensation. J. genomics *3*, 1–19.

24. Disteche, C.M. (2012). Dosage compensation of the sex chromosomes. Annu. Rev. Genet. *46*, 537–560.

25. Ohno, S. (1966). Sex Chromosomes and Sex-Linked Genes (Berlin, Heidelberg: Springer Berlin Heidelberg).

26. Muyle, A., Zemp, N., Fruchard, C., Cegan, R., Vrana, J., Deschamps, C., Tavares, R., Hobza, R., Picard, F., Widmer, A., *et al.* (2018). Genomic imprinting mediates dosage compensation in a young plant XY system. Nat. Plants *4*, 677–680.

27. Martin, H., Carpentier, F., Gallina, S., Godé, C., Schmitt, E., Muyle, A., Marais, G.A.B., and Touzet, P. (2019). Evolution of Young Sex Chromosomes in Two Dioecious Sister Plant Species with Distinct Sex Determination Systems. Genome Biol. Evol. *11*, 350–361.

28. Muyle, A., Zemp, N., Deschamps, C., Mousset, S., Widmer, A., and Marais, G.A.B. (2012). Rapid De Novo Evolution of X Chromosome Dosage Compensation in Silene latifolia, a Plant with Young Sex Chromosomes. PLoS Biol. *10*, e1001308.

29. Veltsos, P., Ridout, K.E., Toups, M.A., González-Martínez, S.C., Muyle, A., Emery, O., Rastas, P., Hudzieczek, V., Hobza, R., Vyskot, B., *et al.* (2019). Early Sex-Chromosome Evolution in the Diploid Dioecious Plant Mercurialis annua. Genetics *212*, 815–835.

30. Wei, K.H.C., and Bachtrog, D. (2019). Ancestral male recombination in Drosophila albomicans produced geographically restricted neo-Y chromosome haplotypes varying in age and onset of decay. PLoS Genet., e1008502.

31. Hough, J., Hollister, J.D., Wang, W., Barrett, S.C.H., and Wright, S.I. (2014). Genetic degeneration of old and young Y chromosomes in the flowering plant Rumex hastatulus. Proc. Natl. Acad. Sci. U. S. A. *111*, 7713–8.

32. Manna, F., Martin, G., and Lenormand, T. (2011). Fitness landscape: an alternative theory for the dominance of mutations. Genetics *189*, 923–937.

33. Martin, G., and Lenormand, T. (2006). A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. Evolution *60*, 893–907.

34. Krasovec, M., Kazama, Y., Ishii, K., Abe, T., and Filatov, D.A. (2019). Immediate Dosage Compensation Is Triggered by the Deletion of Y-Linked Genes in Silene latifolia. Curr. Biol., 2214–2221.

35. Gu, L., Reilly, P.F., Lewis, J.J., Reed, R.D., Andolfatto, P., and Walters, J.R. (2019). Dichotomy of Dosage Compensation along the Neo Z Chromosome of the Monarch Butterfly. Curr. Biol., 4071–4077.

36. Dolgin, E.S., and Charlesworth, B. (2006). The fate of transposable elements in asexual populations. Genetics *174*, 817–827.