

Muller's ratchet of the Y chromosome with gene conversion

Takahiro Sakamoto * and Hideki Innan*

Department of Evolutionary Studies of Biosystems, SOKENDAI, The Graduate University for Advanced Studies, Hayama, Kanagawa 240-0193, Japan

*Corresponding authors: Graduate University for Advanced Studies, Hayama, Kanagawa 240-0193, Japan. Email: sakamoto_takahiro@soken.ac.jp (T.S.); innan_hideki@soken.ac.jp (H.I.)

Abstract

Muller's ratchet is a process in which deleterious mutations are fixed irreversibly in the absence of recombination. The degeneration of the Y chromosome, and the gradual loss of its genes, can be explained by Muller's ratchet. However, most theories consider single-copy genes, and may not be applicable to Y chromosomes, which have a number of duplicated genes in many species, which are probably undergoing concerted evolution by gene conversion. We developed a model of Muller's ratchet to explore the evolution of the Y chromosome. The model assumes a nonrecombining chromosome with both single-copy and duplicated genes. We used analytical and simulation approaches to obtain the rate of gene loss in this model, with special attention to the role of gene conversion. Homogenization by gene conversion makes both duplicated copies either mutated or intact. The former promotes the ratchet, and the latter retards, and we ask which of these counteracting forces dominates under which conditions. We found that the effect of gene conversion is complex, and depends upon the fitness effect of gene duplication. When duplication has no effect on fitness, gene conversion accelerates the ratchet of both single-copy and duplicated genes. If duplication has an additive fitness effect, the ratchet of single-copy genes is accelerated by gene duplication, regardless of the gene conversion rate, whereas gene conversion slows the degeneration of duplicated genes. Our results suggest that the evolution of the Y chromosome involves several parameters, including the fitness effect of gene duplication by increasing dosage and gene conversion rate.

Keywords: Y chromosome; Muller's ratchet; gene conversion; population genetics

Introduction

Nonrecombining chromosomes such as Y chromosome often degenerate rapidly because deleterious mutations accumulate irreversibly. This process called Muller's ratchet (Muller 1964) has been investigated extensively in many theoretical studies (Haigh 1978; Stephan *et al.* 1993; Gessler 1995; Charlesworth and Charlesworth 1997; Gordo and Charlesworth 2000a, 2000b; Jain 2008; Rouzine *et al.* 2008; Waxman and Loewe 2010; Neher and Shraiman 2012; Goyal *et al.* 2012; Metzger and Eule 2013). Although these studies used models of single-copy genes, recent sequencing of the Y chromosome has revealed that many genes have acquired multiple copies through gene duplication. These theoretical results therefore cannot be directly applied to the Y chromosome, because the evolution of duplicated genes is not as simple as that of single-copy genes. Duplicated genes are likely to undergo concerted evolution, during which the duplicated copies coevolve by exchanging their DNA sequences with each other by gene conversion (Ohta 1983; Arnheim 1983). The major effect of gene conversion during concerted evolution is that, under neutrality, the level of divergence between the duplicated copies is kept low, while the level of polymorphism within each copy is increased (Innan 2002, 2003; Teshima and Innan 2004). It has been theoretically demonstrated that the effect of selection is enhanced in duplicated genes; deleterious mutations are more efficiently removed, and beneficial mutations are more likely to

become fixed in both of the duplicated copies (Mano and Innan 2008). However, the way in which gene conversion affects the degeneration of the Y chromosome, in which single-copy genes and duplicated genes coexist with complete linkage, is not fully understood.

The aim of this study was to theoretically understand the degeneration process of duplicated genes with special interest in the Y chromosome. Y chromosomes have a unique evolutionary history. They usually evolve from an autosome, on which a sex-determining locus arises. After recombination with the X chromosome is suppressed, the Y chromosome gradually loses functional genes by the accumulation of deleterious mutations (Charlesworth and Charlesworth 2000; Bachtrog 2013). During this process, many genes undergo gene duplication (Skaletsky *et al.* 2003; Hughes *et al.* 2010, 2012, 2020; Soh *et al.* 2014). Duplicated gene copies are either in large palindrome structures, as found in primates (Skaletsky *et al.* 2003; Hughes *et al.* 2010, 2012), or arranged in tandem, as found in mouse (Soh *et al.* 2014), bull (Hughes *et al.* 2020), and fruit fly (Bachtrog *et al.* 2019). Since homologous copies often show high sequence identity, frequent gene conversion should have occurred between the copies (Rozen *et al.* 2003; Hallast *et al.* 2013; Skov *et al.* 2017). These findings indicate that the degeneration of the Y chromosome involves both single-copy and duplicated genes. In this work, we develop a

theory to address the following questions: (1) After gene duplication, will the degeneration rate become faster or slower? and (2) How does gene duplication affect the rate of degeneration of linked single-copy genes?

To investigate these questions, we used a model of Muller's ratchet. In a broad sense, Muller's ratchet is a process by which deleterious mutations are fixed irreversibly in the absence of recombination (Muller 1964). In theoretical reports, it is commonly assumed that all mutations have the same effect on fitness [but see Söderberg and Berg (2007)]. Under these conditions, the fitness of an individual depends only upon how many deleterious mutations it has, therefore individuals in the population can be classified based on the number of deleterious mutations (d), as illustrated in Figure 2A, in which all haploid individuals have four functional genes, represented by different colored boxes. The class $d=0$ consists of individuals with no deleterious mutations, which have the highest fitness in the population (i.e., the least-loaded class in this situation). The second class is that of individuals with one deleterious mutation ($d=1$), and those who have two deleterious mutations belong to the class $d=2$. Because we assume that all mutations are irreversible (i.e., back mutation is ignored), the class of an individual can shift down (e.g., $d=0 \rightarrow 1 \rightarrow 2$) as it accumulates mutations. Under these assumptions, the population can be structured according to d , and Muller's ratchet proceeds as these classes turn over and over. In practice, if the least-loaded class ($d=0$ in Figure 2A) goes extinct, the ratchet clicks and the class of $d=1$ becomes the least-loaded class.

We here extend the model of Muller's ratchet to the case of duplicated genes. Figure 2B illustrates an example in which all haploid individuals have four genes which have been duplicated. Again, the population can be structured based on the number of deleterious mutations, and the process of Muller's ratchet proceeds along the turnover of the classes. The major difference is that a new deleterious mutation which has occurred in the duplicated genes is not "irreversible" because gene conversion could remove it. If the original part of the intact copy without the corresponding mutation is transferred, the mutated copy will lose the mutation. However, if gene conversion occurs in the opposite direction—from the mutated copy to the intact version—the mutation becomes irreversible in this individual, because gene conversion cannot remove it any more, under the assumption of no back mutation. We refer to the former and latter types of mutations as reversible and irreversible mutations, respectively (presented by yellow and red circles in Figure 2B). Thus, when a new mutation arises in one copy, its fate is not determined, and we can consider that the mutation can contribute to an irreversible click of Muller's ratchet when the mutation is shared in both copies. In this model, we can simplify the process of Muller's ratchet if we assume that all mutations are recessive and have the same fitness effect. Individuals can therefore be classified according to the number of irreversible mutations they carry. Note that reversible mutations have no fitness effect. Intuitively, gene conversion should have two opposite effects on the degeneration process (Graves 2004). If gene conversion mutates both copies, producing an irreversible mutation, the degeneration process is accelerated. If, however, gene conversion removes the mutation and restores both copies to the original form, the speed of degeneration is slowed. We used this model to explore the way in which gene conversion affects the degeneration of duplicated genes. We were particularly interested in the interactions between the two counteracting effects of gene conversion. We provide analytical expressions of the speed of Muller's ratchet

under this simplified model with a constant selection coefficient. We also consider the effect of variable selection coefficients and the degree of dominance, using mathematical analysis and simulations.

Several studies have investigated the effect of gene conversion on the degeneration of Y chromosomes (Connallon and Clark 2010; Marais et al. 2010), but their focuses have been different from those of the present study. Connallon and Clark (2010) investigated the role of gene conversion on the conservation of duplicated pairs. Because their model assumes that deleterious mutations have lethal effects, a mutation cannot be shared by both duplicated copies, because this situation would be lethal. Therefore, there is no gene loss through the accumulation of deleterious mutations. Marais et al. (2010) investigated the evolution of the human Y chromosome using simulations, in which gene conversion between duplicated genes was taken into account. Their focus was on the process of fixation of a modifier of the gene conversion rate, rather than the long-term degeneration process.

Model

General model

We used a discrete-generation Wright–Fisher model of a haploid population with size N . Each chromosome consists of L_1 single-copy genes and L_2 pairs of duplicated genes (Figure 1A). We assume no inter-chromosomal recombination, or crossing-over, so all genes on the same chromosome are completely linked. A chromosome therefore behaves as a single haploid individual. We are interested in the way in which the genes lose their functions through Muller's ratchet, resulting in a reduction in the number of functional genes on the chromosome. To this end, we applied a simple loss-of-function model to each gene and gene pair. For the i th single-copy gene, we assume the fitness effect of losing the gene function to be s_i . We only consider loss-of-function mutations, so that one mutation is sufficient to make a gene a pseudogene, with loss of the gene function. Throughout this article, we say a gene is "lost" when it loses the function. The rate of loss-of-function mutation is u per copy per generation, and no back mutation is allowed. Every mutation therefore

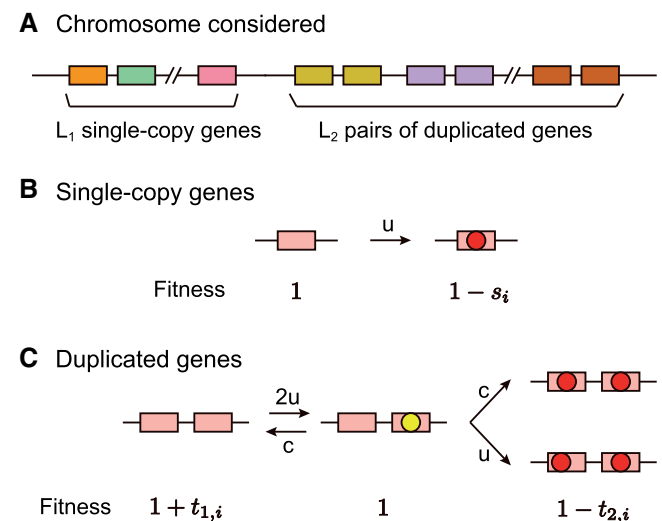


Figure 1 Summary of the model. (A) Chromosome considered in the model. (B, C) Fitness models for a single-copy gene (B) and a pair of duplicated gene (C). Red and yellow circles are, respectively, irreversible and reversible mutations.

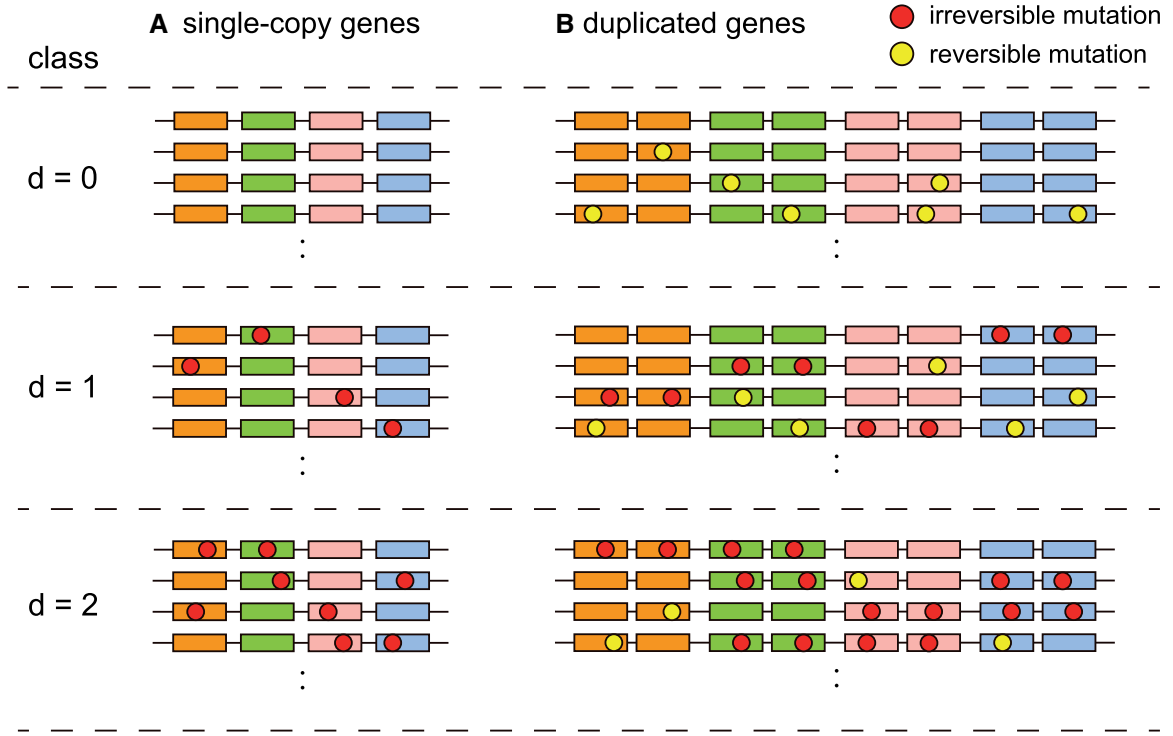


Figure 2 Illustration of the population structure based on d , the number of irreversible mutations for (A) the case of all single-copy genes and (B) all duplicated genes. Boxes in different colors represent genes, or genic regions. Boxes in the same color in (B) indicate duplicated gene pairs. Red and yellow circles are, respectively, irreversible and reversible mutations.

results in an irreversible loss of the gene (see the right state with a red circle in Figure 1B).

For a pair of duplicated genes (Figure 1C), we set the fitness as follows. To be comparable with the case of single-copy genes, the fitness of the state with one functional copy is 1 (middle in Figure 1C). The other copy is inactivated by a loss-of-function mutation, which is shown by a yellow circle in Figure 1C because it is not irreversible, but “reversible”. A reversible mutation can disappear when the intact sequence is transferred from the other functional copy by gene conversion. If gene conversion occurs in the opposite direction, the loss-of-function mutation is transferred to the functional copy, resulting in a state in which the loss-of-function mutation is shared by both copies (the right state with red circles in both copies in Figure 1C). In this state, the gene has completely lost its function, because the mutation became irreversible, since gene conversion cannot rescue the gene function anymore, and the fitness is given by $1 - t_{2,i}$ for the i -th pair of duplicated genes. The fitness of the state with two intact copies (left in Figure 1C) is given by $1 + t_{1,i}$. Having two copies therefore confers a selective advantage by $t_{1,i}$. It is well documented that Y chromosomes have a number of duplicated genes, some of which seem to provide an advantage by increasing the dosage of the gene product (Hughes et al. 2010, 2012, 2020; Soh et al. 2014). Under this model, if we assume $t_{1,i} = 0, t_{2,i} > 0$, the functional allele is in complete dominance, whereas if $t_{1,i} = t_{2,i} > 0$ is assumed, the dominance effect is additive. Loss-of-function mutations arise at a rate of u per copy per generation, and no back mutation is allowed. Gene conversion occurs at a rate of c per copy per generation in both directions between the duplicated copies. We assume that a gene conversion event transfers the entire genic region (represented by a single box in Figure 1C).

It should be noted that two independent mutations can cause a loss of the gene function, as illustrated in the lower case of the right part of Figure 1C. This situation is more complex, because the gene function is lost but the two mutations are still reversible. Nevertheless, we treat this situation as if the gene function is irreversibly lost, which is true in our model, in which a gene conversion event transfers the entire genic region.

The fitness of a haploid individual (chromosome) is determined by the multiplicative effect of all genes. That is,

$$f = \prod_{i=1}^{L_1} w_{1,i} \prod_{i=1}^{L_2} w_{2,i}, \quad (1)$$

where $w_{1,i}$ and $w_{2,i}$ are determined as follows:

$$w_{1,i} = \begin{cases} 1 & \text{(if } i\text{th single-copy gene is functional)} \\ 1 - s_i & \text{(if } i\text{th single-copy gene is lost)} \end{cases},$$

and

$$w_{2,i} = \begin{cases} 1 + t_{1,i} & \text{(if both copies of the } i\text{th duplicated genes are functional)} \\ 1 & \text{(if one copy of the } i\text{th duplicated genes is lost)} \\ 1 - t_{2,i} & \text{(if both copies of the } i\text{th duplicated genes are lost)} \end{cases}.$$

Based on the fitness of all individuals in the current population, the next generation is generated following the Wright-Fisher model.

Simplification for mathematical analyses

Since the general model described above is too complicated for mathematical analysis, we make the following two simplifying

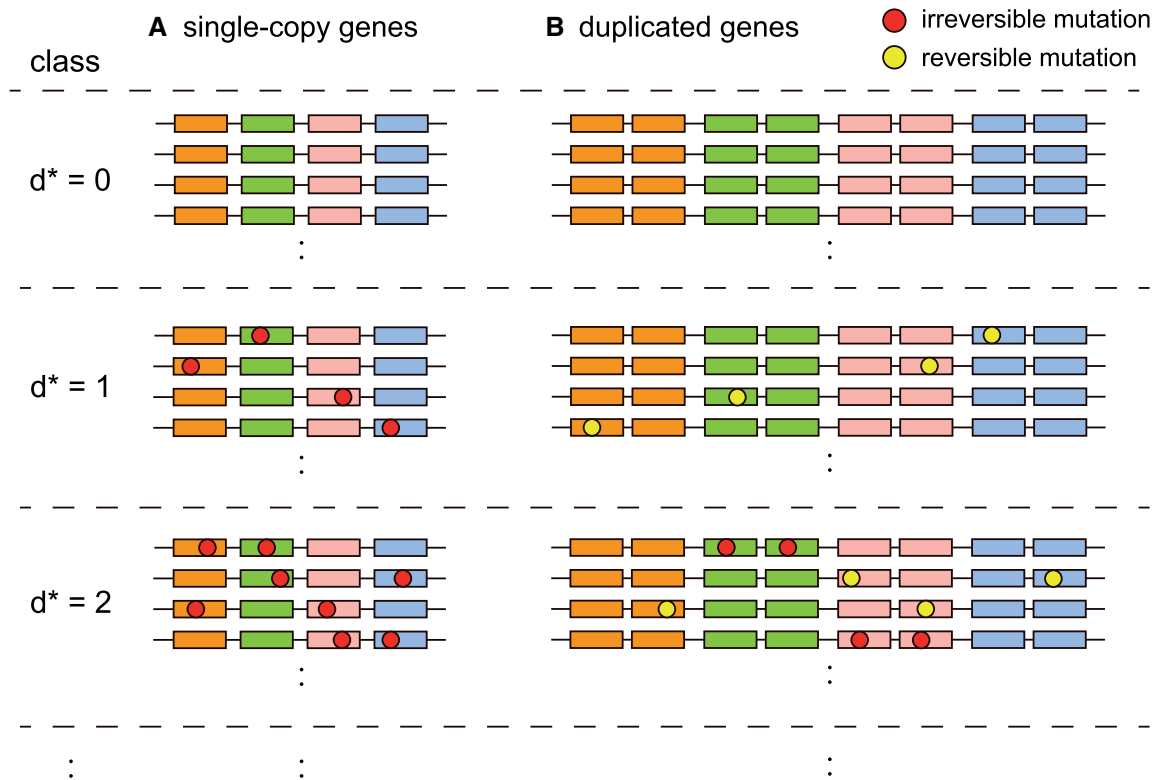


Figure 3 Illustration of the population structure based on d^* , the number of deleterious mutations for (A) the case of all single-copy genes and (B) all duplicated genes. Boxes in different colors represent genes, or genic regions. Boxes in the same color in (B) indicate duplicated gene pairs. Red and yellow circles are, respectively, irreversible and reversible mutations.

assumptions. First, we assume that all genes have the same fitness effect. Therefore, $s_i = s$, $t_{1,i} = t_1$, $t_{2,i} = t_2$ for all i . Second, we assume that the fitness effect of losing the function is the same for a single-copy gene and a pair of duplicated genes ($s = t_2$). Under these assumptions, we derive the rate of gene loss in two special cases: one in which the functional copy is in complete dominance ($t_1 = 0$), and one in which it is additive ($t_1 = t_2$). In the complete dominance case, the fitness of an individual depends on the number of irreversible mutations, d , and the population can be structured based on d (Figure 2). In the additive case, the fitness of an individual depends on the total number of mutations, d^* , and individuals can be classified based on d^* (Figure 3).

Results

We investigated the way in which a nonrecombining chromosome loses functional genes through Muller's ratchet. Our model assumes a chromosome carrying L_1 single-copy genes and L_2 pairs of duplicated genes, and we compared the rates of gene loss for single-copy genes and duplicated genes. We theoretically consider the speed of gene loss per gene, conditional on L_1 and L_2 . The most part of the following is a theoretical analysis, to which the simplifying assumptions detailed above apply. These assumptions are relaxed in the simulation-based analysis, where we mention it.

Duplication with no fitness effect of dosage

We first consider the case in which the presence of two copies, which have arisen by duplication, confers no selective advantage in comparison with having only one copy ($t_1 = 0$). In this case, the fitness of each individual is specified only by the number of

irreversible mutations, d , as defined in Figure 2, and reversible mutations can be ignored. We first derive the equilibrium size of the least-loaded class ($d = 0$), N_e , conditional on L_1 and L_2 , which is the most important quantity with which to determine the evolutionary dynamics of the population. Let U be the total rate of production of irreversible mutations per chromosome (haploid individual). Irreversible mutations arise at rate uL_1 in single-copy genes, while the rate for duplicated genes is $(u + c)H_j$, where H_j is the number of reversible mutations in the j th individual. If we ignore the variance of H_j among individuals, U is given by $U = uL_1 + (u + c)\bar{H}$ where \bar{H} is the average of H_j . We can ignore irreversible mutations segregating in the population, because they exist at low frequencies under the conditions we consider in this work, $u \ll U \sim s$. Following Haigh (1978), N_e is then expressed as

$$N_e = N \exp(-U/s), \quad (2)$$

indicating we need to derive \bar{H} to obtain N_e . It is obvious that \bar{H} largely depends on the frequency of individuals having one reversible mutation in each gene pair, which is denoted by p_i for the i th gene pair. p_i is mainly determined by the frequency dynamics in the least-loaded class, and the deterministic change of p_i per generation in the least-loaded class is given by

$$E[\Delta p_i] = \frac{p_i + 2u(1 - p_i) - (u + 2c)p_i}{1 - (u + c)p_i} - p_i \approx 2u(1 - p_i) - cp_i - (u + c)p_i(1 - p_i), \quad (3)$$

where the term $(u + c)p_i$ in the denominator represents the proportion of individuals that are removed from the least-loaded class by an irreversible mutation, by additional mutation or gene

conversion that occurred in individuals that already had one mutation. Therefore, by using Wright's formula (Wright 1931), the stationary distribution of p_i , $\phi(p_i)$, is given by

$$\phi(p_i) = Ce^{-2N_e(u+c)p_i} p_i^{4N_e u-1} (1-p_i)^{2N_e c-1}, \quad (4)$$

where C is a constant determined such that $\int_0^1 \phi(p_i) dp_i = 1$. Then, considering L_2 pairs of duplicated genes, \bar{H} can be approximately as

$$\begin{aligned} \bar{H} &\approx L_2 \int_0^1 p \phi(p) dp \\ &= L_2 \frac{2u}{2u+c} \frac{{}_1F_1(4N_e u+1; 4N_e u+2N_e c+1; -2N_e(u+c))}{{}_1F_1(4N_e u; 4N_e u+2N_e c; -2N_e(u+c))}, \end{aligned} \quad (5)$$

where ${}_1F_1(\cdot)$ is a confluent hypergeometric function. We now have \bar{H} as a function of N_e . We can therefore obtain N_e as a solution of Equation (2) where \bar{H} in U is replaced by Equation (5).

With N_e from Equation (2), it is straightforward to derive T , the expected waiting time for the first click, conditional on L_1 and L_2 by using the common approach in the theory of Muller's ratchet for single-copy genes (Gordo and Charlesworth 2000a; Jain 2008; Rouzine et al. 2008). When $N_e s > 1$ (i.e., slow ratchet regime), T is approximated by

$$T \approx \frac{e^{\alpha s N_e}}{\alpha s} \sqrt{\frac{\pi}{\alpha s N_e}} \quad (6)$$

(Jain 2008), where $\alpha = \sqrt{s/U}$ (Goyal et al. 2012). When $N_e s < 1$ (i.e., fast ratchet regime), following Rouzine et al. (2008), T is approximately given by

$$T \approx \frac{1}{Uv}, \quad (7)$$

where v is a solution of the following equation:

$$\begin{aligned} \alpha^2 \log(NU\alpha^3) &= \left(1 - \frac{v}{2} \left((1 - \log v)^2 + 1\right)\right) \\ &- \alpha^2 \log \left[\sqrt{\frac{v^3}{1-v}} \frac{1 - \log v}{1 - v(1 - \log v) + 5\alpha^2/6} \right]. \end{aligned}$$

Then, the gene loss rate per gene for a single-copy gene and that for a pair of duplicated genes (R_1 and R_2 , respectively) can be derived as

$$\begin{aligned} R_1 &= \frac{uL_1}{U} \frac{1}{L_1 T} \\ &= \frac{u}{U T} \\ R_2 &= \frac{(u+c)\bar{H}}{U} \frac{1}{L_2 T}, \end{aligned} \quad (8)$$

because these rates are proportional to $1/T$ and the rates of irreversible mutation.

To verify the performance of Equation (8), we carried out forward simulations, and part of the result is shown in Figures 4 and 5. Our simulation assumed a Wright-Fisher population with N haploid individuals with L_1 single-copy genes and L_2 pairs of duplicated genes. Mutation and gene conversion rates and their fitness effects were as described in the Model section. We assumed that $N=10,000$ and $u=1.0 \times 10^{-5}$. Then, in every generation, a random N haploid individuals were generated based on the fitness of the individuals in the previous generation [see Equation

(1)]. The selection parameters were assumed such that $s=0.01$ for a strong selection case, and $s=0.0025$ for a weak selection case. The purpose of this simulation was to obtain \bar{T}_1 and \bar{T}_2 , the average time required for one click of the ratchet at single-copy genes and duplicated genes, respectively, conditional on L_1 and L_2 . From them, the two gene loss rates per gene, R_1 and R_2 , can be computed as $R_1 = \frac{1}{L_1 \bar{T}_1}$ and $R_2 = \frac{1}{L_2 \bar{T}_2}$, respectively. However, if we simply run a simulation, L_1 and L_2 decreases along the run, making it difficult to evaluate T conditional on a specified pair of L_1 and L_2 . To solve this problem, in our simulation, we used an ad hoc method, in which L_1 and L_2 were kept constant by adding an intact gene (or a pair of duplicated genes) when we observed a loss of a gene (or a gene pair). This treatment allowed us to approximately obtain the expectation of T conditional on L_1 and L_2 . See Appendix A, in which we demonstrate that this heuristic treatment worked quite well. In each simulation run, after a burn-in period of 100N generations, we scored the waiting time T for every click for both classes of gene, and the run was terminated when we observed 10,000 clicks or 10,000N generations had passed.

Figure 4 compares the results of two extreme cases: In one, the simulated chromosome consists only of duplicated genes ($L_1 = 0$ and $L_2 = 2,000$), and in the other, only single-copy genes are present in the chromosomes ($L_1 = 2,000$ and $L_2 = 0$). In the case of all duplicated genes, three levels of dominance were considered ($t_1 = 0, t_2/2$, and t_2 presented in blue, green and red, respectively), and the gene conversion rate was changed from 10^{-7} to 10^{-3} . The result for the case of all single-copy genes is shown by the black broken line in Figure 4.

Let us focus on the simulation result of no dominance presented by blue circles ($t_1 = 0$), to which Equation (8) (blue line) is applicable. Equation (8) is in a good agreement with the simulation result, except when the gene conversion rate is very large ($c > 10^{-4}$). In the strong selection case, where the ratchet proceeds so slowly that each click occurs almost independently, R_2 in the case of all duplicates is largely affected by the gene conversion rate (Figure 4A). When the gene conversion rate is very small ($c = 10^{-7}$), R_2 is almost identical to R_1 in the case of all single-copy genes. This is because the ratchet process in duplicated genes is quite similar to that of single-copy genes. The pattern illustrated, with two independent mutations, in Figure 1C, applies to this case. In a pair of duplicated genes, a first mutation itself is neutral and rarely transferred to the other copy, because of a very small c , and only a secondary mutation contributes to the ratchet, behaving as if it occurs in a single-copy gene. R_2 increases as the gene conversion rate increases. This can be explained by the increase of U in Equation (2). Because Equation (5) is approximately given by $\bar{H} \approx \frac{2uL_2}{2u+c}$ when $N_e u, N_e c \ll 1$, U becomes:

$$U \approx u(L_1 + L_2) + \frac{uc}{2u+c} L_2. \quad (9)$$

This equation indicates that, when the mutation and gene conversion rates are not large ($N_e u, N_e c \ll 1$), U for the case of all duplicates is the same as that for the case of all single-copy genes when $c=0$, and R_2 increases as c increases. For a very large c ($N_e c \sim 1$), R_2 decreases with increasing c , because a high c keeps p_i very low and reduces U [see Equation (4)]. At an extremely high gene conversion rate, the fate of a reversible mutation is determined very quickly. In such a case, a reversible mutation arises at a rate of $2u$, and half become irreversible mutations. These mutations then behave like mutations in a single-copy gene that arises at rate u . Therefore, R should be as low as the expectation

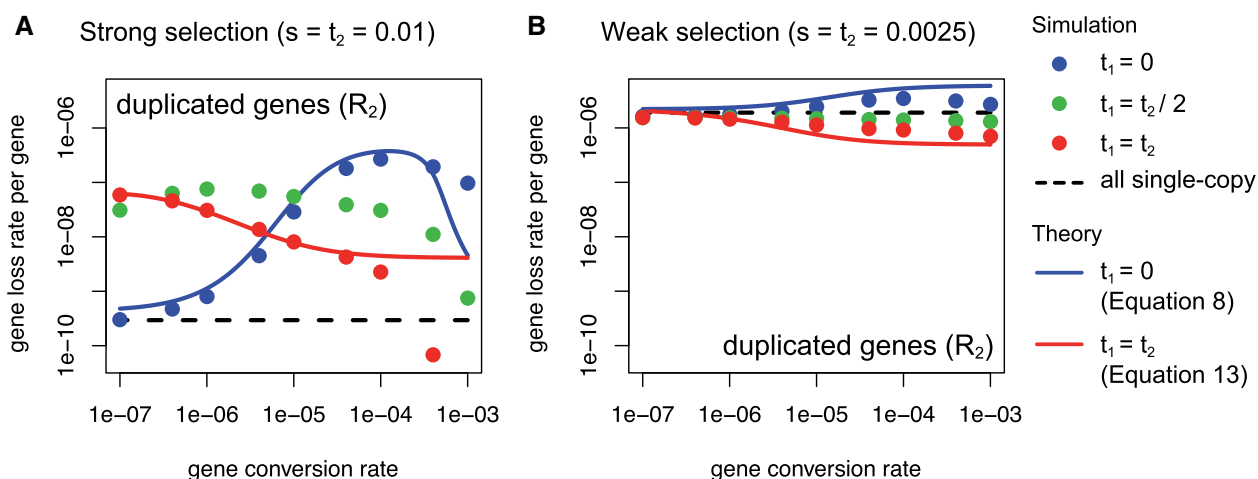


Figure 4 Gene loss rates per duplicated gene pair for (A) the strong selection case and (B) the weak selection case when all genes are single-copy ($L_1 = 2,000, L_2 = 0$) or all genes are duplicated genes ($L_1 = 0, L_2 = 2,000$). The closed circles are the simulation results for the case of all duplicated genes (R_2), while the black dashed lines are the simulation results for the case of all single-copy genes (R_1). The solid lines are the theoretical results, which are available for $t_1 = 0$ (blue) and $t_1 = t_2$ (red). $N = 10,000$, $u = 1.0 \times 10^{-5}$, and $s = t_2$ are assumed.

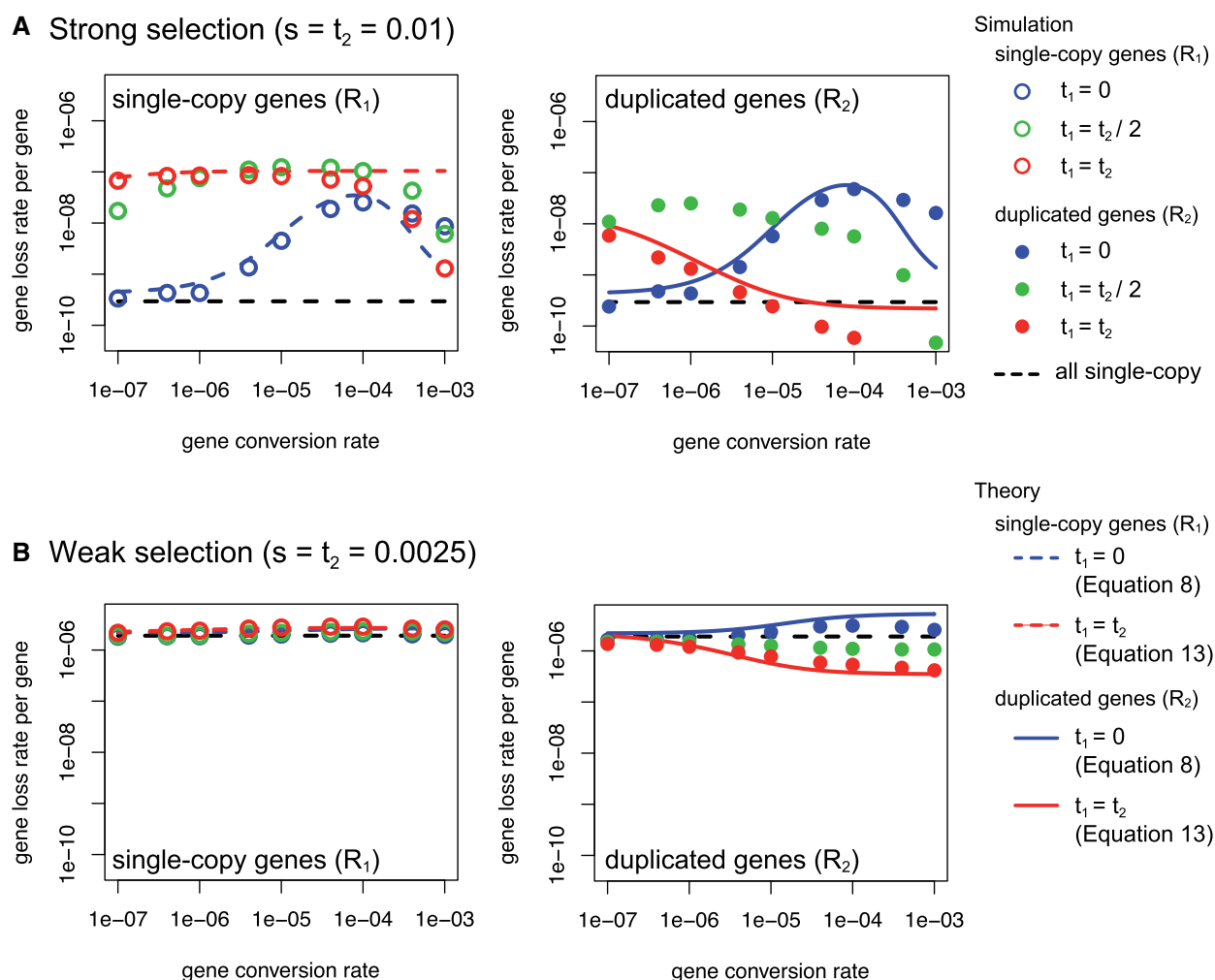


Figure 5 Gene loss rates per gene (or gene pair) for (A) the strong selection case and (B) the weak selection case when single-copy genes and duplicated genes coexist ($L_1 = L_2 = 1,000$). The left panels show the gene loss rate of single-copy genes (R_1) and the right panels show that of duplicated genes (R_2). The open and closed circles are the simulation results, and the colored lines show the theoretical results, which are available for $t_1 = 0$ (blue) and $t_1 = t_2$ (red). The black dashed lines are the simulation results for the case of all single-copy genes. $N = 10,000$, $u = 1.0 \times 10^{-5}$ are assumed.

for a single-copy gene (the dashed line). This explains why R decreases as c becomes very high. Since the size of the least-loaded class, N_e , is reduced significantly as U increases, the gene loss rate is quite sensitive to the change in c in this regime.

Figure 4B shows the result for the weak selection case, in which selection is so weak that multiple clicks occur in a sequential manner with overlapping fixation processes. The qualitative effect of gene conversion on R_2 appears similar to that in Figure 4A: If c is very small, R_2 is almost identical to R_1 in the case of all single-copy genes (black dashed line). With increasing c , R_2 increases up to an intermediate c , and then decreases. The overall quantitative effect of gene conversion is smaller in comparison with Figure 4A, because N_e is always very small in this regime, and the effect of U on R_2 is small.

Figure 5 assumes that single-copy genes and duplicated genes coexist. In this simulation, it was assumed that $L_1 = L_2 = 1,000$, and the other parameters were the same as those used in Figure 4. The gene loss rates are shown by open circles for single-copy genes (R_1) and by filled circles for duplicated genes (R_2). Let us focus on the results of no dominance (blue circles). Equation (8) (blue dashed line in the left panel for R_1 , blue solid line for R_2 in the right panel in Figure 5) agrees well with the simulation results, unless the gene conversion rate is very large.

In the strong selection case (Figure 5A), R_1 and R_2 are very similar to each other. The increase in R_1 and R_2 from that for the case of all single-copy genes (black dashed line) is smaller than that in Figure 4A, which is merely due to the smaller L_2 assumed in Figure 5. R_2 is slightly larger than R_1 because the rate of generation of irreversible mutations is larger for duplicated genes [see Equation (9)].

In the weak selection case (Figure 5B), R_1 is less sensitive to c and almost identical to R_1 for the case of all single-copy genes, while R_2 shows a similar behavior to that for the case of all duplicates in Figure 4B. This is because weak selection causes a small N_e , a situation in which random genetic drift dominates. In such a case, the gene loss rate is roughly proportional to the rate of generation of irreversible mutations, which is constant at u in single-copy genes for any value of c .

We next consider the ratchet process over the long term, where a chromosome gradually loses its functional genes. Let $L_1(\tau)$ and $L_2(\tau)$ be the number of remaining single-copy and duplicated genes at time τ , respectively. Using Equation (8), the differential equations for $L_1(\tau)$ and $L_2(\tau)$ are given by

$$\begin{aligned} \frac{dL_1(\tau)}{d\tau} &= -L_1R_1, \\ \frac{dL_2(\tau)}{d\tau} &= -L_2R_2, \end{aligned} \quad (10)$$

from which we can numerically compute $L_1(\tau)$ and $L_2(\tau)$. To check the performance of Equation (10), we ran forward simulations, and the results are shown in Figure 6. For this purpose, we simply ran the above simulation without keeping $L_1(\tau)$ and $L_2(\tau)$ constant (see above for details), so that the number of genes monotonically decreased during the process. We assumed that each chromosome initially had $L_1(0)$ intact single-copy genes and $L_2(0)$ intact pairs of duplicated genes. It was also assumed that $N = 10,000$, $L_1(0) = L_2(0) = 1,000$, $u = 1.0 \times 10^{-5}$ and $s = t_2 = 0.0025$, and three levels of dominance were used ($t_1 = 0, t_2/2, t_2$). Two different gene conversion rates were considered: $c = 1.0 \times 10^{-4}$ (deep green) and $c = 1.0 \times 10^{-6}$ (orange). In each panel, the result is compared with that for the case of all single-copy genes presented by black circles (i.e., $L_1(0) = 2,000$, and $L_2(0) = 0$), which was obtained by additional simulations.

As we consider the case of no dominance in this section, let us focus on Figure 6A. Equation (10) (solid lines) is in good agreement with the simulation results. When the gene conversion rate is low ($c = 10^{-6}$), the gene loss rates for both single-copy genes and duplicated genes are very similar to those in the case of all single-copy genes, consistent with the results in Figure 5. When the gene conversion rate is high ($c = 10^{-4}$), the gene loss process is slightly accelerated in duplicated genes, consistent with Figure 5B. We also consider the case where s_i and $t_{2,i}$ are heterogeneous among loci in Appendix B and obtained qualitatively similar patterns (see Appendix B for details).

Duplication with an additive fitness effect of dosage

We consider the case of $t_1 = t_2 = s$, in which duplication has an additive effect on fitness. In this case, the fitness of individuals can be specified by the sum of the number of reversible mutations and irreversible mutations, d^* . The treatment for single-copy genes is the same as in the previous section, while some modifications are needed for duplicated genes. Unlike the previous section, a single irreversible mutation in duplicated genes should be counted as two deleterious mutations, because having an irreversible mutation is as deleterious as having two reversible mutations. Based on d^* , the Muller's ratchet process is illustrated in Figure 3. This ratchet is different from that in the previous section (see Figure 2) because a ratchet click can occur in either direction. Let us consider the situation where the least-loaded class has $d^* = d_0$ mutations. A ratchet click occurs in the forward direction when the class $d^* = d_0$ goes extinct, as in the previous section. We also need to consider a click in the backward direction, when an individual with $d^* = d_0 - 1$ mutations arises by gene conversion, and its descendants become the majority of the population, thereby constituting the new least-loaded class, with $d^* = d_0 - 1$. Clicks in both directions simply change the number of reversible mutations, and only a part of the forward clicks can cause gene loss. The analytical approach we use here is quite different from that in the previous case, but is similar to that of Goyal et al. (2012), who incorporated back mutations into the model of Muller's ratchet for single-copy genes. Following Goyal et al. (2012), we consider the click process in each direction separately.

We first consider clicks in the forward direction, which increases d^* . We derive the expected waiting time until the next click in the forward direction, T_F , from which the rate of gene loss will be derived. Let \bar{H} be the average number of reversible mutations per individual, and denote the total rate of production of deleterious mutations per chromosome as U^* . We need to incorporate the following three types of deleterious mutation to derive U^* : irreversible mutations that arise in single-copy genes at the rate uL_1 , irreversible mutations that arise in duplicated genes at the rate $(u + c)\bar{H}$, and reversible mutations that arise in duplicated genes at the rate $2u(L_2 - \bar{H})$. U^* is then given by

$$U^* = uL_1 + (u + c)\bar{H} + 2u(L_2 - \bar{H}) \quad (11)$$

as a sum of the three types, and the size of the least-loaded class is

$$N_e^* = N \exp(-U^*/s). \quad (12)$$

From N_e^* , T_F can be obtained in a similar manner to the previous section. When $N_e^*s > 1$ (i.e., the slow ratchet regime), T_F can be obtained from Equation (6) by substituting U with U^* and T with

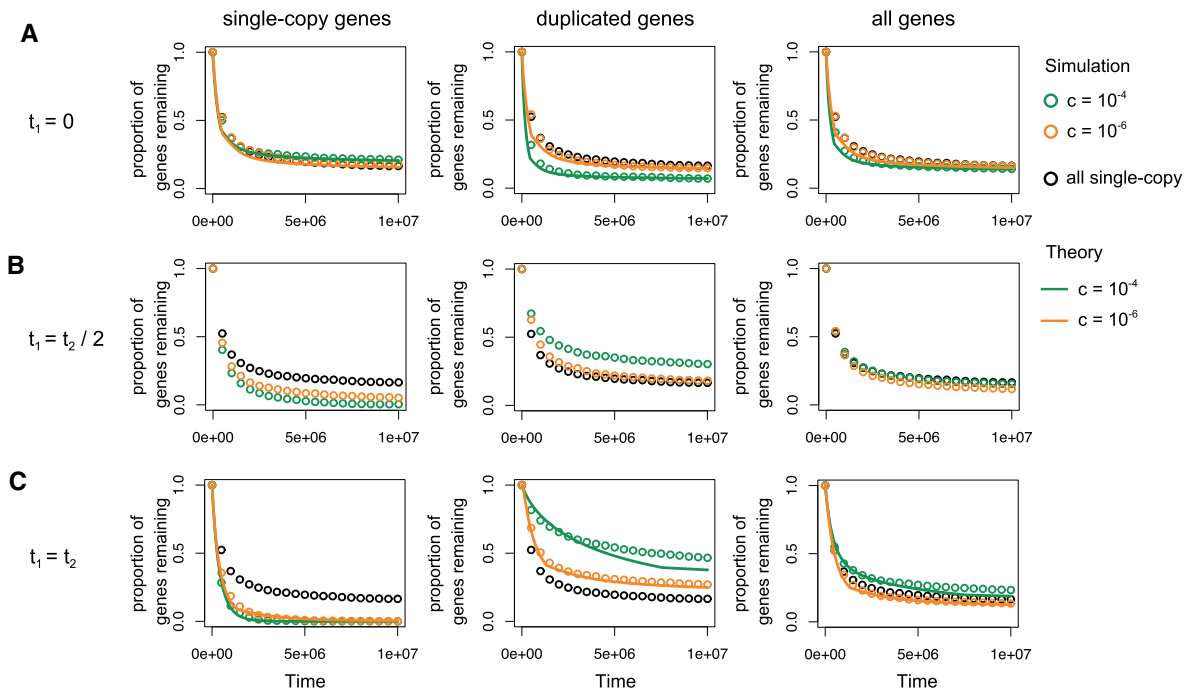


Figure 6 Long-term degeneration process with different dominance: (A) $t_1 = 0$, (B) $t_1 = t_2/2$, and (C) $t_1 = t_2$ are assumed. The open circles represent simulation results. The solid lines are theoretical predictions, which are available for (A) and (C). Other parameters are $N = 10,000$, $u = 1.0 \times 10^{-5}$, $s = t_2 = 0.0025$, and $L_1 = L_2 = 1,000$.

T_F . When $N_e^*s < 1$ (i.e., the fast ratchet regime), T_F is given by Equation (7) by substituting U with U^* and T with T_F . Therefore, the gene loss rate per gene for a single-copy gene and that for a pair of duplicated genes (R_1 and R_2 , respectively) can be derived as

$$\begin{aligned} R_1 &= \frac{uL_1}{U^*} \frac{1}{L_1 T_F} \\ &= \frac{u}{U^*} \frac{1}{T_F}, \\ R_2 &= \frac{(u+c)\bar{H}}{U^*} \frac{1}{L_2 T_F}, \end{aligned} \quad (13)$$

because these rates are proportional to $1/T_F$ and the rates of mutation.

It should be noted that R_1 and R_2 are functions of \bar{H} , which can be obtained as follows. The dynamics of \bar{H} involve both forward and backward processes. That is, if we let ΔH_F and ΔH_B be the expected changes in \bar{H} per generation in the forward and backward processes, respectively, the expected change in \bar{H} per generation is given by

$$\frac{d\bar{H}}{d\tau} = \Delta H_F + \Delta H_B. \quad (14)$$

We here explain how \bar{H} behaves when a forward click occurs. \bar{H} increases by one when a forward click occurs by the fixation of a reversible mutation in duplicated genes, and decreases by one when it occurs by the fixation of an irreversible mutation in duplicated genes. Because these events occur at rates $\frac{2u(L_2 - \bar{H})}{U^* T_F}$ and $L_2 R_2$, respectively, ΔH_F is given by

$$\Delta H_F = \frac{2u(L_2 - \bar{H})}{U^* T_F} - L_2 R_2. \quad (15)$$

We will below obtain ΔH_B in the following derivation for the backward process.

We next consider clicks in the backward direction, where \bar{H} decreases by one. We first derive the expected waiting time until the next click in the backward direction, T_B . A backward click usually occurs when a reversible mutation is removed by gene conversion in an individual in the least-loaded class and its descendants become the majority of the population, resulting in a new least-loaded class. When $N_e^*s > 1$, the treatment of the new least-loaded class is relatively simple because its fate is determined quickly. Suppose a new least-loaded class with $d^* - 1$ arises by gene conversion in the population where the least-loaded class is d^* . Because of strong selection, the new least-loaded class spreads in the population quickly with probability $\approx 2s$ (Haldane 1927), otherwise becomes extinct. Therefore, T_B is given by

$$T_B \approx (2sc\bar{H}N_e^*)^{-1}, \quad (16)$$

because gene conversion creates a new least-loaded class at rate $c\bar{H}N_e^*$ per generation.

When $N_e^*s < 1$, the treatment of the new least-loaded class is not straightforward. Because selection is weak, the frequency of the newly arisen least-loaded class with $d^* - 1$ fluctuates because of genetic drift. The newly arisen least-loaded class could be maintained in the population for a while and become extinct by genetic drift, or it could increase in frequency to some extent when a backward click occurs (but it does not necessarily mean that the new least-loaded class becomes the majority). It is technically difficult to distinguish the two situations. Rouzine et al. (2008) proposed that the least-loaded class can be considered to be lost when its frequency is smaller than $\frac{1}{5^*N}$, where $5^* = U^*v(1 - \log v)$. Following this proposition, we assume that the frequency of the new least-loaded class is significantly increased (i.e., it is considered as a backward click) when the frequency reaches $\frac{1}{5^*N}$, an event which occurs with probability 5^* . Then, because the expected size of the least-loaded class is

approximately given by $N_e^* \approx \frac{\exp(S^*T_F)-1}{S^*T_F}$ (see Rouzine et al. 2008), T_B is approximately given by

$$T_B \approx \left(\frac{\exp(S^*T_F)-1}{S^*T_F} c\bar{H} \right)^{-1} = \frac{1 - \log v}{(e/v - 1)c\bar{H}}, \quad (17)$$

where e is Euler's number. From T_B , it is straightforward to obtain ΔH_B . Because a backward click always reduces \bar{H} by one, which occurs at rate $1/T_B$, we have

$$\Delta H_B = -1/T_B. \quad (18)$$

Together with Equations (14) and (15), we are now ready to compute $\frac{d\bar{H}}{d\tau}$. This treatment is general, in that we can obtain the temporal change in \bar{H} from any initial condition.

To check the performance of Equation (13), we compare it with simulation results in Figures 4 and 5 (see above for details about the simulations). Although we can compute R_1 and R_2 using Equation (13) for any initial values of \bar{H} , we use a treatment with no initial conditions specified, because we are interested in R_1 and R_2 conditional on L_1 and L_2 in a steady state, in which the initial condition is relatively unimportant. To obtain R_1 and R_2 conditional on L_1 and L_2 , we assume that the frequency of reversible mutation is in equilibrium. Then, \bar{H} is determined from Equations (13) and (14) such that $\frac{d\bar{H}}{d\tau} = \frac{1}{L_2} \frac{d\bar{H}}{d\tau} + \bar{H} R_2 = 0$. Using this \bar{H} , we derived R_1 and R_2 by Equation (13).

Figure 4 shows the results of two extreme cases: In one case, the chromosome consists only of duplicated genes ($L_1 = 0, L_2 = 2,000$), and in the other case, it consists only of single-copy genes ($L_1 = 2,000, L_2 = 0$). In the case of all duplicates, since Equation (13) is applicable to the case of $t_1 = t_2$, let us focus on the result represented by the red circles. The result for the case of all single-copy genes is represented by black broken lines. Equation (13) agrees well with the simulation results, unless the gene conversion rate is very high.

In the strong selection case (Figure 4A), R_2 is strongly affected by the gene conversion rate. When the gene conversion rate is very low ($c = 10^{-7}$), R_2 is much higher than R_1 in the case of all single-copy genes, because U^* is elevated due to the increase in copy number caused by duplication [see Equation (11)]. As the gene conversion rate increases, R_2 decreases, and then drops dramatically when $c \sim 10^{-4}$, producing a bad fit between Equation (13) and the simulation result. This situation arises because reversible mutations at duplicated genes are quickly removed by gene conversion in this regime. Individuals with a reversible mutation in the second most loaded class are quickly transferred into the least-loaded class, which increases the size of the least-loaded class and retards its extinction. In this situation, d^* in the population does not follow a Poisson distribution, explaining why the simulation result is not well explained by our derivation [Equation (12)], in which a Poisson distribution is used for the distribution of d^* .

In the weak selection case (Figure 4B), R_2 is relatively robust to the gene conversion rate. In this regime, N_e^* is small enough for random genetic drift to dominate, and R_2 is roughly proportional to the rate of generation of irreversible mutations per gene, $(u + c)\bar{H}/L_2$ [see Equation (13)]. When the gene conversion rate is very low ($c = 10^{-7}$), R_2 is almost identical to R_1 in the case of all single-copy genes (Figure 4B). This phenomenon can be explained by considering the fate of a newly arisen mutation, as shown in Figure 1C, in which gene conversion should be ignored. If a

mutation arises in one copy, given a small N_e^* , the mutation can become fixed in one copy with a specific probability. Once it is fixed, as our model does not allow back mutation, the state with one reversible mutation (middle in Figure 1C) is prolonged, because this mutation cannot be removed (if gene conversion is ignored). The next event that could happen is that an independent mutation fixes in the other copy, causing a loss of the duplicated genes (the lower case in the right part in Figure 1C). Thus, if we consider a chromosome with L_2 pairs of duplicates in a steady state, it is likely that most pairs would be in this state ($\bar{H} \sim L_2$), because genes with their functions already lost are out of the system. Give this situation, the rate at which another mutation causing a gene loss is generated is approximately u per gene, which is identical to that of the all single-copy case, explaining the similar gene loss rates in the two cases. As the gene conversion rate increases, R_2 decreases because gene conversion removes reversible mutations to some extent, resulting in $\bar{H} \ll L_2$.

In Figure 5, we consider a chromosome in which single-copy genes and duplicated genes coexist. It was assumed that $L_1 = L_2 = 1,000$ and other parameters were the same as those used for Figure 4. In the additive selection case (red circles), Equation (13) is in a good agreement with the simulation results unless the gene conversion rate is very large.

In the strong selection case, R_2 exhibits behavior similar to that shown in Figure 4A, but R_1 is quite different. When the gene conversion rate is very low ($c = 10^{-7}$), R_1 is much higher than that of the all single-copy case, because U^* elevated by gene duplication reduces N_e^* , so that the ratchet clicks in the forward direction occur more frequently [see Equations (11)–(13)]. Unless the gene conversion rate is very high ($c < 10^{-4}$), R_1 is quite robust to c , because gene conversion in duplicated genes should not have a direct effect on mutations in single-copy genes. When the gene conversion rate is very high ($c > 10^{-4}$), R_1 starts decreasing with increasing c . The fit of Equation (13) to the simulation result is not good for the same reason as that in Figure 4A.

When selection is weak, R_2 is very similar to that in Figure 4B. R_1 is less affected by the gene conversion rate, because N_e is small enough for random genetic drift to dominate, and the gene loss rate is roughly proportional to the rate of generation of irreversible mutations, which is constant for u in single-copy genes.

Next, we focus on the long-term degeneration process. $L_1(\tau)$ and $L_2(\tau)$ are the numbers of remaining single-copy genes and duplicated gene pairs, respectively, and $\bar{H}(\tau)$ is the average number of reversible mutations per individual at time τ . By using Equations (13) and (14), the differential equations for $L_1(\tau)$, $L_2(\tau)$ and $\bar{H}(\tau)$ are given by

$$\begin{aligned} \frac{dL_1(\tau)}{d\tau} &= -L_1 R_1, \\ \frac{dL_2(\tau)}{d\tau} &= -L_2 R_2, \\ \frac{d\bar{H}(\tau)}{d\tau} &= \Delta H_F + \Delta H_B, \end{aligned} \quad (19)$$

from which we can compute the way in which $L_1(\tau)$, $L_2(\tau)$, and $\bar{H}(\tau)$ change over time.

To check the accuracy of Equation (19), we performed simulations (Figure 6C). In the initial state of the simulation, each chromosome had 1,000 intact single-copy genes and 1,000 intact duplicated pairs of genes ($L_1(0) = L_2(0) = 1,000$ and $\bar{H}(0) = 0$). Two gene conversion rates ($c = 10^{-4}, 10^{-6}$) were considered. It was found that, at both of the two gene conversion rates,

single-copy genes decreased faster and duplicated genes decreased more slowly than in the case of all single-copy genes (black circles in Figure 6C), because the rapid decrease in the number of single-copy genes slows ratchet clicks in the forward direction. As a consequence, more duplicated genes remain functional than in the case of all single-copy genes. The deviation from the case of all single-copy genes is larger when the gene conversion rate is higher. A very similar result was obtained when the assumption of constant selection coefficient was violated. See Appendix B for details.

Duplication with an intermediate fitness effect of dosage

Finally, we consider the case of an intermediate degree of dosage effect, where $t_1 = t_2/2$ is assumed. Since it was difficult to obtain analytical results, we investigated this case using simulations. The green circles in Figure 4 show R_2 , the gene loss rate in the case of all duplicated genes ($L_2 = 2,000$). In the strong selection case, R_2 generally decreases as c increases, whereas in the weak selection case, R_2 is almost identical to R_1 in the case of all single-copy genes. When both single-copy genes and duplicated genes coexist ($L_1 = L_2 = 1,000$, in green in Figure 5), the pattern is generally similar to that in Figure 4. In the strong selection case, R_1 and R_2 show a similar pattern to the additive case. Figure 6B shows the long-term degeneration pattern. In Appendix B, we consider the case where s_i and $t_{2,i}$ are heterogeneous among loci. Overall, the behavior when $t_1 = t_2/2$ seems to be intermediate between those of the cases of no fitness effect of dosage ($t_1 = 0$) and additive effect ($t_1 = t_2$).

Discussion

Muller's ratchet is a process in which a nonrecombining chromosome irreversibly accumulates deleterious mutations (Muller 1964). Muller's ratchet has been considered to play an important role in the evolution of Y chromosome (Bachtrog 2008). Previous theories regarding Muller's ratchet considered only single-copy genes, and the way in which Muller's ratchet works in duplicated genes has not been fully understood. Because there are a number of duplicated genes on the Y chromosome in many species (Skaletsky et al. 2003; Hughes et al. 2010, 2012, 2020; Soh et al. 2014; Bachtrog et al. 2019; Peichel et al. 2020), in this work we developed a theory for the process of Muller's ratchet on a nonrecombining chromosome in which single-copy and duplicated genes coexist. Mutations in duplicated genes can be considered to be a kind of epistatic mutations, because the strength of selection on a mutation depends on whether the other copy already has a mutation or not. Several studies have investigated the effect of epistasis on the process of Muller's ratchet (Charlesworth et al. 1993; Kondrashov 1994; Butcher 1995; Jain 2008). However, these studies have focused on more complex epistatic interactions, in which the fitness effect of a mutation depends upon mutations at all other loci.

This work focuses on the role of gene conversion between duplicates, which has two opposite effects on the degeneration process. Degeneration is promoted if gene conversion leads to the mutation of both copies, while degeneration is retarded if gene conversion restores both copies to the original state. Our theoretical results demonstrate that the effect of gene conversion is complex, depending on the fitness effect of dosage change by gene duplication. When duplication has no fitness effect by dosage, gene conversion increases the rates of loss of both single-copy and duplicated genes (see Figure 5). In the case of an additive

dosage effect on fitness, the gene loss rate of single-copy genes is elevated by gene duplication, while gene conversion prevents duplicated genes from losing their functions (see Figure 5). These patterns are more clearly observed when selection is strong.

This complex nature of the effect of Muller's ratchet on the Y chromosome has not previously been identified. Most of the previous studies have considered only the positive effects of gene conversion. For example, Connallon and Clark (2010) showed that gene conversion works positively in the conservation of essential duplicated genes, assuming mutations are lethal when present in both copies. Our work demonstrates that this is not the case when duplicated genes are not essential (see the Introduction for the difference in the models). Mano and Innan (2008) considered a "single" pair of duplicated gene in a Wright-Fisher population, so the effect of linked genes was ignored, and demonstrated that deleterious mutations could be efficiently removed by gene conversion, which suggests that gene conversion could slow the degeneration process of duplicated genes [similar to Connallon and Clark (2010)]. However, we found that, when multiple genes are completely linked and multiple mutations exist simultaneously, gene duplication can accelerate the degeneration process in some cases.

Our results suggest that understanding the way in which the Y chromosome evolves requires the consideration of a number of parameters, including the number of duplicated genes and their fitness effect through dosage increase. Unfortunately, there are very few empirical data that allow us to estimate those parameters. Theory suggests that dosage increase of many duplicated genes may be beneficial because duplication is more likely to be fixed in the population when it has a direct selective advantage (Clark 1994; Connallon and Clark 2010; Innan and Kondrashov 2010). If so, our theory predicts that duplicated genes are well-conserved by gene conversion, while linked single-copy genes are lost rapidly. However, the situation should be much more complicated, particularly in the early stages of the evolution of the Y chromosome, when duplication-rich Y chromosomes may be developed. Duplicated genes with no dosage effect on fitness may be fixed with a linked gene which has a selective advantage due to a dosage effect. More data with reliable estimates of those selection parameters will give us deeper insights into the evolution of the Y chromosome.

Another important parameter is the gene conversion rate, which has been relatively well estimated in the palindrome regions of the human Y chromosome. Rozen et al. (2003) estimated the gene conversion rate of human palindrome as $2c = 2.2 \times 10^{-4}$ per nucleotide per generation, based on the amount of divergence between duplicates. Hallast et al. (2013) used a phylogenetic approach, and reported a similar but slightly smaller value ($2c = 2.9 - 8.4 \times 10^{-5}$). At such a high gene conversion rate, our theory predicts that gene conversion plays a significant role in the degeneration of the Y chromosome. Although the gene conversion rate in other species has not been as well investigated, high sequence identity between duplicated copies is observed in many species (Soh et al. 2014; Hughes et al. 2020), suggesting that our theory would apply to a wide range of species.

Data availability

The authors state that all data necessary for confirming the conclusions presented in the manuscript are represented fully within the manuscript. Codes used for numerical analyses and simulations are available at <https://github.com/TSakamoto-evo/Y-ratchet>.

Supplementary material is available at GENETICS online.

Funding

This work was partly supported by grants from the Graduate University for Advanced Studies, SOKENDAI, and Japan Society for the Promotion of Science (JSPS) to H.I. and T.S. (JSPS KAKENHI Grant Numbers JP19H03207, JP20J21760).

Conflicts of interest

The authors declare that there is no conflict of interest.

Literature cited

- Arnheim N. 1983. Concerted evolution of multigene families. In: M. Nei, RK, Koehn, editors. *Evolution of Genes and Proteins*. Sunderland, MA: Sinauer. p. 38–61.
- Haigh J. 1978. The accumulation of deleterious genes in a population—Muller's ratchet. *Theor Popul Biol.* 14:251–267.
- Muller HJ. 1964. The relation of recombination to mutational advance. *Mutat Res.* 1:2–9.
- Bachtrog D, Mahajan S, Bracewell R. 2019. Massive gene amplification on a recently formed *Drosophila* Y chromosome. *Nat Ecol Evol.* 3:1587–1597.
- Bachtrog D. 2008. The temporal dynamics of processes underlying Y chromosome degeneration. *Genetics.* 179:1513–1525.
- Bachtrog D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat Rev Genet.* 14: 113–124.
- Butcher D. 1995. Muller's ratchet, epistasis and mutation effects. *Genetics.* 141:431–437.
- Charlesworth B, Charlesworth D. 1997. Rapid fixation of deleterious alleles can be caused by Muller's ratchet. *Genet Res.* 70:63–73.
- Charlesworth B, Charlesworth D. 2000. The degeneration of Y chromosomes. *Philos Trans R Soc Lond B Biol Sci.* 355:1563–1572.
- Charlesworth D, Morgan M, Charlesworth B. 1993. Mutation accumulation in finite outbreeding and inbreeding populations. *Genet Res.* 61:39–56.
- Clark AG. 1994. Invasion and maintenance of a gene duplication. *Proc Natl Acad Sci U S A.* 91:2950–2954.
- Connallon T, Clark AG. 2010. Gene duplication, gene conversion and the evolution of the Y chromosome. *Genetics.* 186:277–286.
- Gessler DD. 1995. The constraints of finite size in asexual populations and the rate of the ratchet. *Genet Res.* 66:241–253.
- Gordo I, Charlesworth B. 2000a. The degeneration of asexual haploid populations and the speed of Muller's ratchet. *Genetics.* 154: 1379–1387.
- Gordo I, Charlesworth B. 2000b. On the speed of muller's ratchet. *Genetics.* 156:2137–2140.
- Goyal S, Balick DJ, Jerison ER, Neher RA, Shraiman BI, et al. 2012. Dynamic mutation–selection balance as an evolutionary attractor. *Genetics.* 191:1309–1319.
- Graves JAM. 2004. The degenerate Y chromosome—can conversion save it? *Reprod Fertil Dev.* 16:527–534.
- Haldane JBS. 1927. A mathematical theory of natural and artificial selection, part V: selection and mutation. *Math Proc Camb Phil Soc.* 23:838–844.
- Hallast P, Balaesque P, Bowden GR, Ballereau S, Jobling MA. 2013. Recombination dynamics of a human Y-chromosomal palindrome: rapid GC-biased gene conversion, multi-kilobase conversion tracts, and rare inversions. *PLoS Genet.* 9:e1003666.
- Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Graves T, et al. 2012. Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes. *Nature.* 483:82–86.
- Hughes JF, Skaletsky H, Pyntikova T, Graves TA, Van Daalen SK, et al. 2010. Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content. *Nature.* 463:536–539.
- Hughes JF, Skaletsky H, Pyntikova T, Koutseva N, Raudsepp T, et al. 2020. Sequence analysis in *Bos taurus* reveals pervasiveness of X–Y arms races in mammalian lineages. *Genome Res.* 30: 1716–1726.
- Innan H, Kondrashov F. 2010. The evolution of gene duplications: classifying and distinguishing between models. *Nat Rev Genet.* 11:97–108.
- Innan H. 2002. A method for estimating the mutation, gene conversion and recombination parameters in small multigene families. *Genetics.* 161:865–872.
- Innan H. 2003. The coalescent and infinite-site model of a small multigene family. *Genetics.* 163:803–810.
- Jain K. 2008. Loss of least-loaded class in asexual populations due to drift and epistasis. *Genetics.* 179:2125–2134.
- Kondrashov AS. 1994. Muller's ratchet under epistatic selection. *Genetics.* 136:1469–1473.
- Mano S, Innan H. 2008. The evolutionary rate of duplicated genes under concerted evolution. *Genetics.* 180:493–505.
- Marais GA, Campos PR, Gordo I. 2010. Can intra-Y gene conversion oppose the degeneration of the human Y chromosome? A simulation study. *Genome Biol Evol.* 2:347–357.
- Metzger JJ, Eule S. 2013. Distribution of the fittest individuals and the rate of Muller's ratchet in a model with overlapping generations. *PLoS Comput Biol.* 9:e1003303.
- Neher RA, Shraiman BI. 2012. Fluctuations of fitness distributions and the rate of Muller's ratchet. *Genetics.* 191:1283–1293.
- Ohta T. 1983. On the evolution of multigene families. *Theor Popul Biol.* 23:216–240.
- Peichel CL, McCann SR, Ross JA, Naftaly AF, Urton JR, et al. 2020. Assembly of the threespine stickleback Y chromosome reveals convergent signatures of sex chromosome evolution. *Genome Biol.* 21:31.
- Rouzine IM, Brunet É, Wilke CO. 2008. The traveling-wave approach to asexual evolution: Muller's ratchet and speed of adaptation. *Theor Popul Biol.* 73:24–46.
- Rozen S, Skaletsky H, Marszalek JD, Minx PJ, Cordum HS, et al. 2003. Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature.* 423:873–876.
- Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, et al. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature.* 423:825–837.
- Skov L, Schierup MH, Consortium DPG, et al.; Danish Pan Genome Consortium. 2017. Analysis of 62 hybrid assembled human Y chromosomes exposes rapid structural changes and high rates of gene conversion. *PLoS Genet.* 13:e1006834.
- Söderberg RJ, Berg OG. 2007. Mutational interference and the progression of Muller's ratchet when mutations have a broad range of deleterious effects. *Genetics.* 177:971–986.
- Soh YS, Alföldi J, Pyntikova T, Brown LG, Graves T, et al. 2014. Sequencing the mouse Y chromosome reveals convergent gene acquisition and amplification on both sex chromosomes. *Cell.* 159:800–813.
- Stephan W, Chao L, Smale JG. 1993. The advance of Muller's ratchet in a haploid asexual population: approximate solutions based on diffusion theory. *Genet Res.* 61:225–231.
- Teshima KM, Innan H. 2004. The effect of gene conversion on the divergence between duplicated genes. *Genetics.* 166:1553–1560.

- Waxman D, Loewe L. 2010. A stochastic model for a single click of Muller's ratchet. *J Theor Biol.* 264:1120–1132.
- Wright S. 1931. Evolution in mendelian populations. *Genetics.* 16: 97–159.

Communicating editor K. Jain

Appendices

Appendix A: On the ad hoc treatment in the simulation to keep L_1 and L_2 constant

In the simulation to obtain R_1 and R_2 conditional on L_1 and L_2 in a steady state, we used an ad hoc treatment in which an intact single-copy gene, or an intact pair of duplicated genes, suddenly appears in all individuals when an irreversible mutation is fixed in the population. The problem is that this method skips a burn-in period, in which some mutations could be accumulated. The performance of this ad hoc method was verified using the following simulation. Let us assume that we would like to obtain R_1 and R_2 conditional on L_1^* and L_2^* . Then, we need to consider a degeneration process of a chromosome on which $L_1^0 = L_1^* + \Delta L_1$ single-copy genes and $L_2^0 = L_2^* + \Delta L_2$ duplicated genes are initially located, where $\Delta L_1, \Delta L_2$ should be sufficiently large. Then, the system waits for the state with the focal pair $(L_1, L_2) = (L_1^*, L_2^*)$, and we can continue the simulation run to obtain the waiting time for the next click. Very few simulation runs hit $(L_1, L_2) = (L_1^*, L_2^*)$, and all other runs are terminated when $L_1 < L_1^*$ or $L_2 < L_2^*$. This method is honest and correct, but requires a very large number of runs to accumulate a reasonable number of simulation runs that hit $(L_1, L_2) = (L_1^*, L_2^*)$. This is why we used the ad hoc

treatment in the main text. We here show how the ad hoc treatment works in comparison with the correct simulation with a limited set of parameters.

We assumed that $N=500$, $u = 2.0 \times 10^{-4}$ and $L_1 = L_2 = 1,000$. The gene conversion rate changed from 2.0×10^{-6} to 2.0×10^{-2} . In [Supplementary Figure S1A](#) we assumed strong selection ($s = t_2 = 0.2$), while in [Supplementary Figure S1B](#), weak selection ($s = t_2 = 0.05$) was assumed. All population scaled parameters (Nu , Nc , Ns , Nt_1 , and Nt_2) are identical to those in [Figure 5](#) if scaled by the population size N . This process demonstrates that the results of the two methods are almost identical, indicating that our ad hoc method works well.

Appendix B: Variable selection coefficients across loci

In this Appendix, we relax the assumption that all genes have identical effects on fitness, and explore the effect of varying fitness selection coefficients across genes on the long-term degeneration process. Three types of dosage effect on duplication are considered: $t_{1,i} = 0$, $t_{1,i} = t_{2,i}/2$, and $t_{1,i} = t_{2,i}$. We performed simulations for these three types of dosage effect, with variation allowed between individual gene pairs. We randomly chose s_i and $t_{2,i}$, assuming they follow an exponential distribution with mean \bar{s} , and they were shared by the simulations for the three types of dosage effect to focus on the effect due the dosage type alone. We simulated the case of all single-copy genes, in which we also used the parameters determined above ($s_i = t_{2,i-L_1}$ is assumed for $L_1 + 1 \leq i \leq L_1 + L_2$). All other parameters are the same as those used in the main text. [Supplementary Figure S2](#) shows the simulation result, which is qualitatively very similar to [Figure 6](#).